Article

Machine Learning Method for Assessing Fire Risk in Wildland-Rural and Urban Interface Regions: The Example of Istanbul's European Side

Ercüment AKSOY [1], Abdulkadir KOCER [2], İsmail YILMAZ [3], Arif Nihat AKÇAL [4] and Kudret AKPINAR [5],*

[1] Department of Geographic Information Systems, Akdeniz University, 07070 Antalya, Türkiye; ercumentaksoy@akdeniz.edu.tr

[2] Department of Electricity and Energy, Akdeniz University, 07070 Antalya, Türkiye; akocer@akdeniz.edu.tr

[3] Remote Sensing and Geographic Information Systems Graduate Program, Institute of Natural and Applied Sciences, Akdeniz University, 07070 Antalya, Türkiye; isoylmz@gmail.com

[4] Department of Construction, Akdeniz University, 07070 Antalya, Türkiye; arifnihatakcal@akdeniz.edu.tr

[5] Department of Civil Defense and Firefighting, Akdeniz University, 07070 Antalya, Türkiye

* Correspondence: akudret@akdeniz.edu.tr; Tel.:+90 242 3106741, +90 505 5881283

Abstract: As in many parts of the world, rural-urban, forest-urban interface areas surrounding urban regions expose the natural areas they interact with to a threat of fire risk that can reach various sizes. This risk has been assessed for various regions of the world using many different methods and numerical models so far. Among these, it is seen that machine learning models have successful applications in risk assessment and risk prediction studies. For the fire risk prediction of Istanbul's yet unurbanized regions, but where the city is anticipated to potentially shift, data was collected using the opportunities provided by Geographic Information Systems and Remote Sensing technologies based on fires that occurred between 2000-2021, and the region was examined. Machine learning methods' Random Forest (RF), Extreme Gradient Boosting (XGB), and Light Gradient Boosting (LGB) models were applied for the classification of factors effective in fire. The best result was given by the RF model with 0.93 accuracy, 0.062 F1 score, and 0.753 Area Under Curve (AUC) value. In the classification of factors in the RF model, the grouping between fire-initiating factors and factors effective in spreading is evident, while this distinction is partially noticeable in the other two models.

Keywords: fire risk, wildland-urban interface, rural-urban interface, machine learning classification, GIS.

1. Introduction

The primary causes of fires are mostly human-induced. In fires occurring in rural areas and forests, while natural reasons for the ignition and spread are extremely few, direct or indirect human factors are much more prevalent [1,2]. In this case, rural and forested areas and areas intertwined or intermixed with human settlements are under the increasing influence of human activities. If human settlements are in close proximity to natural and human-derived biomass, the main cause of forest, rural, and biomass fires is largely human. It has been observed that there has been an increase of approximately 12% in fires caused by lightning strikes due to global warming in the United States [3]. Even with this increase, it is still far below the approximately 80% rate of human-induced fires within all fires [4]. In fires occurring in areas within the blend of wildland areas and human settlements, the people living there and their properties are also endangered [5]. It is very important for city planning experts and related institutions to take into account studies related to fire to minimize these bilateral risks. Because as cities expand towards natural areas, they also implement various transformations in the field to provide the resources

they need (water, energy, transportation, etc.). For example, opening roads, extending power lines, building dams, laying pipelines, opening plantation areas, etc.

In the assessment of risks in rural, forest, and vegetation fires, the selected criteria generally include slope, distance to roads, elevation, distance to water bodies, aspect, type of vegetation, land use, population, climate features, distance to settlements, and fuel type [6-9]. Factors such as slope, elevation, aspect, distance to water bodies, and climate features are directly related to geography. Factors such as population, distance to settlements, distance to roads, and land use are directly human-induced, and vegetation, flammable material type, and even climate features can be considered factors related to humans, albeit indirectly. Vegetation cannot exist solely through natural development, it is being destroyed in many places by humans or its qualities are changed through agricultural and forestry activities, thus it is under anthropogenic domination. Thus, humans partly influence what type of fuel will be for fires in rural and forested areas. The negative human impact on climate change is not just carbon released into the air as a result of industrial activities. For example, it is known that urbanization affects microclimate features in proportion to the width of the dam in places where it is dense and spread over large areas or a dam is built in a region. With the expansion of urban areas, the natural water cycle changes in the areas they cover, significant differences occur in the quality and quantity of sunlight reflection and absorption from the lands, and also in wind speed [10]. In addition to these, anthropogenic heat fluxes and air pollution created lead to local microclimate changes first and then permanent climate changes in the region. The effect of dams on the microclimate is very important for humid and semi-humid regions. It is a fact that dams directly affect the evaporation amount in the region depending on heat, but they also indirectly affect the microclimate because they improve the vegetation on their shores [11].

The commonly used criteria in modeling studies conducted on fire risk assessment in forest and rural area fires are given in Table 1, based on a review of relevant research [7-9, 12-34].

Table 1. Criteria Used in Fire Risk Assessment

Slope	Species composition	Topographic Wetness Index (TWI)
Aspect	Development Stage	Canadian Forest Fire Weather Index (FWI)
Elevation	Solar Radiation	Distance to Fire Response Teams
Distance to settlement	Fire regimes (TSF-FR)	Distance to Fire Watch Towers
Distance to road	Tree Species Composition	Visibility from Fire Watch Towers
Distance to water	Topo-morphology	Distance from the Anti-poaching Camp Shed
Population	Land Use	Distance to Previous Fire Points
Precipitation	Stand type	Topographic Position Index (TPI)
Vegetation Density	Stand age	Tree Stages
Temperature	Stand canopy density	Fuel Type
Vegetation type	Distance to fields	Humidity
Distance from Agricultural Land	Forest Cover	Forest Type
Wind speed	Tree Species	Bare soil index
Stand Crown closure	Land Surface Temperature	Distance to Tourist Spots

It is a known fact that rural areas and forests are under significant encroachment and pressure due to rapid population growth. This has led to an increase in forest and rural area fires. Despite the developments in firefighting technologies in recent years, increased efficiency in prevention activities, reforestation and improvement efforts, and heightened awareness against fires, it has been observed in a report by the Food and Agriculture Organization of the United Nations that the reduction of forests has not been prevented [35]. These total forest losses at a global scale should be considered as partly due to fires and partly due to loss of forest qualities because of land use. However,

regardless of the reason, the decrease in forests and vegetation means less absorption of sunlight and atmospheric carbon, which clearly will contribute to global warming. The rising global warming brings more drought, particularly in the Northern Hemisphere areas close to the equator, due to climate change and causes conditions of low humidity, making forests more susceptible to fire, and once a fire starts, it spreads faster and becomes more difficult to extinguish. This increases the importance of risk assessment studies in terms of taking precautions against fires, planning, and early detection.

The probability of a fire spreading and the damage it will cause is defined as fire risk [8,9]. At the end of the 20th century, the first remote infrared scanners were used to locate a forest fire with detection applications [36]. Today, with the repetition interval and wide area imaging capability of satellite data, it has been possible to gain important information about various fires. Technologies such as remote sensing and geographic information systems contribute significantly to fire prevention by providing capabilities for data collection, analysis, and mapping [8]. To model fire risk, it is necessary to identify factors that influence the probability of fire [37]. Altitude, aspect, slope, weather conditions, vegetation, and human factors are the main reasons affecting the occurrence and spread of fire in a region [38]. In addition, many criteria such as distance to roads, distance to settlements, land use cover are used in risk models [9].

In this study, the forest and rural areas of the European part of Istanbul and the settlements intertwined with these areas were chosen as the research field. There are settlements spread over rural areas and agricultural activity areas in a significant part of the area covered by the city's boundaries. The city, having the highest population nationwide, its geopolitical importance, history, natural resources, economic importance, etc., have led to the selection of this field for research. Istanbul is also the richest city in terms of academic studies. With a population of approximately 15 million, having a larger population than many countries in the world is among the most important factors directing the examination of risks in the region. There are 2500 people per square kilometer [39]. Istanbul has been an important city throughout history. Despite losing its capital status years later, it has never lost its characteristic as a strategic city in terms of economic, political, and commercial importance. Rapid population growth, infrastructure needs, and emerging migration problems make this city an attraction center of Turkey [40]. The city's change reflex has been in the form of growth in recent years, but this has happened uncontrollably. The transformation of fields into building groups, the transformation of the city center into an investment weight, etc., urban growth trends have emerged, but it has been a growth that cannot be controlled and future planning cannot be made [41]. Istanbul also has the feature of being one of the important financial centers of the world. Considering Eastern Europe and Central Asia, it ranks 77th in The Global Financial Centers Index (GFCI) 32 ranking. With a Gross Domestic Product of 232 billion dollars, it is larger than many national economies. It houses half of the country's industrial assets. It has a large share in economic parameters such as the country's businesses, employment, etc., ranging from 30% to 50%. For these reasons, Istanbul has a high disaster vulnerability due to its economic importance [42].

The historical importance of the city stems from its 2500-year-old heritage. Hosting historical and cultural areas such as the Historic Peninsula, the Golden Horn, the Bosphorus, etc., makes the city very valuable. The city was chosen as the European Capital of Culture in 2010. The city's significant place in the world heritage is one of the factors directing the study to this area [43].

Istanbul has experienced many destructive disasters over time. The expectation of an earthquake is frequently mentioned by the scientific community today. Studies are being conducted on a city scale regarding the pre and post-disaster situations. The fact that the city carries this risk is one of the important reasons directing the study to this

area. It is known that fire is expected to occur along with an earthquake [44]. However, the area of interest of this study will not only be urban fires, but also the fire risk created by the natural area within the city and around the city.

In Turkey, fires tend to increase due to the effects of global warming and uncontrolled urbanization. The year 2021 has gone down in history as a very bad year for Turkey in terms of fire. It has been observed that large fires have occurred in 11 provinces in the last 50-year period. According to the records of the General Directorate of Forestry (OGM), 117,734 fire cases have occurred from 1937 to the present, and a very large forest area of 1,851,476 hectares has been lost. This corresponds to approximately 15.73 ha for each fire. In the last 10 years, Muğla (2716), Antalya (2446), Izmir (1649), and then Istanbul (1493), which is the subject of this study, are the four largest provinces in terms of the number of fires. These data are based on the OGM 2022 records [45].

Fire risk maps of a region are created by combining factors that can cause fire [18]. For example, methods such as geographic information systems (GIS), analytic hierarchy process (AHP), fuzzy logic, goal programming (GP), artificial neural networks (ANN), random forest, logistic regression and other machine learning methods are used to create risk maps in forest fires [9].

In this study, the machine learning method, which has recently begun to be used in areas such as fire forecasting, fire risk assessment, creating prediction maps of hazard risk susceptibility, modeling fire behavior, has been used [46-48]. Classification with machine learning methods is one of the methods used in many areas recently. Many studies in the literature show that machine learning methods are efficient in classification problems with high accuracy and minimal deviation [49–54].

Lu et al. [49] have developed a classification model aimed at predicting stadium fire risk. Fire risk data of smart stadiums were used in the model. As a result of the study, it was emphasized that the best performance was in the Gradient Boosting model with a F1 score of 81.9% and an accuracy of 93.2%.

Pang et al. [50] collected fire hotspots, meteorological conditions, land, vegetation, and socioeconomic data from different sources in their study on forest fire prediction. Using these data, they established an artificial neural network, a radial basis function network, a support-vector machine, and a random forest model to determine the thirteen main causes of forest fires in China. As a result of the study, they reported that the prediction accuracies of the four forest fire prediction models were between 75.8% and 89.2%, and the AUC values were between 0.840 and 0.960.

Kalantar et al. [55] conducted forest fire susceptibility prediction based on remote sensing data using resampling algorithms in machine learning models. As a result of the study, they stated that the boosted regression tree model performed better with the highest AUC value of 0.91 compared to other models. They also emphasized that the resampling process improved the prediction performance of all models.

2. Materials and Methods

2.1. Research Site and Data Collection

Istanbul is located at the junction of the European and Asian continents. It consists of two peninsulas, Çatalca and Kocaeli. It is situated on the coast of the Black Sea in the north and the Sea of Marmara in the south. It shares borders with the provinces of Tekirdağ and Kocaeli. It has a geographical location at 41.0122 latitude and 28.976 longitude. (Figure 1). The city has a green vegetation cover. The vegetation consists of forests, maquis, and various tree communities. Due to the climatic influence, dry plant species are prevalent in the northern parts. This feature is something to be considered in terms of fire risk [56].
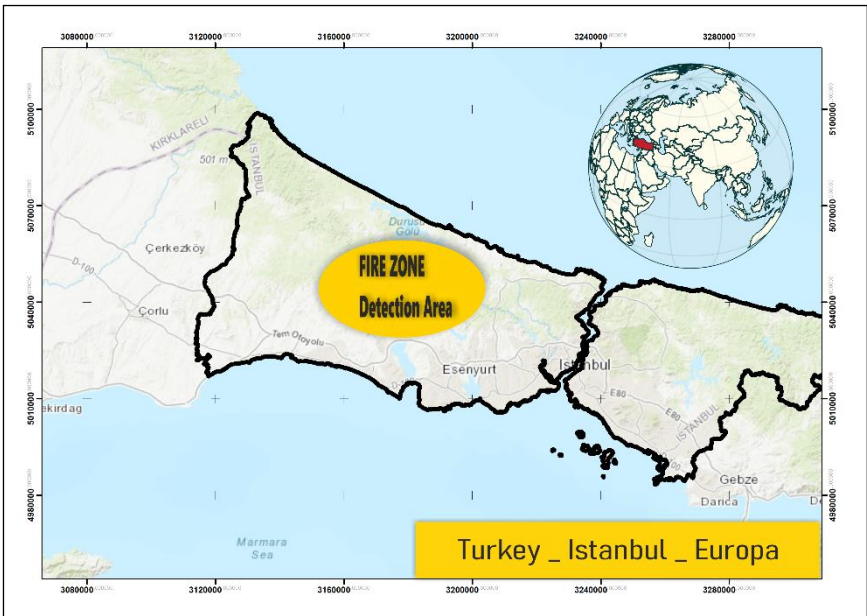
Figure 1. Geographic location of Istanbul (produced using QGIS)

Geographic information systems significantly facilitate the evaluation and interpretation of spatial studies. Therefore, in this study, data obtained with remote sensing techniques were used along with geographic information systems (Figure 2). In order to determine the criteria to be used in the study, previous studies were first examined. As a result of this review, it was determined which factors were preferred as criteria, as shown in Table 2.

The region is densely populated towards the south on the coast of the Marmara Sea and inland from the coast of the Bosphorus in the east. The remaining areas appear to be intermixed between rural-urban, wildland-urban, agricultural space, peri-urban regions, and farms, sometimes forming surface boundaries. In addition to these, it can be seen that there are forests and similar wild areas extending towards the Black Sea to the north.
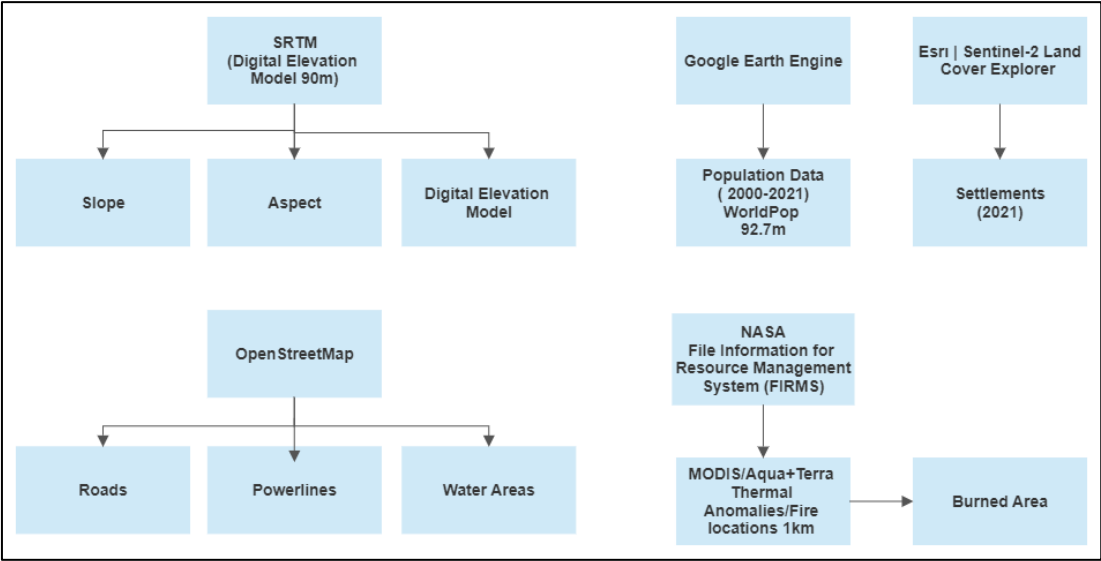
Figure 2. Collected data and its sources

The region is densely populated towards the south on the coast of the Marmara Sea and inland from the coast of the Bosphorus in the east. The remaining areas appear to be intermixed between rural-urban, wildland-urban, agricultural space, peri-urban regions,

and farms, sometimes forming surface boundaries. In addition to these, it can be seen that there are forests and similar wild areas extending towards the Black Sea to the north.

Table 2. Factors used in the literature studies

Factors	12	13	14	15	16	17	8	18	19	7	20	9	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Slope	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Aspect	x		x		x	x		x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x
Elevation	x		x	x	x	x		x	x	x		x	x	x	x	x	x			x		x	x	x	x	x
Distance to settlements	x	x			x	x	x	x	x	x	x	x	x	x		x	x	x		x		x	x	x	x	
Distance to roads	x	x	x		x	x	x	x			x	x	x		x	x	x		x			x	x		x	
Distance to water bodies			x		x			x	x			x			x		x							x	x	
Land use			x	x		x	x							x						x		x		x		x
Precipitation			x		x				x			x		x	x					x						x
Vegetation density										x		x				x	x				x					
Temperature			x											x	x					x						x
Plant type		x				x			x															x	x	
Distance from agricultural land			x						x							x	x									
Wind speed										x			x	x												x
Stand Crown Closure										x	x			x	x				x							
Population										x					x											
Topographic Wetness Index								x				x			x											
Canadian Forest Fire Weather Index (FWI)										x								x								
Tree Stages										x									x							
Fuel Type													x					x								
Humidity									x																	x
Forest type	x																									
Distance to tourist places					x																					
Distance from the anti-poaching Camp Shed					x																					
Distance to fields								x																		
Forest cover								x																		
Distance to previous fire points									x																	
Tree species										x																
Topographic Position Index (TPI)																										
Land surface temperature											x															
Bare soil index													x													
Species composition														x												
Development stage															x											
Solar radiation															x											
Fire regimes (TSF-FR)																		x								
Tree species composition															x				x							
Topo-morphology																					x					
Soil use																					x					
Distance to fire response teams																								x		
Distance to fire watch towers																								x		
Visibility from fire watch towers																								x		
Stand type																									x	
Stand age																									x	
Stand canopy density																									x	
Human Index																					x					

Even in areas that appear to be entirely wild, there are various sized points that are used by humans, referred to as anthropic space. In forest and rural area fires, factors such as topography, climate, and vegetation type are considered crucial. However, for areas affected by human activities within the natural environment, factors such as distance to roads, population, building materials, and power lines or urban infrastructures are also taken into account [5]. Based on this information, the most preferred layers in the studies have been determined, in addition to the ones most necessary in terms of the characteristics of our study area. As a result, seven factors have been identified: slope,

aspect, population, distance to roads, distance to power lines, distance to settlements, and distance to water bodies. The statistical parameters of these factors are presented in Table 3. However, the table also displays eight parameters, including the data of the areas burned in the region between 2000 and 2021.

Table 3. Statistic parameters of dataset

	Count	Mean	Std	Min	25%	50%	75%	Max
Slope (SL) (°)	3455	6.34	6.54	0.00	1.30	4.28	9.09	48.60
Aspect (AS) (°)	3455	134.16	109.03	-1.00	0.00	119.48	234.88	356.55
Digital elevation model (DEM) (*m*)	3455	114.56	73.68	1.00	57.00	104.00	164.00	428.00
Distance to powerlines (DP) (*m*)	3455	4175.25	3678.12	0.00	1315.08	3196.81	6098.97	21801.60
Population (PO) (person*)	3455	16.13	50.24	0.03	0.20	0.94	5.38	470.96
Distance to roads (DR) (*m*)	3455	145.61	183.30	0.00	22.36	80.62	206.03	2046.85
Distance to water areas (DW) (*m*)	3455	1939.42	1396.11	0.00	853.52	1672.00	2773.01	8547.64
Distance to settlements (DS) (*m*)	3455	543.93	799.76	0.00	0.00	257.10	730.31	5734.47
Fire Status (FS) (-)	3455	0.07	0.26	0.00	0.00	0.00	0.00	1.00

The layers created with the data of the factors considered in the study have different characteristics. Detailed feature information is given in Table 4 with the name of the factor data, its source, and resolution information for a better understanding of the data sets. The main source is four data sources. These are Open Street Map (OSM), United States Geological Survey (USGS), Google Earth Engine (GEE), Living Atlas of the World (ArcGIS-Sentinel-2), NASA's File Information for Resource Management System (FIRMS).

Table 4. Data set features

Platform	Data	Source	Resolution
OSM	Road	*http://overpass-turbo.eu*	
	Water Areas	*http://overpass-turbo.eu*	
	Power Line	*http://overpass-turbo.eu*	
USGS	SRTM	*http://earthexplorer.usgs.gov/*	90 m
ArcGIS	Land Cover	*https://livingatlas.arcgis.com/ landcoverexplorer*	10 m - 2021
GEE	WorldPop	*ee.ImageCollection("WorldPop/ GP/100m/pop")*	92.7 m
FIRMS	MODIS+Aqua Terra Thermal Anomalies (Fire Locations)	*https://firms.modaps.eodis.nasa.gov/*	1 km

2.2. Method

The GIS operation steps of the study are shown in Figure 3. Open-source software and platforms were preferred in each GIS operation step.

The Istanbul-European side map was created in the software as a vector data type with a resolution of 1 km. It has been transformed from vector data type to raster data structure using the "polygon to raster" tool from GIS tools. The "raster to point" tool was used to create point data at 1 km intervals.

Another data source, Open Street Map (OSM) platform products, Distance to roads, Water areas, and Power Lines data were obtained from an open access web platform [56] using the code blocks given in Appendix A. The findings obtained are shown in Figure 4. The images created are in the Keyhole Markup Language (KML) file format. It was created as a layer in the GIS software with the "KML to Layer" operation.
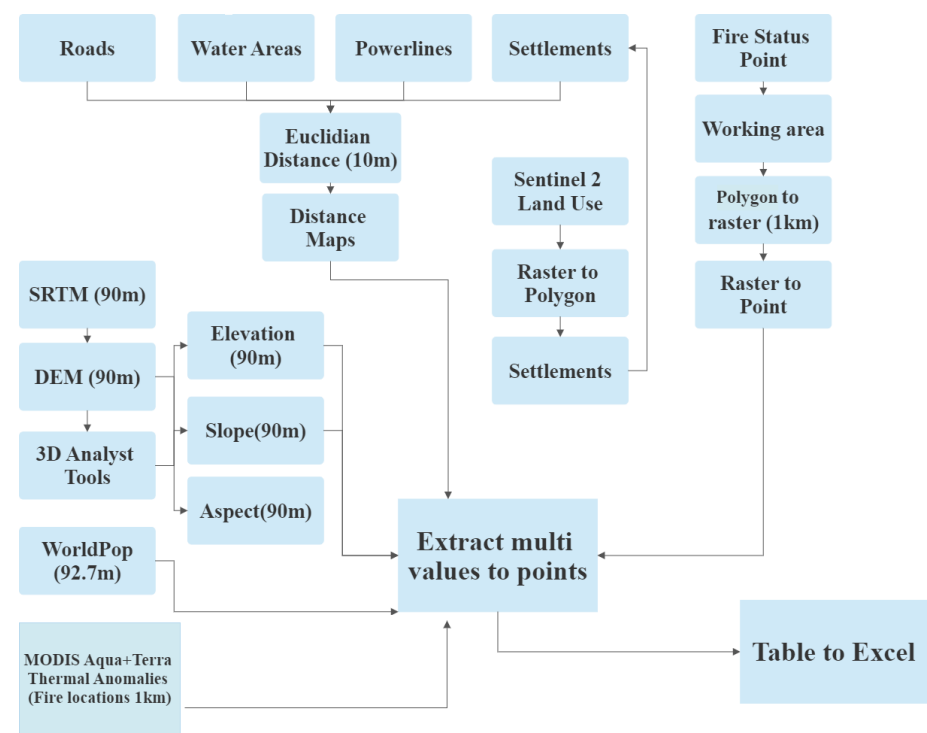
Figure 3. GIS Operating Flow Chart
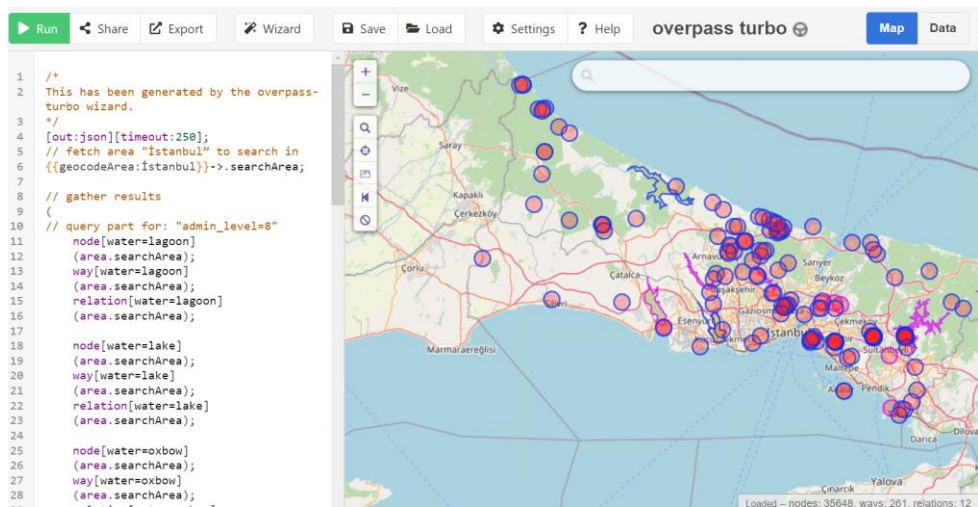
Figure 4. OSM map platform

Another data source used in the study is the Shuttle Radar Topography Mission (SRTM), which provides topographic data. The DEM (Digital Elevation Model) data obtained from this source was used to create a map in the GIS software environment. The map is presented in Figure 5.

The "3D Analyst Tools" tool in the GIS software was used to obtain the slope and aspect data of the study area. The elevation data was used to calculate the slope value. With the help of the DEM data, an aspect map representing the interaction direction of the terrain surfaces with the sun was obtained. Figure 6 shows the slope map, and Figure 7 shows the aspect map.

To identify settlement areas, land use data obtained from Sentinel-2 imagery, provided by Esri, was used for detecting the settlement areas for the year 2021 [58]. The

obtained data was utilized in creating distance maps using the Euclidean Distance method, which measures the distance between two points. The resulting map in Figure 8 was generated through this process. The procedure involved converting the ESRI Sentinel 2 land use data (2021) from raster to polygon format using the "raster to polygon" tool, and selecting the settlement areas. Subsequently, the "euclidean distance" tool was employed to obtain distance maps with a 10-meter interval.
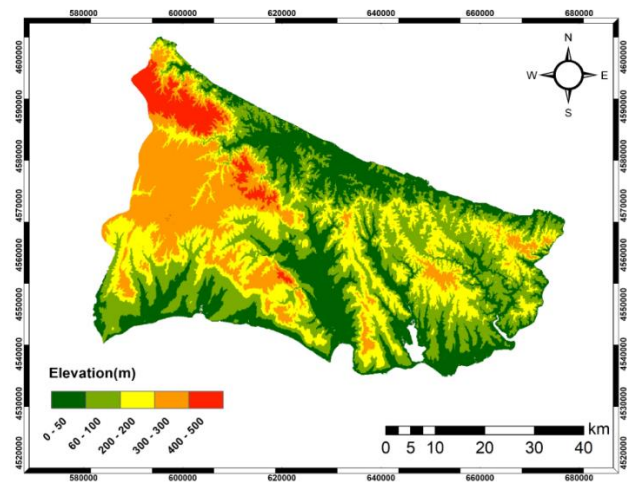
Figure 5. Digital Surface Model (DSM), (DEM)
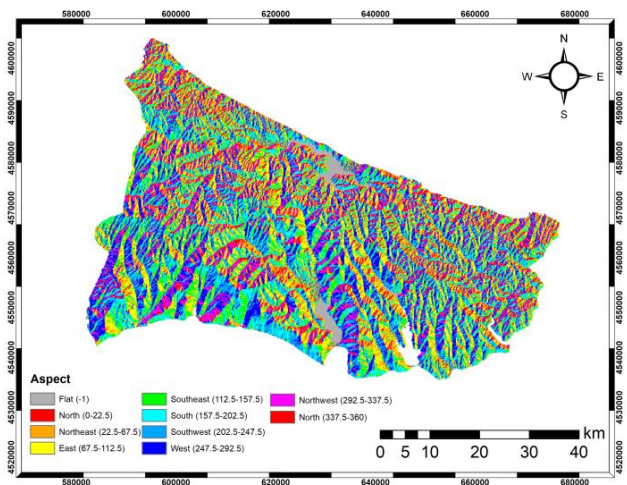
Figure 6. Aspect
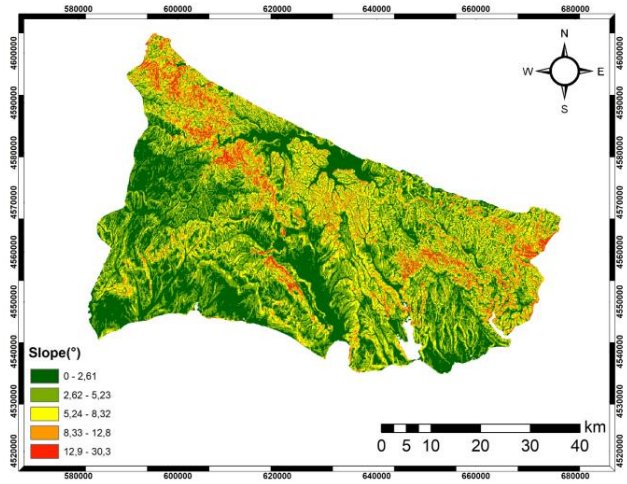
Figure 7. Slope
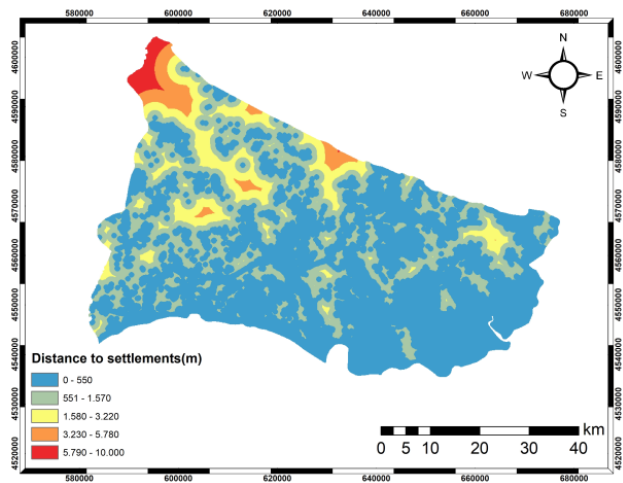
Figure 8. Distance to Settlements

The distance maps for road proximity, water areas, and power lines were obtained using the "euclidean distance" function after extracting the corresponding layers from the OSM dataset. The resulting distance maps are shown in Figures 9 to 11.

The WorldPop dataset, which covers the years 2000 to 2021, was used to map the global population distribution for population density analysis. This dataset was obtained using the Google Earth Engine (GEE) platform, another open-source platform. Based on this data, the map shown in Figure 12 was generated to visualize the population density.
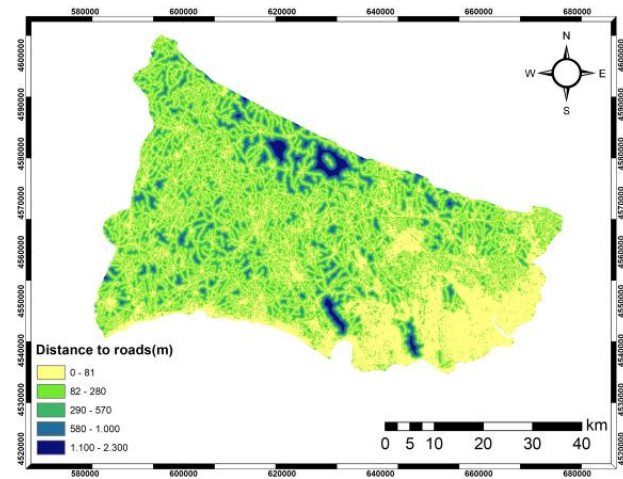
Figure 9. Distance to roads
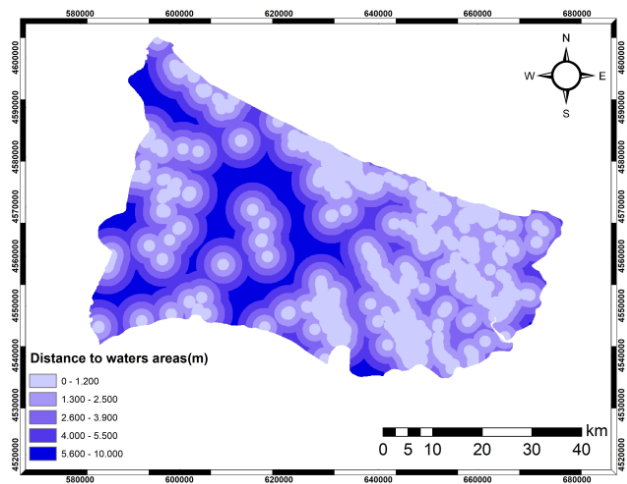
Figure 10. Distance to water areas
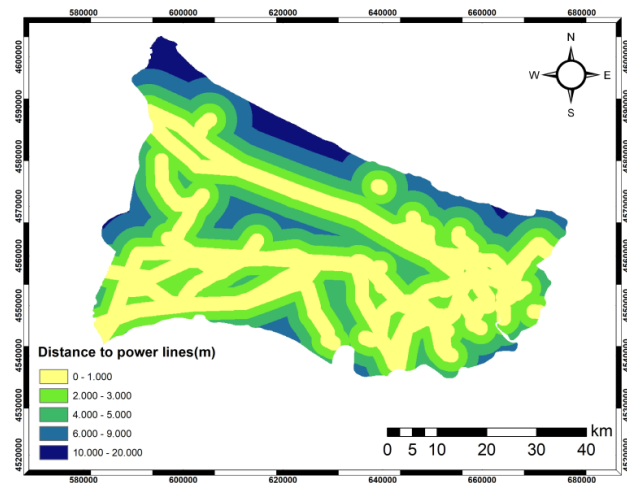
Figure 11. Distance to powerlines
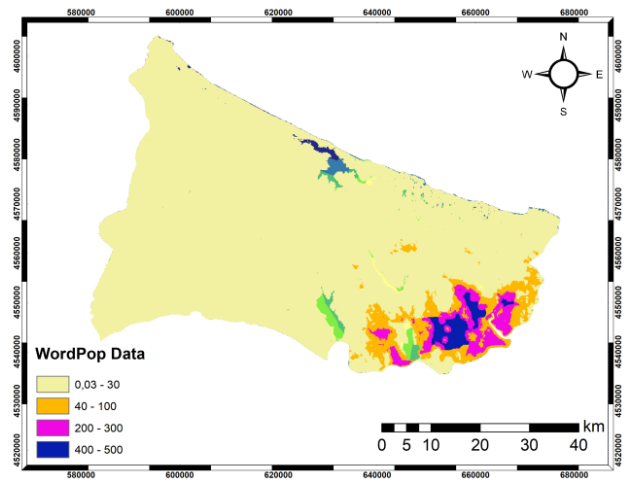
Figure 12. Population

The "MODIS/Aqua+Terra Thermal Anomalies/Fire locations 1km" dataset from NASA's File Information for Resource Management System (FIRMS) platform was used for the detection of burned areas covering the years 2000-2021. These data are obtained from Terra and Aqua satellite images. The created map of the burned area is presented in Figure 13.

Figure 13. Fire locations (between 2000–2021)

Using point data at 1km intervals, the Digital Number (DN) values of pixels were obtained through the "extract multi values to points" tool. An Excel table was created using the "Table to Excel" tool. The DN value matrix of each layer obtained was used as a data source in the subsequent processing step, which involves the Machine Learning method. The aim was to display fire areas from each layer for visual analysis and inferences of the relationship between the layers created and the fire layers. This process was carried out by demonstrating the relevant layer together with the fire layer (Figure 14).

Figure 14. Fire and relationship maps

2.3. Machine learning models

The dataset created within the scope of this study has been randomly divided into a training set (70%) and a test set (30%). Numerous classification models have been established on the training set for fire risk prediction. These include Extra Trees, Random Forest, Light Gradient Boosting, Gradient Boosting, Extreme Gradient Boosting, Decision Tree, AdaBoost, and K-Neighbors machine learning classification methods. The best results were obtained with the Random Forest, Extreme Gradient Boosting, and Light Gradient Boosting methods. A schematic view of the classification based on machine learning is given in Figure 15.

Figure 15. Schematic view of machine learning-based modeling

2.3.1. Random Forest (RF)

The Random Forest method, a community-based classifier developed by Breiman [59], is frequently used in classification and regression problems. The aim of the Random Forest method, an integrated algorithm of the bagging method, is to combine the decisions of a series of classifiers through weighted or unweighted voting [59].

2.3.2. Extreme Gradient Boosting (XGB)

The XGB method, a machine learning method based on Gradient Boosting developed by Chen and Guestrin [60], uses gradient descent in the decision tree to create an optimal model. As a method based on ensemble learning, XGB sequentially builds multiple decision trees considering the impact of high-performance decision tree models, aiming to minimize the errors made by previous decision trees with subsequent decision trees [54].

2.3.3. Light Gradient Boosting (LGB)

It is an improved version of the gradient learning framework based on decision trees and the idea of "weak" learners. Since LGB was developed by Microsoft in 2017, it has been widely applied in many fields as a result of its high prediction accuracy, fast computing speed, and excellent ability to minimize overfitting problems [61].

2.3.4. K-fold cross-validation

In the k-fold cross-validation method used in the study, the data are randomly grouped and divided into "k" subgroups. One of these is used for testing and the remaining "k-1" are used for training. This process is repeated "k" times. The average of the results determines the accuracy of the method [53]. The diagram of the k-fold cross-validation is shown in Figure 16.
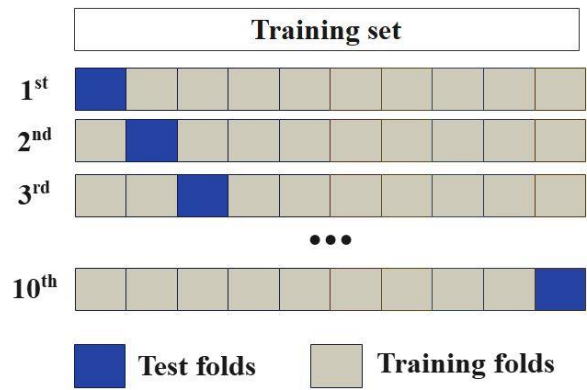
Figure 16. K-fold cross-validation diagram

2.3.5. Model evaluation

In this study, to evaluate the performance of the trained model, the Accuracy, Precision, Recall, F1 score, AUC score, and Receiver Operating Characteristic (ROC) curve, calculated according to the confusion matrix widely used and expressed in Table 5 using equations 1-4, have been used [49]. The ROC curve is the curve of the true positive rate and false positive rate at different classification thresholds. It starts at (0,0) and ends at (1,1). A good model produces a curve that rapidly goes from 0 to 1. The AUC (Area under the ROC curve) summarizes the ROC curve with a single number. The AUC value varies between 0.5 and 1. The highest AUC value indicates an excellent measure of separability, while the lowest AUC value indicates the worst measure of separability.

Table 5. Classification confusion matrix

		Predicted value	
		Fire (class 1)	Non-Fire (class 0)
Actual value	Fire (class 1)	True Positive (TP)	True Negative (TN)
	Non-Fire (class 0)	False Positive (FP)	False Negative (FN)

$$Accurarcy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

$$Recall = \frac{TP}{TP + FN} \qquad (3)$$

$$F1 = 2\frac{Precision \ x \ Recall}{Precision \ + \ Recall} \qquad (4)$$

3. Results and Discussion

This research presents a model for the probability of fire occurrence specific to the European side of Istanbul. This region has a diversity of land use, including rural-urban and natural-urban amalgamations and areas used for agricultural purposes, farms, and cultivated fields. Therefore, predicting fire risk is essential. As the city is considered to be shifted northward to distance it from the fault line passing through the Marmara Sea

due to the expected earthquake, incorporating fire risk into new settlement plans is crucial for developing sustainable residential areas.

It is emphasized in the literature that different machine learning algorithms are quite successful due to their ability to learn and model from data, and this typically yields better results than traditional statistical approaches [62]. In this study, the results obtained from the Random Forest, Extreme Gradient Boosting, and Light Gradient Boosting models are compared with the existing data. The quality parameters used in model validation are given in Table 6. Examining the Accuracy and AUC values, which are among the most important validation parameters, the best result was obtained from the Random Forest model among the three models. The worst results were given by the LGB model for the accuracy parameter and the XGB model for the AUC parameter. However, the difference between them is negligible. But, with the lowest recall value and F1 score, the RF model stands out significantly in terms of accuracy from the XGB and LGB models, which are close to each other.

Table 6. Evaluating parameters for classification models

Model	Accuracy	AUC	Recall	Precision	F1
Random Forest (RF)	0.9293	0.7528	0.0346	0.4000	0.0622
Extreme Gradient Boosting (XGB)	0.9189	0.7409	0.1546	0.3570	0.2119
Light Gradient Boosting (LGB)	0.9107	0.7508	0.1732	0.3454	0.2138

Rodriguez and Riva (2014) used machine learning models to assess human-induced wildfires in Spain between 1988 and 2007 and to predict fire risk. When the whole country is considered, it doesn't entirely consist of natural areas and forests but includes regions with different characteristics like Wildland-Urban Interface (WUI), Rural-Urban Interface (RUI), and Wildland-Agricultural Interface (WAI). The study considered factors like population density, energy lines, railways, and agricultural vehicle density [63]. Except for the railway and agricultural vehicle density, the factors and the manner of field examination are quite similar to this study. As a result of the study, it was revealed that the RF model was the best model with an AUC value of 0.746. In this present study, the RF model also provided the best result with an AUC value of 0.753.

The 10-fold cross-validation method was used in validating the model. The average value of the obtained results was taken into account. The k-fold validation results for all three models are given in Table 7. Upon reviewing the table, it is observed that the standard deviation (std) values for Accuracy and AUC are low. This is true for all three models.

Table 7. K-fold validation accuracy and AUC predictions for classification models

Fold	RF		XGB		LGB	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
0	0.9298	0.6373	0.9050	0.7278	0.9174	0.7244
1	0.9339	0.7425	0.9174	0.7746	0.9132	0.7569
2	0.9298	0.7586	0.9132	0.7762	0.8967	0.7710
3	0.9421	0.8424	0.9091	0.8434	0.9050	0.8165
4	0.9256	0.8203	0.9215	0.7619	0.9091	0.7793
5	0.9215	0.7975	0.9050	0.7371	0.9050	0.7470
6	0.9256	0.8259	0.9132	0.7956	0.9256	0.8058
7	0.9256	0.7240	0.9174	0.7612	0.9132	0.7269
8	0.9253	0.6611	0.9170	0.6799	0.9087	0.6893
9	0.9253	0.7064	0.9087	0.6197	0.9170	0.6657

Mean	0.9285	0.7516	0.9127	0.7478	0.9111	0.7483
Std	0.0056	0.0669	0.0054	0.0590	0.0077	0.0456

The fire risk prediction capacity of classification models was tested using ROC analysis. The AUCs of the ROC graphs for RF, XGB, and LGB models are 0.72, 0.70, and 0.69 respectively for the test data (Figures 17 - 19).
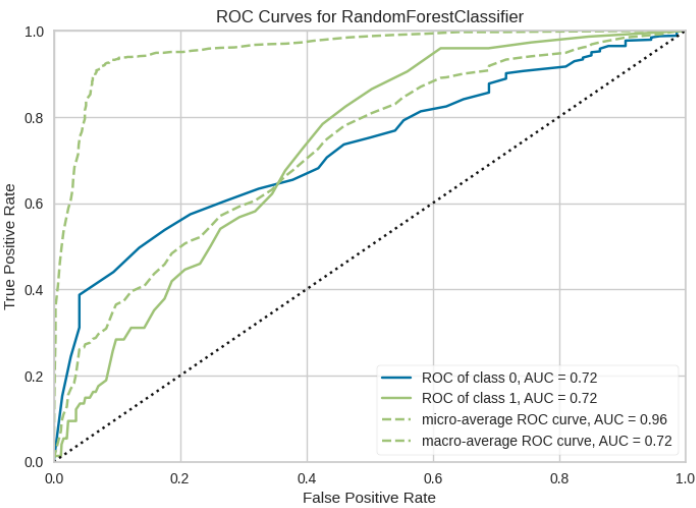
Figure 17. ROC Curves for RF Classifier
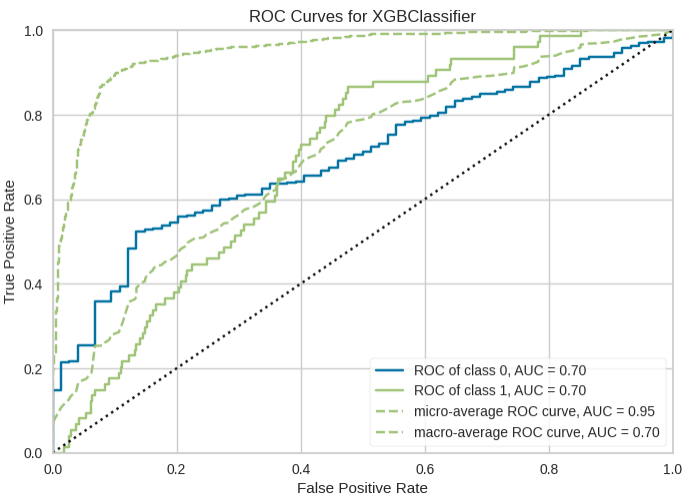
Figure 18. ROC Curves for XGB Classifier
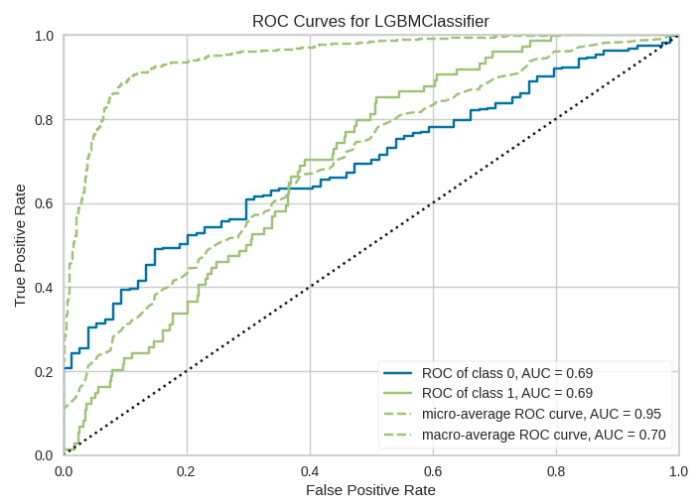
Figure 19. ROC Curves for LGB Classifier

The RF and XGB models have shown the importance of the population factor most significantly (Figures 20, 21). The LGB model, on the other hand, has brought forward the proximity to energy lines (Figure 22). The proximity to energy lines was the second most important factor in the RF model. In the XGB model, the distance to settlements factor took a place of importance close to the value of energy lines in the RF model. The LGB model determined the second most important factor as elevation, at a value close to the first most important factor, proximity to energy lines.
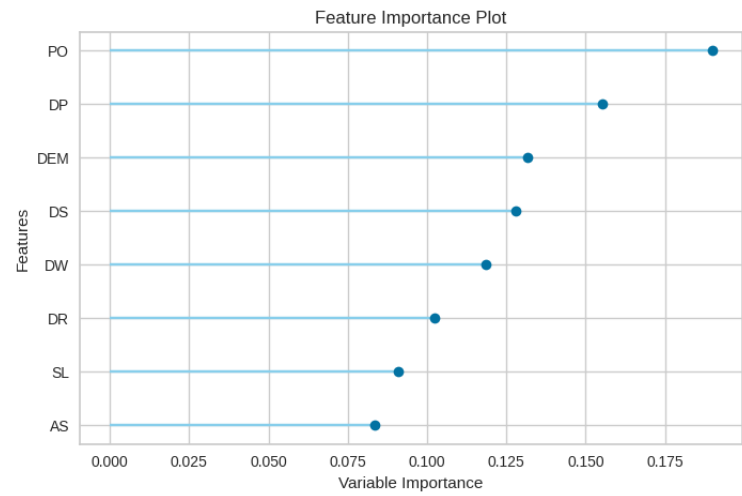
Figure 20. Future importance of RF Classifier
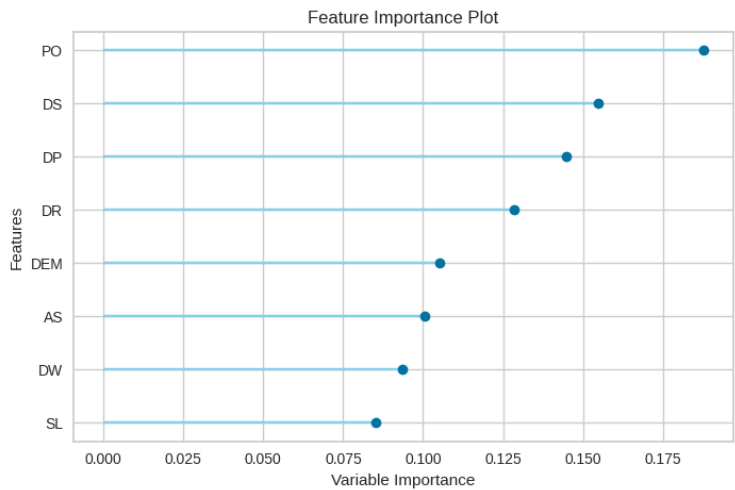
Figure 21. Future importance of XGB Classifier
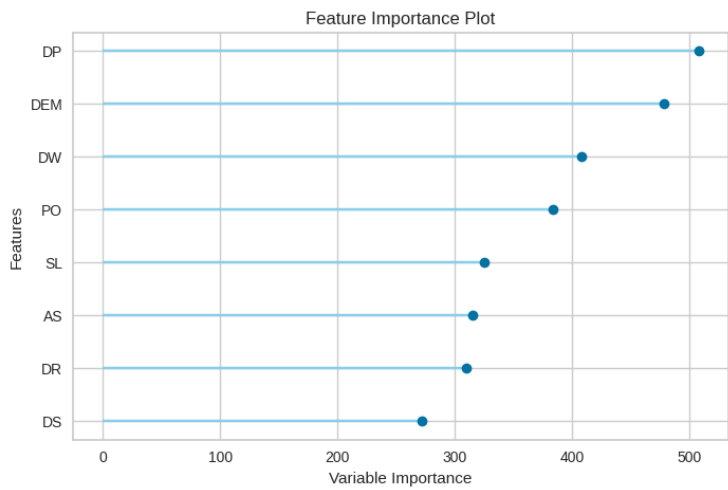
Figure 22. Future importance of LGB Classifier

Here, in the classification of the RF and XGB models, the most important factor was population and the second most important factors were energy lines and distance to settlements, and these are factors that are more commonly found as initiators in fires. On the contrary, in the LGB model, while the energy lines, which are the most important factor, are an initiator, the height, which is the second most important factor, is not among the initiating effects in fires. In fact, as a topographic feature, elevation has a direct effect on temperature, humidity, and wind. As elevation increases, the probability of precipitation usually increases, thus reducing fire intensity [64]. The LGB model placed a factor that has a negative relationship with fire probability as the second most important in the importance ranking. This factor is in the third place in the importance ranking in the RF model and does not have the value in the LGB model's classification graph. Generally, the importance ranking of factors in the classification of RF and XGB models appears to be closer to each other. The RF model is a model that can generally provide high accuracy. It also achieves this accuracy with fewer factor variables. Fewer parameters actually mean easier calibration [63, 65]. Nevertheless, since no single model can ever be perfect enough to always generate correct predictions, it should be evaluated together with other models.

Human population and structures that develop with the population, such as energy lines, roads, and settlements, are determined as factors of high importance for rural area and forest fires, and it is clear that these factors are also important parameters in

planning cities. Based on this, the fire risk in planning settlements that tend to expand towards natural areas can be evaluated with this study model made for the European side of Istanbul and can be safely used in decision-making processes in urban planning. The proposition that this study puts forward for designs in urban planning works is to pull back the approach of determining construction and infrastructure opportunities aimed at reaching sufficient capacities to meet a population size as much as possible. It shows that the dominant influence of the population and its related factors (DS, DP, DR) in planning studies should be reduced and should be included as parameters kept under control.

4. Conclusions

The classification of factors that play a role in fires that occurred in a certain time period has been made in terms of importance through machine learning algorithms. Among the RF, XGB, and LGB models, which have different bases used in these classifications, it was observed that community-based RF performed this with high accuracy and prediction power, and XGB based on decision trees gave results close to RF. Thus, reliable results have been obtained with the determinations of different models that support each other in determining effective factors. This is a study carried out using accessible resources by everyone with data collection from completely open-source platforms instead of data that generally takes a long time to access and collect, using software languages and programs. Therefore, it can be examined in every aspect and it is possible to access data, software, and method details. While it gives results compatible with some similar studies done before, it has revealed higher values in parameters such as AUC, F1 score, accuracy among similar studies where the same models were used.

The model used was successful because the factors that are effective in the start of the fire and the factors that are effective in its spread can be realistically separated with the parameters selected regarding the region where the study was conducted. Therefore, parameters related to the proposal for a change of approach in the planning to be made for other areas of the city that will be opened to settlement and development have emerged with high accuracy.

Author Contributions: Conceptualization, E.A. and K.A; software, E.A. and A.K.; validation, A.K.; investigation, K.A., I.Y., and E.A.; resources, E.A. and I.Y.; data curation, E.A. and A.K.; formal analysis, E.A. and A.K.; writing—original draft preparation, K.A., A.N.A.; writing—review and editing, K.A., A.K. and A.N.A.

Conflicts of Interest: The authors declare no conflict of interest

Appendix A. OSM Code Blocks

Road	Water Areas	Power Line

```*/```
```[out:json][timeout:250];```
```// fetch area "İstanbul" to search in```
```{{geocodeArea:İstanbul}}-```
```>.searchArea;```
```// gather results```
```(```
```// query part for: "admin_level=8"```
```   node[highway=motorway]```
```   (area.searchArea);```
	```*/```
	```[out:json][timeout:250];```
	```// fetch area "İstanbul" to search in```
	```{{geocodeArea:İstanbul}}-```
	```>.searchArea;```
	
	```// gather results```
	```(```
	```// query part for: "admin_level=8"```
	``` node[water=lagoon]```
		```*/```
		```[out:json][timeout:250];```
		```// fetch area "İstanbul" to search in```
		```{{geocodeArea:İstanbul}}-```
		```>.searchArea;```
		```// gather results```
		```(```
		```// query part for: "admin_level=8"```
		``` node[power=line]```
		``` (area.searchArea);```

```
        way[highway=motorway]
        (area.searchArea);
        relation[highway=motorway]
        (area.searchArea);

);
// print results
out body;
>;
out skel qt;

* The layers below, such as "lagoon"
and "lake", have been obtained with the
same method used in the code block
above.
node[highway=trunk]      node[high-
way=primary]  node[highway=second-
ary]
node[highway=tertiary]
node[highway=unclassified]
node[highway=residential]
```

```
        (area.searchArea);
        way[water=lagoon]
        (area.searchArea);
        relation[water=lagoon]
        (area.searchArea);

        node[water=lake]
        node[water=oxbow]

 );
// print results
out body;
>;
out skel qt;

* The layers below, such as "lagoon"
and "lake", have been obtained with the
same method used in the code block
above.
node[water=oxbow]
node[water=rapids]
node[water=river]
node[water=stream]
node[water=river]
node[water=stream_pool]
node[water=reservoir]
node[water=drain]
```

```
        way[power=line]
        (area.searchArea);
        relation[power=line]
        (area.searchArea);
);
// print results
out body;
>;
out skel qt;
```

References

1. Sari, F. Identifying anthropogenic and natural causes of wildfires by maximum entropy method-based ignition susceptibility distribution models. *J. Forestery Res.*, (2022) 1-17. https://doi.org/10.1007/s11676-022-01502-4

2. Vacchiano, G., Foderi, C., Berretti, R., Marchi, E., & Motta, R. Modeling anthropogenic and natural fire ignitions in an inner-alpine valley. *Nat. Hazard. Earth Sys.*, (2018) *18*(3), 935-948. https://doi.org/10.5194/nhess-18-935-2018

3. Romps, D. M., Seeley, J. T., Vollaro, D., & Molinari, J. Projected increase in lightning strikes in the United States due to global warming. *Science*, (2014) *346*(6211), 851-854.   https://doi.org/10.1126/science.1259100

4. Balch, J. K., Bradley, B. A., Abatzoglou, J. T., Nagy, R. C., Fusco, E. J., & Mahood, A. L. Human-started wildfires expand the fire niche across the United States. *P. Natl. A. Sci.*, (2017) *114*(11), 2946-2951.   https://doi.org/10.1073/pnas.1617394114

5. Bento-Gonçalves, A., & Vieira, A. Wildfires in the wildland-urban interface: Key concepts and evaluation methodologies. *Sci. Total Environ.*, (2020) 707, 135592. https://doi.org/10.1016/j.scitotenv.2019.135592

6. Hoang, T.V., Chou, T.Y., Fang, Y.M., Nguyen, N.T., Nguyen, Q.H., Canh, P.X., Toan, D.N.B., Nguyen, X.L. and Meadows, M.E. Mapping Forest Fire Risk and Development of Early Warning System for NW Vietnam Using AHP and MCA/GIS Methods. *Appl. Sci.*, (2020) 10, 4348-4366. https://doi.org/10.3390/app10124348

7. Kumari, B. and Pandey, A.C. Geo-imformatics based multi-criteria decision analysis (MCDA) through analytic hierarchy process (AHP) for forast fire risk mapping in Palamau Tiger Reserve, Jharkhand state, India. *J. Earth Syst. Sci.*, (2020) 129 204. https://doi.org/10.1007/s12040-020-01461-6

8. Fekir, Y., Hamadouche, M. A., & Anteur, D. Integrated approach for the assessment of forest fire risk and burn severity mapping using GIS, AHP method, and Google Earth Engine in Western Algeria. *Euro-Mediterranean J. Environ. Integ.*, (2022) 1-14. https://doi.org/10.007/s41207-022-00338-y

9. Novo, A., Fariñas-Álvarez, N., Martínez-Sánchez, J., González-Jorge, H., Fernández-Alonso, J. M., & Lorenzo, H. Mapping forest fire risk—a case study in Galicia (Spain). *Remote Sens.* (2020) *12*(22), 3705. https://doi.org/10.3390/rs12223705

10. Liu, Y., Lu, H., Tian, P., & Qiu, L. Evaluating the effects of dams and meteorological variables on riparian vegetation NDVI in the Tibetan Plateau. *Sci. Total Environ.* (2022) *831*, 154933. https://doi.org/10.1016/j.scitotenv.2022.154933

11. Vinayak, B., Lee, H. S., Gedam, S., & Latha, R. Impacts of future urbanization on urban microclimate and thermal comfort over the Mumbai metropolitan region, India. *Sustain. Cities Soc.* (2022) *79*, 103703. https://doi.org/10.1016/j.scs.2022.103703

12. Ajin, R., Loghin, A.-M., Vinod, P., & Jacob, M. Forest Fire Risk Zone Mapping Using RS and GIS Techniques: A Study in Achankovil Forest Division, Kerala, India. *J. Earth, Environ. Health Sci.* (2016) *2*(3), 109. https://doi.org/10.4103/2423-7752.199288

13. Akay, A. E. Integrated Use of AHP and GIS Techniques for Generating Forest Fire Risk Map in Karacabey Flooded Forest. *Environ. Sci. Proc.,* (2021) 3(1), 62. <u>https://doi.org/10.3390/iecf2020-08061</u>

14. Akay, A. E., & Şahin, H., Forest Fire Risk Mapping by using GIS Techniques and AHP Method: A Case Study in Bodrum (Turkey). *European J. Forest Eng.*, June (2019) 25–35. https://doi.org/10.33904/ejfe.579075

15. Akbulak, C., Tatlı, H., Aygün, G., & Sağlam, B. Forest fire risk analysis via integration of GIS, RS and AHP: The Case of Çanakkale, Turkey. *J. Hum. Sci.* (2018) 15(4), 2127. https://doi.org/10.14687/jhs.v15i4.5491

16. Cetin, M., Isik Pekkan, Ö., Ozenen Kavlak, M., Atmaca, I., Nasery, S., Derakhshandeh, M., & Cabuk, S. N. GIS-based forest fire risk determination for Milas district, Turkey. *Nat. Hazards* (2022) *0123456789*. https://doi.org/10.1007/s11069-022-05601-7

17. Chavan, M. E., Das, K. K., & Suryawanshi, R. S. Forest fire risk zonation using Remote Sensing and GIS in Huynial watershed , Tehri Garhwal district , UA. *Int. J. Basic Appl. Res.* (2012) *02*, 6–12. ISSN - 2249-3352

18. Jaiswal, R. K., Mukherjee, S., Raju, K. D., & Saxena, R. Forest fire risk zone mapping from satellite imagery and GIS. *Int. J. Appl. Earth Observ. Geoinf.* (2002) *4*(1), 1–10. https://doi.org/10.1016/S0303-2434(02)00006-5

19. Kayet, N., Chakrabarty, A., Pathak, K., Sahoo, S., Dutta, T., & Hatai, B. K. Comparative analysis of multi-criteria probabilistic FR and AHP models for forest fire risk (FFR) mapping in Melghat Tiger Reserve (MTR) forest. *J. Forestry Res.* (2020) *31*(2), 565–579. https://doi.org/10.1007/s11676-018-0826-z

20. Lamat, R., Kumar, M., Kundu, A., & Lal, D. Forest fire risk mapping using analytical hierarchy process (AHP) and earth observation datasets: a case study in the mountainous terrain of Northeast India. *SN Appl. Sci.* (2021) *3*(4), 1–15. https://doi.org/10.1007/s42452-021-04391-0

21. Mohammadi, F., Bavaghar, M. P., & Shabanian, N. Forest Fire Risk Zone Modeling Using Logistic Regression and GIS: An Iranian Case Study. *Small-Scale For.* (2014) *13*(1), 117–125. https://doi.org/10.1007/s11842-013-9244-4

22. Naskar, S., Rahaman, A., & Biswas, B. Forest Fire Susceptibility Mapping of West Sikkim District, India using MCDA techniques of AHP & TOPSIS model. *Res. Square* (2022) 1–23. https://doi.org/10.21203/rs.3.rs-1753672/v1

23. Nikhil, S., Danumah, J. H., Saha, S., Prasad, M. K., Rajaneesh, A., Mammen, P. C., Ajin, R. S., & Kuriakose, S. L. Correction to: Application of GIS and AHP Method in Forest Fire Risk Zone Mapping: a Study of the Parambikulam Tiger Reserve, Kerala, India (Journal of Geovisualization and Spatial Analysis, (2021), 5, 1, (14), 10.1007/s41651-021-00082-x). *J. Geovisual. Spatial Anal.* *5*(1). https://doi.org/10.1007/s41651-021-00083-w

24. Nuthammachot, N., & Stratoulias, D. Multi-criteria decision analysis for forest fire risk assessment by coupling AHP and GIS: method and case study. *Environ. Dev. Sustain.* (2021) *23*(12), 17443–17458. https://doi.org/10.1007/s10668-021-01394-0

25. Nuthammachot, N., & Stratoulias, D. A GIS- and AHP-based approach to map fire risk: a case study of Kuan Kreng peat swamp forest, Thailand. *Geocarto Int.* (2021) *36*(2), 212–225. https://doi.org/10.1080/10106049.2019.1611946

26. Özşahin, E. Forest Fire Susceptibility Analysis Using Gis and Ahp: the Case of Antakya Forestry Operation Directorate, Cbs ve Ahs Kullanılarak Orman Yangını Duyarlılığı Analïzï: Antakya Orman Işletme Müdürlüğü Örneğï. *Route Edu. Social Sci. J.* (2014) *1*(3), 50-71.

27. Pandey, K., & Ghosh, S. K. Modeling of parameters for forest fire risk zone mapping. *Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci. - ISPRS Archives* (2018) *42*(5), 299–304. https://doi.org/10.5194/isprs-archives-XLII-5-299-2018

28. Qtaishat, K., Shatnawi, N., & Habib, M. Forest Fire Risk Zonation Using Remote Sensing and GIS Technology "Case study in Jordan". *Forest Fire Int. J. Environ. Global Climate Change* (2017) 5(2), 1-8. ISSN 2310-6743

29. Rahmani, S., & Benmassoud, H. Modelling of forest fire risk spatial distribution in the region of Aures, Algeria. *Geoadria* (2019) *24*(2), 79–92. https://doi.org/10.15291/ geoadria.2846

30. Setiawan, I., Mahmud, A. R., Mansor, S., Mohamed Shariff, A. R., & Nuruddin, A. A. GIS-grid-based and multi-criteria analysis for identifying and mapping peat swamp forest fire hazard in Pahang, Malaysia. *Disaster Prev. Manag: An Int. J.* (2004) *13*(5), 379–386. https://doi.org/10.1108/09653560410568507

31. Sivrikaya, F., & Küçük, Ö. Modeling forest fire risk based on GIS-based analytical hierarchy process and statistical analysis in Mediterranean region. *Ecol. Inform.*, (2022) *68* (September 2021). https://doi.org/10.1016/j.ecoinf.2021.101537

32. Soydan, O. Determination of Forest Fire Risk Using GIS: A Case Study in Nigde, Turkey. *Bartın Orman Fakültesi Dergisi* (2022) *24*(1), 77–94. https://doi.org/10.24011/barofd.1078642

33. Suryabhagavan, K. V., Alemu, M., & Balakrishnan, M. Gis-based multi-criteria decision analysis for forest fire susceptibility mapping: A case study in Harenna forest, southwestern Ethiopia. *Trop. Ecol.* (2016) *57*(1), 33–43.

34. Van Hoang, T., Chou, T. Y., Fang, Y. M., Nguyen, N. T., Nguyen, Q. H., Canh, P. X., Toan, D. N. B., Nguyen, X. L., & Meadows, M. E. Mapping forest fire risk and development of early warning system for NW vietnam using AHP and MCA/GIS methods. *Appl. Sci. (Switzerland)* (2020) *10*(12), 1–19. https://doi.org/10.3390/app10124348

35. The State of the World Forests - Forests, Biodiversity and People, 2020, Food and Agriculture Organization of the United Nation. Available online: *https://www.fao.org/3/ ca8642en/ca8642en.pdf* (accessed in 03 May 2023)

36. Chuvieco, E., and Congalton, R., Application of Remote Sensing and Geographic Information Systems to Forest Fire Hazard Mapping, *Remote Sens. Environ.* (1989) 29:147-159. https://doi.org/10.1016/0034-4257(89)90023-0

37. Eskandari, S.; Chuvieco, E. Fire danger assessment in Iran based on geospatial information. *Int. J. Appl. Earth Observ. Geoinf.* (2015) 42, 57–64. https://doi.org/10.1016/j.jag.2015.05.006

38. Yakubu, I.; Mireku-Gyimah, D.; Duker, A.A. Review of methods for modelling forest fire risk and hazard. *Afr. J. Environ. Sci. Technol.* (2015) 9, 155–165. https://doi.org/10.5897/AJEST2014.1820

39. Turkish Statistical Instutie (TUIK), Population and demography database, (2022), Available online: https://data.tuik.gov.tr (accessed in 10 May 2023)

40. Gürel, A. and Balta, Y. İstanbul' un göç olayı ve etnik hayat üzerine. *Marmara Sosyal Araştırmalar Dergisi* (2011) 1: 1-15.

41. Kubat, A. S. and Hazar D. İstanbul'un Çeper Kuşak Gelişim Süreci. *Türkiye Kentsel Morfoloji Ağı* (2018) 693-711. ISBN: 978-605-80820-1-4

42. İstanbul Ekonomisi, (2021) İstanbul Büyükşehir Belediyesi, İstanbul Planlama Ajansı, İstanbul İstatistik Ofisi. Available online: https://ipa.istanbul/wp-content/uploads/2021/08 /son-verilerle-i-stanbul-ekonomisi-2021-1-c-eyrek.pdf (accessed in 13 May 2023) (In Turkish)

43. Dinçer, İ., Enlil, Z., and Evren, Y. İstanbul' un koruma alanlarının değerlendirilmesi. *Megaron* (2009) *3*(3).

44. Üstün, A. K., and Angün, A. S. İstanbul'un ilçelerinin afet yönetimi açısından önem ağırlıklarının analitik hiyerarşi süreci kullanılarak belirlenmesi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi* (2016) 31(1) https://dx.doi.org/10.17341/gummfd.8342

45. Atmiş, E., Tolunay, D. and Erdönmez, C. Bölüm I-II, Orman Yangınlarının Sayısal Analizi, Orman Yangınları, Kavgacı, A., Başaran, M.A. (Editors), *Türkiye Ormancılar Derneği Yayını* (2023) 422 s. Ankara, ISBN: 978-605-71791-4-2

46. Xie, Y., & Peng, M. Forest fire forecasting using ensemble learning approaches. *Neural Comput. Appl.* (2019) *31*, 4541-4550. https://doi.org/10.1007/ s00521-018-3515-0

47. Pourghasemi, H. R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., & Cerda, A. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Sci. Rep.* (2020) *10*(1), 3203. https://doi.org/ 10.1038/s41598-020-60191-3

48. Rubí, J. N., de Carvalho, P. H., & Gondim, P. R. Application of machine learning models in the behavioral study of forest fires in the Brazilian Federal District region. *Eng. Appl. Artif. Intel.* (2023) *118*, 105649. https://doi.org/10. 1016/j.engappai.2022.105649

49. Lu Y, Fan X, Zhao Z, Jiang X. Dynamic Fire Risk Classification Prediction of Stadiums: Multi-Dimensional Machine Learning Analysis Based on Intelligent Perception. *Appl. Sci.* (2022) 12(13), 6607. https://doi.org/10.3390/rs12244157.

50. Pang Y, Li Y, Feng Z, Feng Z, Zhao Z, Chen S, et al. Forest Fire Occurrence Prediction in China Based on Machine Learning Methods. *Remote Sens.* (2022) *14*(21), 5546. https://doi. org/10.3390/rs14215546.

51. Mohajane M, Costache R, Karimi F, Bao Pham Q, Essahlaoui A, Nguyen H, et al. Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area. *Ecol. Indic.* (2021) 129:107869. https://doi.org/10.1016/ J.ECOLIND.2021.107869.

52. Dwiasnati, S. and Devianto, Y. Classification of forest fire areas using machine learning algorithm. *World J. Adv. Eng. Technol. Sci.* (2021) *3*(1), 008-015. https://doi.org/10.30574/ wjaets.2021.3.1.0048.

53. Ercan, U. Estimating of Households Shopping on The Internet Using Random Forest Method. *Kafkas Üniversitesi İktisadi ve İdari Bilim Fakültesi Derg*isi (2021); 12:728–52. https://doi.org/10.36543/kauiibfd.2021.030.

54. Ercan, U. Determining Airline Customer Satisfaction with Ensemble Learning Methods. *Alanya Acad. Rev.* (2022) 6:2763–74. https://doi.org/10.29023/ALANYAAKADEMIK.1095574.

55. Kalantar B, Ueda N, Idrees MO, Janizadeh S, Ahmadi K, Shabani F. Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sens.* 2020, *12*(22), 3682, https://doi.org/10.3390/ rs12223682.

56. Gürel, A., and Gündüz, A. E. İstanbul' un Ekolojik Yapısı Üzerine Bir Araştırma. *Marmara Sosyal Araştırmalar Dergisi* (2011) (1), 1-12.

57. Overpass turbo, (2022), Available online: *http://overpass-turbo.eu* (accessed in 10 May 2023)

58. Esri, Sentinel-2 Land Cover Explorer, (2022), Available online: *https://livingatlas.arcgis. com/landcoverexplorer/* (accessed in 10 May 2023)

59. Breiman L. Random forests. *Mach. Learn.* (2001) 45:5–32. https://doi.org/10.1023/A: 1010933404324.

60. Chen, T., & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016, August) 785-794. https://dx.doi.org/10.1145/2939672.2939785.

61. Fan J, Ma X, Wu L, Zhang F, Yu X, Zeng W. Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agr. Water Manag.* (2019) 225:105758. https://doi.org/10.1016/J.AGWAT.2019.105758.

62. Jain, P., Coogan, S. C., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. A review of machine learning applications in wildfire science and management. *Environ. Rev.* (2020) *28*(4), 478-505. https://doi.org/10.1139/er-2020-0019

63. Rodrigues, M., & De la Riva, J. An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environ. Modell. Softw.* (2014) *57*, 192-201. https://doi.org/10.1016/j.envsoft.2014.03.003

64. Tehrany, M. S., Jones, S., Shabani, F., Martínez-Álvarez, F., & Tien Bui, D. A novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using LogitBoost machine learning classifier and multi-source geospatial data. *Theor. Appl. Climatol.* (2019) *137*, 637-653. http://doi.org/10.1007/s00704-018-2628-9

65. Michael, Y., Helman, D., Glickman, O., Gabay, D., Brenner, S., & Lensky, I. M. Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series. *Sci. Total Environ.* (2021) *764*, 142844. https://doi.org/10.1016/j.scitotenv.2020.142844