ENSEMBLING OF DECISION TREES, KNN, AND LOGISTIC REGRESSION WITH SOFT-VOTING METHOD FOR WILDFIRE SUSCEPTIBILITY MAPPING

A. Rezaei Barzani[1], P. Pahlavani[1*], O. Ghorbanzadeh[2]

[1] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran – (ali.rezaeibarzani, pahlavani@ut.ac.ir)

[2] Institute of Advanced Research in Artificial Intelligence (IARAI), Landstraßer Hauptstraße 5,1030 Vienna, Austria – omid.ghorbanzadeh@iarai.ac.at

Commission IV, WG IV/3

KEY WORDS: Forest fire, Bushfire, Machine learning, Ensemble model, Soft-voting.

ABSTRACT

As a result of climate change, climatic catastrophes, such as wildfires, are likely to increase. Wildfires continue to occur frequently and spread with greater intensity due to extreme weather conditions. In recent years, explosive fire growths have been reported in the United States, Australia, and other parts of the world. A combination of climate change and human activity has caused the semi-arid forestry areas in Iran's northern provinces to become more desiccated, leading to an increase in wildfires. The accuracy of the resulting fire susceptibility maps (FSMs) will directly be related to the performance of the method classifier. In this study, we use an ensemble classifier to model the FSM for a selected forestry case study area in one of the northern provinces of this country. Therefore, FSM is generated based on established criteria using the ensemble model. With Decision Trees, K nearest neighbor, and Logistic Regression, the ensemble model was created using the soft-voting method. A forest fire inventory data is created based on data collected over five years using GPS and the MODIS thermal anomalies product for training and testing the applied approach. The K-fold method was used for validation, and the resulting FSM was validated using five accuracy assessment metrics. The best result from the area under the curve (AUC) yields 93% for fold 9, and the mean AUC for ten folds yields 88%.

1. INTRODUCTION

The threat of wildfire disasters has increased due to global warming. Consequently, forest ecosystems are increasingly at risk (Papathoma-Köhle et al., 2022). However, a wildfire may have natural, accidental, or even criminal origins (Sayad et al., 2019). As fire is one of the significant threats to forestry areas, this hazard, whether natural or manufactured, can negatively impact the environment and the economy of local communities. Forests are vital resources worldwide, and in Iran, they provide an essential source of income to local communities in the northern Provinces. Globally, fire is a significant contributor to the development and degradation of forests. The climate of Iran ranges from arid to semi-arid, and wildfires are increasingly occurring across a large portion of the country's forests in the northern provinces (Adab et al., 2015). Moreover, this region is becoming warmer and drier due to climate change. Hence, there is an urgent need for studies on the spatial prediction of wildfires through modeling fire susceptibility maps (FSMs) to mitigate wildfire's adverse impacts. The modeling FSMs depends on a conditional wildfire criterion including anthropological, vegetation, topographical, meteorological, and related hydrological criteria. Wildfire analysis and susceptibility mapping are significantly enhanced by remote sensing satellite imagery and GIS application (Karimi et al., 2021; Pourghasemi, 2016; Pradhan et al., 2007). Moreover, machine learning (ML) classifiers have proven superior for spatial hazard predictions during the last two decades (Jaafari et al., 2019; Tien Bui et al., 2019; Watson et al., 2019). Two factors have led to the wide use of ML classifiers; first, the higher availability of thermal anomaly products from satellites like the moderate resolution imaging spectrometer (MODIS) to generate wildfire inventory data sets for training and to test the classifiers, and second, the technology evolution in computing platforms. The use of ML classifiers for wildfire spatial prediction and susceptibility mapping is currently gaining much attention. Therefore, several studies have examined them for modeling the FSMs in forest areas worldwide. Iban et al. (Iban and Sekertekin, 2022) evaluated seven ML classifiers, including random forest (RF), AdaBoost (AB), and support vector machine (SVM), for modeling FSMs for two provinces of Adana and Mersin in Turkey. In their accuracy assessment process, the lowest and highest accuracy scores for the applied seven different ML classifiers were 0.734 and 0.812. Sharma et al. (Sharma et al., 2022) used four ML classifiers of the Boosted Regression Tree, Extreme Gradient Boosting, Fuzzy Forest (FF), and the RF for wildfire severity prediction in different regions of Victoria, Australia, and the FF showed the highest accuracy. They used six wildfire severity criteria: soil temperature and moisture, air temperature and pressure, relative humidity, and vertical wind. For the Northern Beaches area of Sydney, Australia, thirty-six wildfire conditional factors were selected and used by Naderpour et al. (2021). Using an optimized deep neural network, they could create an FSM with an accuracy of more than 95% ROC. In another case study area in Australia, Hosseini and Lim et al. Hosseini and Lim (2022) applied eight classifiers like logistic regression and SVM for bushfire susceptibility mapping for New South Wales. They selected eight conditional factors according to a literature review of studies that have been done elsewhere, such as Huichang County, China; the Liguria region in Italy; the Zagros Mountains, Iran; and Wuyishan Scenery District, China. Their selected criteria included the elevation, aspect, slope, annual temperature and precipitation, normalized difference vegetation

https://doi.org/10.5194/isprs-annals-X-4-W1-2022-647-2023

index (NDVI), land cover, and distance to roads. The wildfire-prone areas of Iran have also been subjected to studies using similar methodologies. The performance of three classifiers of the RF, artificial neural network (ANN), and SVM were evaluated by Ghorbanzadeh et al. (2019) for spatial prediction of wildfires in Mazandaran Province, Iran. Their accuracy assessment was based on four-fold cross-validation (CV), and RF resulted in the best prediction by 88% ROC. In another study, Tavakkoli et al. (2022) employed the Google Earth Engine (GEE) software to evaluate coarse and medium spatial resolution data that influenced the ML models for FSM modeling for the forests of the Guilan Province, Iran. The researchers also use the Dempster-Shafer theory (DST) and 14 wildfire conditioning criteria to fuse predictions from different resolutions. Golestan Province, located in northeast Iran, was also selected as the case study area for FSM modeling by Eskandari et al. (2021). Several classifiers have been compared using a training data set that often included criteria of the elevation, slope, plan curvature, topographic wetness index (TWI), annual rainfall mean, annual temperature mean, wind effect, distance to urban areas, distance to streams, and distance to roads. The conditional criteria were selected based on a review of relevant research findings from other regions, such as the Hyrcanian ecoregion, Iran; Dayu County, China; the Western Mazandaran Province, Iran; and forest areas of Vietnam. Since forest areas have different topographies, climate zones, settlement densities, human activity, etc. In this study, using an ensemble model for generating FSM would be investigated.

2. STUDY AREA AND DATA

2.1 Study area

The Alborz Mountains are part of the much larger Alpide belt and run from Ardabil Province to the western and southern coasts of the Caspian Sea until the northern parts of Khorasan Province in northeast Iran. In this study, we selected a forestry region that mostly covered the northern part of the mountainous regions of Amol County in the Mazandaran Province. The study area is located in the Central Alborz region on the southern coast of the Caspian Sea (see Figure 1). There are different agroecosystems in Amol County, including plains, forests, and rainforests, and forests in mountainous areas of this county are selected for further analyses and FSM modeling.

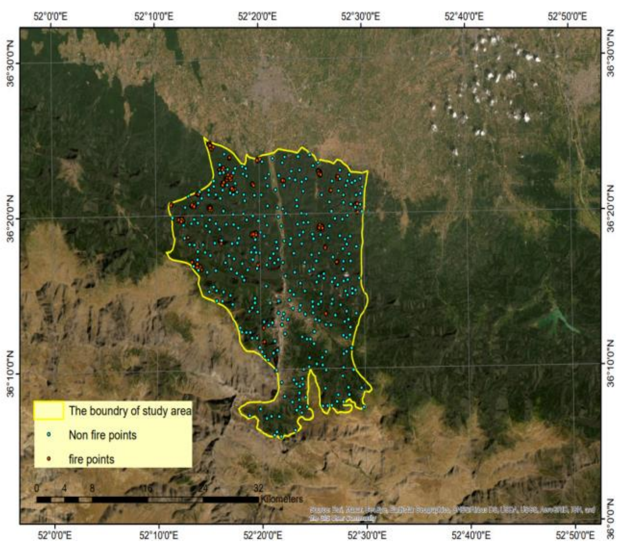

Figure 1. Location map of the study area.

2.2 Wildfire data

A fire inventory is an important first step in generating FSMs using ML classifiers. For any ML classifier, accurate and adequate inventory data is crucial for training the classifier. In this study, the freely available MODIS moderate resolution imaging spectroradiometer (MODIS) thermal anomalies product is used to generate wildfire inventory data for 2012 to 2017. Over the study period, 34 fire polygons were detected covering 17,420 pixels. Moreover, GPS data collected through field surveys were utilized to evaluate and manually correct the detected polygons. In this way, GPS data and MODIS data were combined to create the wildfire inventory data. The manual corrections were made using ArcGIS software. Except for the fire polygons, non-forest fire points were also randomly distributed within the entire study area for training our machine learning models. Conditioning criteria play a vital role in training machine learning models to generate FSMs. This study intently selects the conditional criteria applied in previous studies that have been done in the same study area or geographically similar regions. The applied criteria of distance to the settlements, recreational areas and roads, and land use are our anthropologically selected criteria in this study; the normalized difference vegetation index (NDVI) is considered the vegetation criterion derived from Landsat-8 https://earthexplorer.usgs.gov/. This research uses the advanced space-borne thermal emission and reflection radiometer (ASTER) freely available from https://asterweb.jpl.nasa.gov/ to generate the elevation, slope, slope aspect, landforms, the topographic wetness index (TWI), and plan curvature (PC). Three criteria of the wind effect, annual temperature, and potential solar radiation were selected as the meteorological criteria. The wildfire conditional hydrological criteria are distance to streams and annual rainfall. All criteria are shown in Table 1.

anthropological	vegetation	topographical	meteorological	Hydrological
Distance to settlements (m)	NDVI	Elevation (m)	Wind effect	Distance to stream (m)
Distance to road (m)		Slope (%)	Annual temperature (C)	Annual rainfall (mm)
Recreation area (m)		Slope aspect	Potential solar radiation	
Land use		Landforms		
		TWI		
		PC (100/m)		

Table1. Listed are the conditional wildfire criteria selected for Amol County.

3. METHODS

This study aimed to prepare a fire susceptibility map. According to our literature review and the fact that the famous SVM and Rf models have been used in this region (Ghorbanzadeh et al., 2019a, 2019b), 3 classifiers were selected in this study, and in order to improve their accuracy, the soft voting method was used to achieve a proper evaluation of these models compared to the famous models. To achieve this goal, first, three machine learning methods, including DT, KNN, and LR were trained and tested on the dataset by the K-fold method. The grid search method extracts the most suitable hyperparameters of each model. These three models were combined with the soft-voting method to create an ensemble model. Then, using accuracy metrics, this model was analyzed. The flowchart of the research path is shown in Figure 2.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W1-2022-647-2023 | © Author(s) 2023. CC BY 4.0 License.

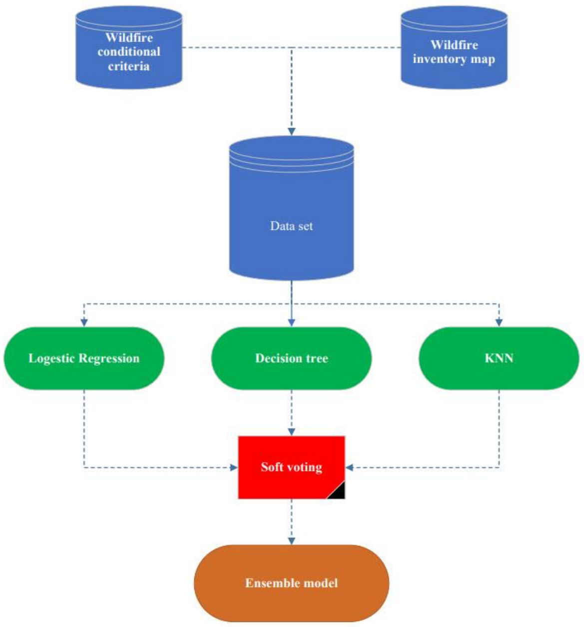

Figure 2. The flowchart signifies the introduced approach for forest fire susceptibility modelling and mapping.

3.1 Logistic regression (LR)

Logistic regression is one of the classification methods in supervised machine learning. Logistic regression is a famous two-class classification model. The LR is considered as one of the most widely used algorithms in FSM modeling (Adab, 2017; Mohajane et al., 2021; Sachdeva et al., 2018). When a LR model is trained, there are many algorithms to optimize its weight parameters. Optimization algorithms in regression logistics struggle to find the best weight that minimizes the cost function. The algorithm used for optimization in regression logistics is the Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm. The norm used for penalization is L2, which might limit the coefficients' size and push the estimated coefficients to zero. Regularizations are used to prevent overfitting in Machine Learning and Deep Learning models. ML models were built using the Scikit-Learn open-source Python library.

3.2 Decision tree (DT)

The decision tree algorithm is a non-parametric machine learning method used in regression and classification problems. This algorithm predicts the target variable value by learning the decision-making rules that infer from the characteristics of the data. The DT is considered one of the most widely used algorithms in FSM modeling (Gholamnia et al., 2020; Pham et al., 2020). Unlike other methods, this method does not require the normalization of data. One of the advantages of the decision tree is that it is readily interpretable. The decision tree is unstable because small changes in the data lead to the production of a completely different tree. The tree is prone to overfitting, so it is practical to avoid mechanisms such as setting the minimum number of samples required per node. The minimum number of samples per node was two by the grid search method. Also, the maximum feature was considered equal to the root of the features.

3.3 K nearest neighbour (KNN)

The KNN classification algorithm was first proposed by Cover and Hart in 1967 (Cover and Hart, 1967). K nearest neighbor algorithm is one of the easiest and fastest machine learning algorithms. The idea behind this algorithm is to calculate the distance between the new sample, and other samples. The distance function in this study is the Manhattan type, which is obtained through the grid search method from three kinds of Euclidean, Manhattan, and Minkowski distances. Then, K selects the nearest points of the training sample, and this value K can be any integer, in which the value K is obtained through the grid search method. Finally, it assigns the data point to the class to which most K data points belong. Therefore, the algorithm does not work well in problems where each sample contains high-dimensional features, making it difficult to calculate the distance in each dimension.

3.4 Ensemble model

Using three models, DT, KNN, and LR and designed our soft-voting ensemble classifier based on these three basic models. The advantages of voting are that, since voting relies on several models, incorrect classification of a model is not an obstacle. Also, the poor performance of a model can be offset by the strong performance of other models. The soft-voting ensemble classifier covers the weakness of individual base classifiers and outperforms the overall results by aggregating the multiple prediction models (Sherazi et al., 2021). In contrast to hard-voting, soft-voting gives better results and performance because it uses the averaging of probabilities (Saqlain et al., 2019). In soft-voting, the probability combination of each prediction in each model is considered, and the prediction is selected with the highest total probability. In soft-voting, the class labels are calculated based on the predicted probabilities p for classifiers (Sherazi et al., 2021):

$$y = \arg\max_{i} \sum_{j=1}^{m} w_j p_{ij} \tag{1}$$

where w_j is the weight that can be assigned to the j^{th} classifier. One of the limitations of voting in creating an ensemble model is that it treats all models similarly, meaning that all models are equally involved in the final prediction. This is the weakness because some models perform well in some situations, and in others, they perform poorly.

4. EXPERIMENTAL RESULTS

This study used a K-fold method to train and test the model. Usually, the value K is 5 or 10, and there is no general rule for determining the k-value. To analyse the performance of each fold, 5 accuracy metrics have been analyzed. 4 metrics parameters for each fold are visible in Table 2. Accuracy, recall, precision, F1 criteria were calculated from Equations 2-5.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

https://doi.org/10.5194/isprs-annals-X-4-W1-2022-647-2023 |

$$precision = \frac{TP}{TP + FP} \tag{4}$$

$$F1 = \frac{2 * precision * recall}{precision + recall} \tag{5}$$

where true positive (TP) is the number of forest fire points categorized correctly as forest fire and true negative (TN) denotes the number of non-forest fire points correctly classified as non-forest fire points. Meanwhile, false positive (FP) and false negative (FN) refer to the number of forest fire points incorrectly classified as a forest fire or non-forest fire points. The fifth parameter is the AUC criterion, which is the area under the ROC curve. The ROC curve is visible in Figure 3 with the AUC value. As is evident from Table 2, the average accuracy is 80.5%, which is an acceptable value and indicates that the model has left a good performance. The highest and lowest accuracy occurred in folds nine and five, respectively. In the most papers, the AUC criteria has been considered. The AUC indicates how well a model can distinguish between classes. The AUC value for all folds is above 80%, an average AUC value was 88%, and the highest AUC value in fold 9 was 93%, and the lowest AUC value in fold five was 80%.

Fold	Accuracy	Precision	Recall	F1
0	0.773	0.75	0.818	0.783
1	0.788	0.743	0.879	0.805
2	0.785	0.721	0.939	0.816
3	0.785	0.743	0.879	0.805
4	0.8	0.778	0.849	0.811
5	0.692	0.71	0.666	0.687
6	0.831	0.769	0.937	0.845
7	0.861	0.812	0.937	0.869
8	0.861	0.828	0.906	0.866
9	0.877	0.833	0.937	0.882
Mean	0.805	0.769	0.875	0.817

Table2. Accuracy metrics of all folds for the ensemble model.

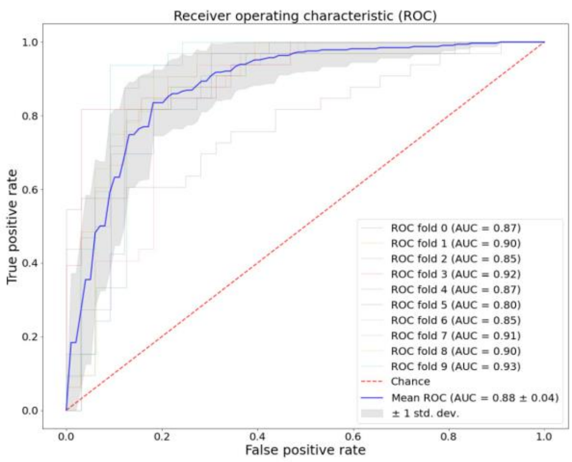

Figure 3. ROC curve and AUC value of each fold.

In this study, a fire susceptibility map has been produced with an ensemble model. This map is displayed in Figure 4. The model generates a number between 0 and 1 for each pixel according to its feature vector. Using a reclassification tool in the Spatial Analyst Tools ArcGIS 10.8 software, each final map cell is classified into five classes (very low, low, moderate, high, and very high) representing the forest fire hazard index, with the natural breaks method, all outcomes are divided into five classes. In Figure 6, the area of each class was calculated and displayed as a pie chart. As shown in Figure 5, the maximum area of the study area is 38% in the very low class. Only 10% of the area is located in very high class and according to the FSM most of the fires are located in areas that are in high and very high class, and this indicates that the model has well identified the fire-prone areas.

According to Figure 5, nearly 63 percent of the area is located in the low class, very low, and nearly 21 percent of the calculated area is located in the high, very high class. By providing this map and considering that most fires occur in 21% of the area, more fire extinguishing equipment, and aerial surveillance can be restricted.

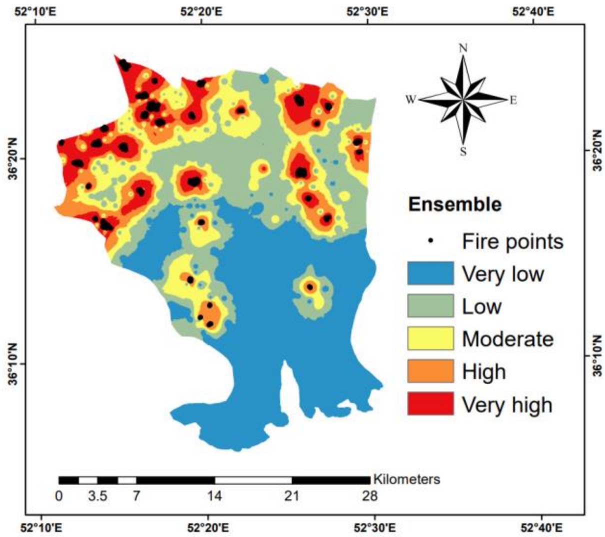

Figure 4. Fire susceptibility map from the ensemble model.

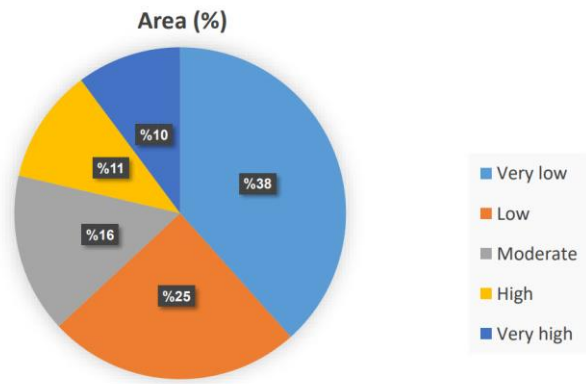

Figure 5. Percentages of the area in different susceptibility levels.

5. DISCUSSION

In this study, three models of KNN, DT, LR which are lighter and faster than the famous RF and SVM models, were selected to construct an ensemble model. This ensemble model was designed based on soft-voting on these three models. On the one hand, according to previous studies, 17 critical factors involved in the fire were identified (see Table 1). In the first step, all three models, i.e. KNN, DT, and LR, were optimized on the train data. Then, an ensemble model was generated using soft-voting method. In the last step, the model was investigated using 5 different accuracy metrics. As can be seen from Table 2, each 4-accuracy metrics evaluation has an average score above 75%. On the other hand, in fold 9, the highest Accuracy and AUC with 93% and 87% values were obtained. Using the final model, the map of fire-prone areas in 5 categories was prepared. Nearly 38% of the study area was classified in very low-risk areas and only 10% of the area of study area was classified as a high-risk area. Ghorbanzadeh et al. (2019) created a fire susceptibility map with the ANN, RF, and SVM models, with an AUC of 0.88 for the ANN model. According to the comparative study, it can be said that acceptable results have been achieved using this ensemble model.

6. CONCLUSION

This study proposed a soft-voting ensemble model used to identify fire-prone areas in the Amol study area. A total of five accuracy metrics were used to analyse this ensemble model, with all five metrics showing remarkable performance. The fire susceptibility map was created for the Amol region and classified into five categories. The susceptibility map created in this study could potentially pave the way for predicting the necessary measures and equipment to reduce the risk of fire in this area for makers and policymakers in this area. Considering that the model has been evaluated in a small area, to investigate the applicability of this model, it can be explored by localizing the factors of other regions according to the model and answering the question of whether the model can be generalized to larger scales and other different areas.

REFERENCES

Adab, H., 2017. Landfire hazard assessment in the Caspian Hyrcanian forest ecoregion with the long-term MODIS active fire data. Nat Hazards 87, 1807–1825. https://doi.org/10.1007/s11069-017-2850-2

Adab, H., Kanniah, K.D., Solaimani, K., Sallehuddin, R., 2015. Modelling static fire hazard in a semi-arid region using frequency analysis. Int. J. Wildland Fire 24, 763. https://doi.org/10.1071/WF13113

Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inform. Theory 13, 21–27. https://doi.org/10.1109/TIT.1967.1053964

Eskandari, S., Pourghasemi, H.R., Tiefenbacher, J.P., 2021. Fire-susceptibility mapping in the natural areas of Iran using new and ensemble data-mining models. Environ Sci Pollut Res 28, 47395–47406. https://doi.org/10.1007/s11356-021-13881-y

Gholamnia, K., Gudiyangada Nachappa, T., Ghorbanzadeh, O., Blaschke, T., 2020. Comparisons of Diverse Machine Learning Approaches for Wildfire Susceptibility Mapping. Symmetry 12, 604. https://doi.org/10.3390/sym12040604

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Aryal, J., 2019a. Forest Fire Susceptibility and Risk Mapping Using Social/Infrastructural Vulnerability and Environmental Variables. Fire 2, 50. https://doi.org/10.3390/fire2030050

Ghorbanzadeh, O., Valizadeh Kamran, K., Blaschke, T., Aryal, J., Naboureh, A., Einali, J., Bian, J., 2019. Spatial Prediction of Wildfire Susceptibility Using Field Survey GPS Data and Machine Learning Approaches. Fire 2, 43. https://doi.org/10.3390/fire2030043

H., Pazdar, S., 2021. Evaluating models and effective factors obtained from remote sensing (RS) and geographic information system(GIS) in the prediction of forest fire risk: A structured review. JGC 4. https://doi.org/10.24294/jgc.v4i2.618

Iban, M.C., Sekertekin, A., 2022. Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin provinces, Turkey. Ecological Informatics 69, 101647. https://doi.org/10.1016/j.ecoinf.2022.101647

Jaafari, A., Zenner, E.K., Panahi, M., Shahabi, H., 2019. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. Agricultural and Forest Meteorology 266–267, 198–207. https://doi.org/10.1016/j.agrformet.2018.12.015

Mohajane, M., Costache, R., Karimi, F., Bao Pham, Q., Essahlaoui, A., Nguyen, H., Laneve, G., Oudija, F., 2021. Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area. Ecological Indicators 129, 107869. https://doi.org/10.1016/j.ecolind.2021.107869

Naderpour, M., Rizeei, H.M., Ramezani, F., 2021. Forest Fire Risk Prediction: A Spatial Deep Neural Network-Based Framework. Remote Sensing 13, 2513. https://doi.org/10.3390/rs13132513

Papathoma-Köhle, M., Schlögl, M., Garlichs, C., Diakakis, M., Mavroulis, S., Fuchs, S., 2022. A wildfire vulnerability index for buildings. Sci Rep 12, 6378. https://doi.org/10.1038/s41598-022-10479-3

Pham, B.T., Jaafari, A., Avand, M., Al-Ansari, N., Dinh Du, T., Yen, H.P.H., Phong, T.V., Nguyen, D.H., Le, H.V., Mafi-Gholami, D., Prakash, I., Thi Thuy, H., Tuyen, T.T., 2020. Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction. Symmetry 12, 1022. https://doi.org/10.3390/sym12061022

Sachdeva, S., Bhatia, T., Verma, A.K., 2018. GIS-based evolutionary optimized Gradient Boosted Decision Trees for forest fire susceptibility mapping. Nat Hazards 92, 1399–1418. https://doi.org/10.1007/s11069-018-3256-5

Sharma, S.K., Aryal, J., Rajabifard, A., 2022. Remote Sensing and Meteorological Data Fusion in Predicting Bushfire Severity: A Case Study from Victoria, Australia. Remote Sensing 14, 1645. https://doi.org/10.3390/rs14071645

Tavakkoli Piralilou, S., Einali, G., Ghorbanzadeh, O., Nachappa, T.G., Gholamnia, K., Blaschke, T., Ghamisi, P., 2022. A Google Earth Engine Approach for Wildfire Susceptibility Prediction Fusion with Remote Sensing Data of Different Spatial Resolutions. Remote Sensing 14, 672. https://doi.org/10.3390/rs14030672

Saqlain, M., Jargalsaikhan, B., Lee, J.Y., 2019. A Voting Ensemble Classifier for Wafer Map Defect Patterns Identification in Semiconductor Manufacturing. IEEE Trans. Semicond. Manufact. 32, 171–182. https://doi.org/10.1109/TSM.2019.2904306

Sherazi, S.W.A., Bae, J.-W., Lee, J.Y., 2021. A soft voting ensemble classifier for early prediction and diagnosis of occurrences of major adverse cardiovascular events for STEMI and NSTEMI during 2-year follow-up in patients with acute coronary syndrome. PLoS ONE 16, e0249338. https://doi.org/10.1371/journal.pone.0249338

Karimi, A., Abdollahi, S., Ostad-Ali-Askari, K., P. Singh, V., Eslamian, S., Heidarian, A., Nekooei, M., Gholami, H., Pazdar, S., 2021. Evaluating models and effective factors obtained from remote sensing (RS) and geographic information system(GIS) in the prediction of forest fire risk: A structured review. JGC 4. https://doi.org/10.24294/jgc.v4i2.618

Pourghasemi, H.R., 2016. GIS-based forest fire susceptibility mapping in Iran: a comparison between evidential belief function and binary logistic regression models. Scandinavian Journal of Forest Research 31, 80–98. https://doi.org/10.1080/02827581.2015.1052750

Pradhan, B., Dini Hairi Bin Suliman, M., Arshad Bin Awang, M., 2007. Forest fire susceptibility and risk mapping using remote sensing and geographical information systems (GIS). Disaster Prevention and Management: An International Journal 16, 344–352. https://doi.org/10.1108/09653560710758297

Sayad, Y.O., Mousannif, H., Al Moatassime, H., 2019. Predictive modeling of wildfires: A new dataset and machine learning approach. Fire Safety Journal 104, 130–146. https://doi.org/10.1016/j.firesaf.2019.01.006

Tien Bui, D., Khosravi, K., Shahabi, H., Daggupati, P., Adamowski, J.F., M.Melesse, A., Thai Pham, B., Pourghasemi, H.R., Mahmoudi, M., Bahrami, S., Pradhan, B., Shirzadi, A., Chapi, K., Lee, S., 2019. Flood Spatial Modeling in Northern Iran Using Remote Sensing and GIS: A Comparison between Evidential Belief Functions and Its Ensemble with a Multivariate Logistic Regression Model. Remote Sensing 11, 1589. https://doi.org/10.3390/rs11131589

Watson, G.L., Telesca, D., Reid, C.E., Pfister, G.G., Jerrett, M., 2019. Machine learning models accurately predict ozone exposure during wildfire events. Environmental Pollution 254, 112792. https://doi.org/10.1016/j.envpol.2019.06.088