

LANDSLIDE SUSCEPTIBILITY MAPPING USING MACHINE LEARNING ALGORITHMS STUDY CASE AL HOCEIMA REGION, NORTHERN MOROCCO

Oussama HIMMY^a, Hassan RHINANE^a

^aGeosciences Laboratory, Faculty of Sciences Aïn Chock, Hassan II University, Casablanca, Morocco

KEYWORDS: Landslide susceptibility mapping, Machine Learning, GIS, Remote Sensing, Python, Random Forest, SVM, XGBoost, Decision Tree, Logistic Regression, Al Hoceima (Morocco).

ABSTRACT:

Landslides are one of the most dangerous natural disasters worldwide. Al Hoceima region, Part of the Moroccan mountain chain of the Rif is not an exception, since it's dominated by relatively young reliefs and marked by its dynamics compared to other regions. The main goal of this study is to assess the performance of Machine learning algorithms and identify the optimal method for the mapping of the area susceptible to landslides, in Al Hoceima city and its periphery. The current study aimed at evaluating the capabilities of six advanced machine learning algorithms including, Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), XGBoost (XGB) and Logistic Regression (LR). A total of 114 landslides were mapped from various sources. 70% of this database was used for model building and 30% for validation. Ten landslide factors are selected to detect the most sensitive areas: altitude, slope, aspect, distance to faults, distance from roads, lithology, curvature, plan curvature, profile curvature, and vegetation index (NDVI). The outcome of the landslide susceptibility analysis was verified using receiver operating characteristics (ROC) curves and precision-recall curves (PRC), acknowledging XGBoost and Random Forest as the optimal methods with AUCROC (96% and 95.5% accordingly), a severe imbalanced classification was detected by the PRC, solved by undersampling the majority class, to obtain major improvement in models performance according to AUCPRC (from 40% to 87% for XGBoost) and slight decrease according to AUCROC (from 96% to 94% for XGBoost). The outcome of this study and the landslide susceptibility maps would be useful for environmental, economic, and social protection and to help formulate suggestions for optimizing landslide risk assessment in areas exposed to this phenomenon.

1. INTRODUCTION

Landslides are the most common geological disasters, causing loss of human life and damage to the economy (Tien Bui et al., 2012) It occurs when natural or man-made slopes become unstable due to geological, hydrological, and geomorphological conditions, heavy rainfall, seismic movements, volcanic eruptions, and human activities leading to destabilization of slopes (Soeters and van Westen 1996).

Landslides susceptibility is the likelihood of a landslide occurring in an area based on local terrain conditions (Clark et al., 1984). It predicts "where" landslides are likely to occur.

Landslide susceptibility maps rely mostly on the amount and quality of available data, the working scale, and the selection of the appropriate methodology of analysis and modeling. The process of creating these maps involves several qualitative or quantitative approaches, qualitative methods depend on expert criteria, simply by using landslide inventories to identify sites of similar geological and geomorphologic properties susceptible to failure. Some quantitative methods involve the idea of ranking and weighting and may evolve to be semi-quantitative, for example, the use of the Analytic Hierarchy Process (AHP) (Saaty, 1980) by Barredol et al., (2000), and Weighted Linear Combination (WLC). However qualitative or semi-quantitative methods are most useful on a regional scale.

Quantitative methods are based on numerical expressions of the relationship between controlling parameters and landslides. Two types are distinguished, deterministic and statistical, Deterministic quantitative methods depend on the engineering of slope instability expressed in terms of the parameter of safety (mathematical relationships between resisting and driving forces). Statistical methods analyze the link between landslide-

controlling parameters and the distribution of landslide sites (Guzzetti et al., 1999). In the literature, various methods have been carried out to assess landslide susceptibility with the aid of GIS and Remote Sensing (key points for landslide susceptibility mapping).

The main advantages of Machine learning methods include their objective statistical basis, reproducibility, ability to quantitatively analyze the contribution of factors to landslide development, and their potential for continuous updating (Youssef and Pourghasemi, 2021).

In Morocco, landslides are a recurrent problem throughout the Rif Mountains because of the dynamics related to their formation (Alpine tectonics). This region is frequently subjected to heavy precipitation accompanied by instabilities related to tectonic movements. A recent study was conducted in this region (precisely the municipality of Oudka) using machine learning methods (Logistic Regression and Artificial Neural Network) by (Benchelha et al., 2019), which led to interesting results and assured the importance of machine learning with GIS and Remote Sensing in landslide susceptibility mapping.

In this study, as continuity of other recent studies, this time in Al Hoceima region, six advanced machine learning algorithms including SVM, Random Forest, XGBoost, Logistic Regression, Naive Bayes, and Decision Trees, were adopted to construct landslides susceptibility maps with the contribution of GIS (Geographic Information System) and Remote Sensing for spatial data management and models building. these models were utilized for several reasons, mainly because they showed good results in various studies (Ayalew and Yamagishi, 2005, 2005; Can et al., 2021; Karakas et al., 2020; Lee et al., 2017; Yeon et al., 2010; Youssef and Pourghasemi, 2021).

2. STUDY AREA

Al Hoceima region is located in the eastern part of the Bokkoya massif in northern Morocco. It covers an area of 42 km². It is framed by the parallels 35.16° and 35.28° N and the meridians 3.87° and 4.05°W. Limited to the North and the East by the Mediterranean Sea (the Alboran Sea and Al Hoceima Bay), it ends in the west with the point of Boussicour and the south by Oued Isli. It is a mountainous region, made up of steep valleys, separated by modest reliefs with strong slopes.

Al Hoceima region as part of the Rif mountains are areas that have been frequently affected by landslides in recent years. This vulnerability is in perpetuity increases due to mountainous topographical features, subjection to heavy precipitation, the intensification of land use, and socio-economic development. The rate of increase has relatively increased during the years 2004-2014, it is estimated at 0.25% in the study area, compared to 0.1% at the national level (HCP, 2014). This is due to internal emigration from the rural communes bordering the city of Al Hoceima. Most of these have built their houses in areas not suitable for construction, near the courtyards, and in or on unstable areas.

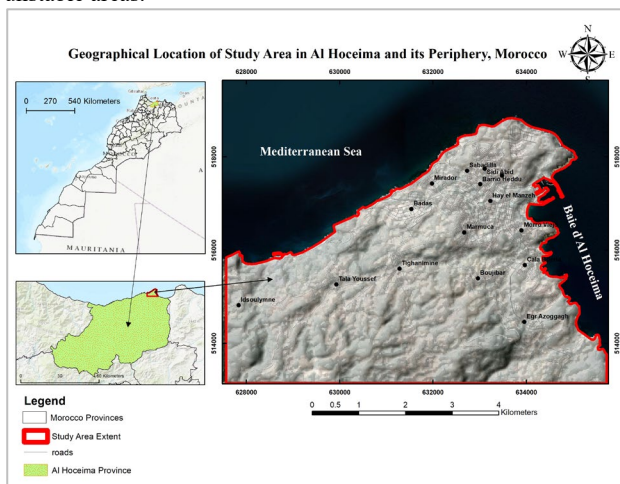


Figure 1. The geographical Location of the study area in Al Hoceima city and its periphery, Morocco

Geologically, the area of interest is constituted by a stack of four layers put in place since the end of the Oligocene, separated from each other by anomalous contacts. The arrangement of these nappes is complex; including Triassic dolomites and flinty limestones which constitute the nappe of Boussicour and surmount the eo-Oligocene nappe formed mainly by marl materials.

The study area is characterized by a Mediterranean, semi-arid climate, marked by winters, temperate winters, and hot summers. The average annual precipitation reaches 385mm, the rainfall characterized by irregularity and brutality, aggravating the action of said rains on the soil. The temperature is influenced by the proximity of the Mediterranean coast, which attenuates the amplitude temperature range, the minimum average of the winter month being mild, it reaches 10°, while the average maximum temperature of the summer month is moderate, reaching 29°.

3. MATERIAL AND METHODS

3.1 Data Preparation

Landslide inventory was obtained based on the previous works of (Byou et al., 2020), for the entire city of Al Hoceima and its

periphery. These events are represented by polygons (114 events as illustrated in Figure 2) based on the landslide inventory and the interpretation of Google Earth satellite images, as well as on GPS mobile mapping (Mobile Mapper) to locate landslides and determine their boundaries. These were delineated into areas of detachment, but some locations, where there is evidence of ground instability, such as tension cracks or leaning trees, were included in the mapping.

After reviewing 12 scientific papers published between 2017 and 2021 in different journals (Ayalew and Yamagishi, 2005; Benchelha et al., 2019; Byou et al., 2020; Can et al., 2021; Karakas et al., 2020; Lee et al., 2017; Muñoz et al., 2020; Onagh et al., 2012; Sevgen et al., 2019; Tien Bui et al., 2012; Yordanov and Brovelli, 2020; Youssef and Pourghasemi, 2021), 10 conditioning factors were selected in terms of disponibility and their higher influence on Landslides occurrence can be grouped into 3 groups: geological factor (including lithology, distance-to-fault), topographical factors (including elevation, slope, aspect, curvature, profile curvature, plan curvature), anthropogenic (including distance to roads) and land used (Normalized Difference Vegetation Index NDVI). The processing of landslide conditioning factors was done by a spatial analysis tool (ArcGIS software) with a common pixel size of 30 m and a common spatial reference (Merchich North Morocco) and the Mediterranean Sea was excluded from the calculation.

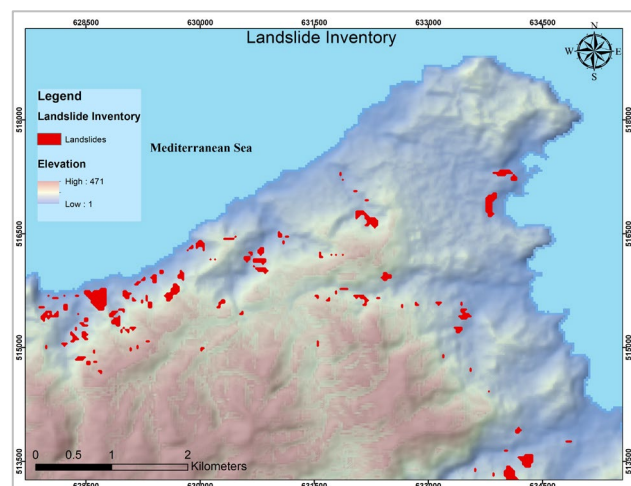


Figure 2. Distribution of landslides occurrence in Al Hoceima and its periphery

The geological parameters, extracted from the Al Hoceima geological map of the study area at 1: 50,000 scale (Edition 1984), such as i) the faults were digitized and distance to faults was calculated using the ArcGIS software (Euclidean distance). ii) The lithology was digitized and split into seven unique classes (figure 3 a). The topographical parameters, derived from the DEM model (ASTER GDEM 30m resolution), such as Elevation, are controlled by various geological, geomorphological, and meteorological factors including, lithological units, weathering, wind action, and precipitation, whose value varies between 1 and 471 m (figure 3 c) Slope is one of the key factors of slope stability ranging from 0 to 59° (figure 3 d). Aspect can be considered as the slope direction, and the values will be the compass direction divided into nine classes: Flat, North, North-East, East, South-East, South, South-West, West, and North-West. Curvature, profile curvature, and plan curvature are also among the commonly used conditioning factors. They influence the material deposition by managing the acceleration or deceleration of these

materials on the slope. The distance to roads parameter reflects the influence of human activities, extracted by calculating the Euclidean distance based on roads imported from the OpenStreetMap database. The land use parameter used is the standardized difference vegetation index NDVI (figure 3 i), which represents the coverage and vigor of forests and woodlands as crops have been harvested and herbaceous vegetation has become withered in late autumn. In general, vegetation, especially trees, can help soil hold water content and improve its mechanical properties through root systems that stabilize slopes. Thus, landslides may arise more likely in unvegetated areas rather than in forests and woodlands. NDVI was derived from Sentinel 2A image (15/10/2021) calculated based on the Formula (1):

$$NDVI = \frac{NIR-R}{NIR+R} \quad (1)$$

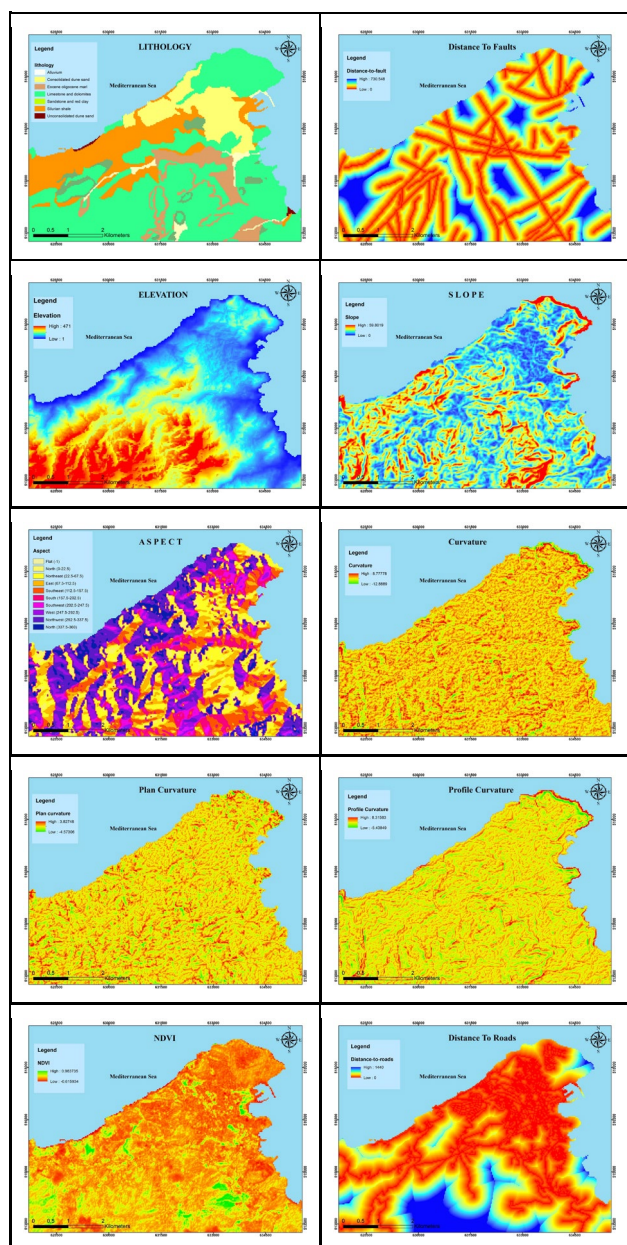


Figure 3. Maps of some landslide conditioning factors: (a) lithology; (b) distance to faults; (c) elevation; (d) slope; (e) aspect; (f) curvature; (g) plan curvature; (h) profile curvature; (i) NDVI; (j) distance to roads

3.2 Modeling approach using machine learning algorithms

Figure 4 presents a diagram illustrating the methodology used for the application and verification of landslide sensitivity models, after making sure that raster data of parameters were overlaid pixel to pixel, the next step is to convert the raster into vector points in which each point has the value of the according to pixel, followed by applying spatial join using Python environment in ArcGIS software to merge all parameters into one single table and then to Excel sheet table. The obtained table is composed of 34084 rows and 11 columns (10 parameters and landslides label). The Excel table was transformed into a Jupyter notebook, to explore the correlation between factors, clean missing values, and divide the dataset into two sets, 70% for training and 30 % for validating and testing. A count of 657 rows represents the positive classes (1.9% of the dataset) where landslides occurred, and 32767 rows represent the negative class (96.1% of the dataset), this indicates a case of imbalanced classification which involve a negative event with most examples and a positive event with a minority of examples.

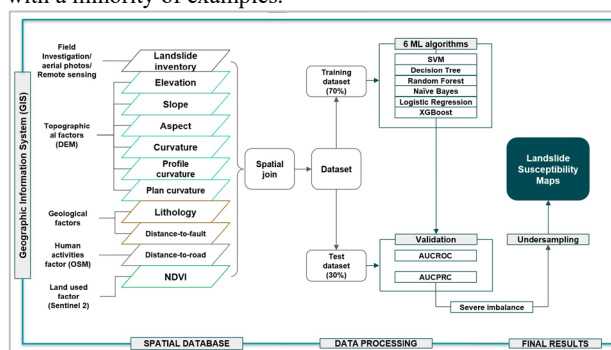


Figure 4. Workflow diagram showing the methodology adopted

In the current study, five advanced machine learning algorithms that vary in their degree of complexity were applied to evaluate their efficacy in landslide susceptibility mapping, including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), XGBoost (XGB) and Logistic Regression (LR) imported from Scikit-learn library in Python environment.

3.2.1 SVM: It is a supervised learning algorithm, derived from statistical learning theory and structural risk minimization principle (Lee et al., 2017; Vapnik, 1995), it deals with the binary classification models. The SVM algorithm uses the training data to generate a separating hyper-plane in the initial space of coordinates between two distinct categories and maximize the margin between them. to make the point more linearly separable, it uses kernel functions to map the initial input space into a high-dimensional feature space, so choosing the adequate kernel function to optimize SVM modeling,

3.2.2 Decision Trees: The decision Tree algorithm is widely used for classification tasks, C4.5, the tree growth begins from a node, which is then split by selecting the attribute that best classifies a set of examples based on an attribute selection measure.

3.2.3 RF: The Random Forest is an ensemble learning method generating many classification trees (Breiman, 2001; Sevgen et al., 2019). The ‘forest’ generated by the random forest algorithm is trained through bagging and bootstrap aggregating which makes RF resistant to overtraining and overfitting (independent trees)

3.2.4 NB: Naive Bayes is a classification method based on the Bayes theorem, it assumes that all factors are independent given the output class, which is called the conditional independence assumption (Soria et al., 2008), the most advantages of the NB technique are including it is robust to noise and irrelevant variables, very easy to apply, and does not need complicated iterative schemes.

3.2.5 XGB: The XGBoost method was used as the supervised classification model. The method originated from the gradient tree boosting algorithm, which is an effective ML method. The main idea of a boosting algorithm is to combine weak learners outputs sequentially to achieve better performance. It uses regularized boosting technique to reduce overfitting and thus increase the model accuracy.

3.2.6 LR: The logistic regression (LR) model is a mathematical method to establish the relationship between independent factors and landslides (Bai et al. 2010; Das et al. 2010; Nandi and Shakoor 2010). It's useful for predicting the presence or absence of a characteristic or outcome based on values of a set of predictors variables. The predicted values range from 0 to 1.

4. RESULTS AND DISCUSSION

4.1 Landslide susceptibility mapping

Before modeling with each algorithm, tuning the hyperparameter was applied using GridSearchCV implemented in Scikit-learn, which gives the best possible hyperparameters to be used. Each ML model was transformed into a map of landslide susceptibility based on predicting the probability of landslide occurrence in each pixel, followed by reclassification into five susceptible zones (very low, low, medium, high, and very high), using the natural breaks (Jenks) classification method (Panahi et al., 2020; Pourghasemi and Kerle, 2016; Tien Bui et al., 2012).

4.2 Accuracy assessment and comparison

4.2.1 Receiver Operating Characteristics (ROC) curves: The accuracy of these landslide susceptibility maps was evaluated to assess the model performance by calculating the relative operating characteristic (ROC), which is widely used in various applications. The problem adopted in this paper reflects a binary classification where the instances are distributed into two main classes – positives and negatives. The outcome of each applied classifier can be summed into a confusion matrix, representing the performance of the algorithm.

The area under the ROC curve (AUC) represents the quality of the probabilistic model (its ability to predict the occurrence or not of an event). According to the AUC method, the variation in model performance among the MLTs was relatively high, the XGBoost (AUCROC = 96%) and the RF (AUCROC = 95.4%) had the highest performances. followed by SVM (AUCROC = 92.5%), Decision Tree (AUCROC = 92.2%), Naive Bayes (AUCROC = 91.9%) and finally Logistic Regression with the lowest performance (AUCROC = 89.8%).

All machine learning techniques used exposed sufficient performance (AUC > 70%). In this study, all landslide sensitivity models were validated using the false positive rate (specificity) and true positive rate (sensitivity) using the validation dataset independent of that used in the landslide model construction process. The ROC curve of this study is illustrated in figure 5.

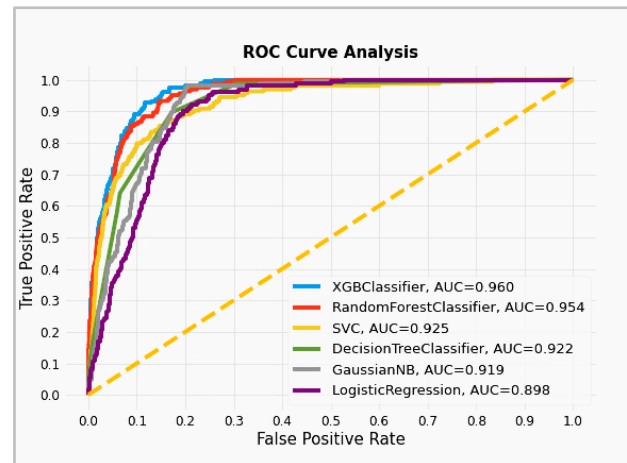


Figure 5. ROC Curve Analysis with AUC score comparison between used ML algorithms

4.2.2 Precision-Recall Curve (PRC): The Precision-Recall curve is another metric that relies on the confusion matrix and the sensitivity (Recall). The difference is that the PRC is a quantifier of the positive classes which in most classifications represent a minority. That makes it an effective diagnostic for imbalanced binary classification models.

To use PRC, the area under the curve can be computed and an area of 1 is considered a perfect classifier, in the case of AUCROC, the threshold for a classifier is 0.5, however, in the case of AUCPRC the threshold is computed based on the ration between the predicted positives and negatives (Eq2)

$$PRC_{THRESHOLD} = \frac{P}{P+N} \quad (2)$$

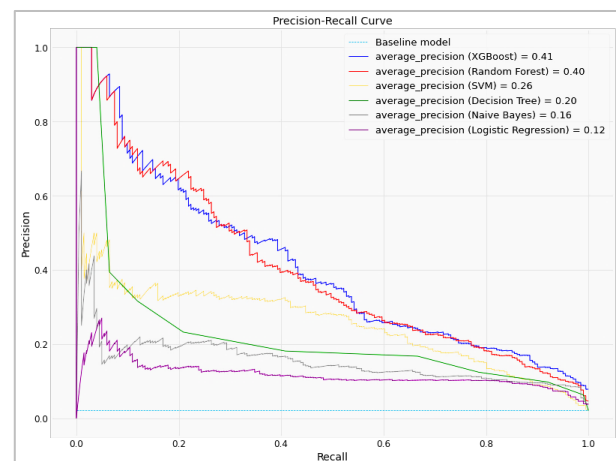


Figure 6. Precision-Recall Curve Analysis with average precision score comparison between used ML algorithms

The results obtained from the average precision and the curve are not promising as obtained by AUCROC, the baseline is close to 0 (0.02), so while ROC AUC showed optimistic results, the PRC AUC on the other hand showed the reality of a severe imbalanced classification with few samples of the minority class as shown in figure 6.

4.3 Undersampling with the majority class

A simple approach to using standard machine learning algorithms on an imbalanced dataset is to change the training dataset to have a more balanced class distribution. This can be achieved by eliminating examples from the majority class, referred to as

"undersampling." A possible downside is that examples from the majority class that are helpful during modeling may be deleted. In this study "RandomUnderSampler class" was used to change the class distribution to a less severe 1:2. Giving 1314 samples for the majority class and the usual 657 samples for the minority class.

The results obtained after undersampling for AUCROC show a very low decrease for most classifiers, as shown in figure 7, with always acknowledging Random Forest and XGBoost as the best models.

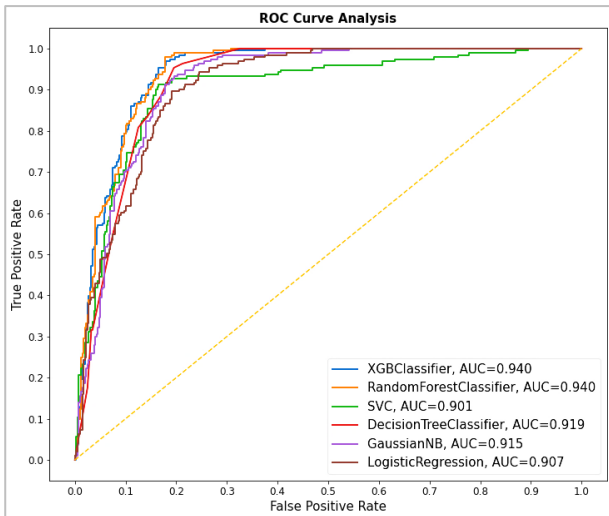


Figure 7. ROC Curve Analysis with AUC score comparison between used ML algorithms after Undersampling

While for AUCPRC, the performance shows a great improvement, results may vary given the stochastic nature of the algorithm or evaluation procedure or differences in numerical precision. By running a few times and comparing the average outcome, the outcome is as follows: The XGBoost (AUCPRC = 87%) and the RF (AUCPRC = 85%) had the highest performances, followed by SVM (AUCPRC = 79%), Decision Tree (AUCPRC = 77%), Naive Bayes (AUCPRC = 80%) and finally Logistic Regression with the lowest performance (AUCPRC = 72%).

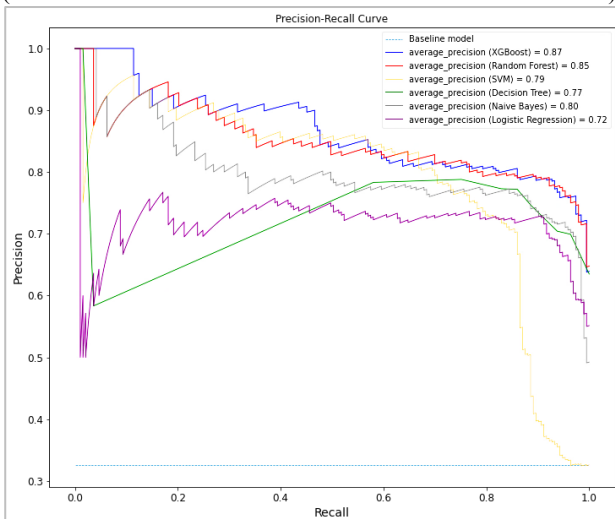


Figure 8. Precision-Recall Curve Analysis with average precision score comparison between used ML algorithms

The results obtained for landslide susceptibility for each model after undersampling look more promising than before, giving more dispersed probability prediction across cells, which makes

it visually possible for an equal interval reclassification as follows: Very low ($0 < LSI \leq 0.2$), Low ($0.2 < LSI \leq 0.4$), medium ($0.4 < LSI \leq 0.6$), high ($0.6 < LSI \leq 0.8$) and very high ($LSI > 0.8$).

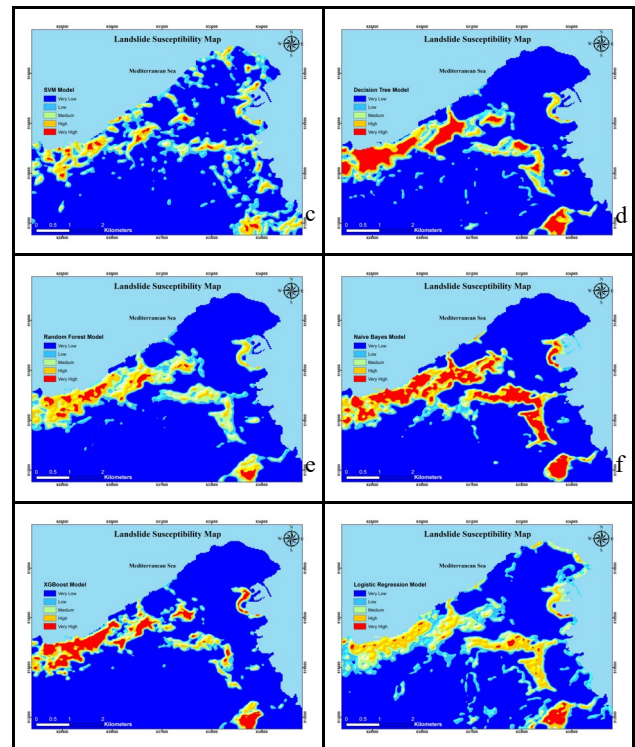


Figure 9. Landslide susceptibility map produced from Machine Learning algorithms after Undersampling: (a) SVM; (b) Decision Tree; (c) Random Forest; (d) Naïve Bayes; (e) XGBoost; (f) Logistic Regression

As a review of the resulting maps of Landslides susceptibility, the quality of different machine learning algorithms is quite promising, and despite the difference in the spatial distribution of the LS probability categories, all models agree on the same spots for high and very high categories for LS.

The percentage of each class's area was computed, and it shows that more than 65 % is designed for the very low class for every model, while the high and very high category, they represent less than 10 % of the whole study area.

5. CONCLUSION

To obtain a good landslide susceptibility model, the selection of factors of these phenomena was an important step in the development of Machine Learning models. Prior knowledge of these factors, responsible for landslide initiation, is necessary for mapping the area of the landslide susceptibility to landslides.

In this study, Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), XGBoost (XGB), and Logistic Regression (LR) were used to analyze landslide susceptibility and create The landslide susceptible maps, generated using MLTs, could be essentially applied by the decision-makers, planners, and engineers, to identify landslide-prone areas, to prevent and mitigate the landslide risks areas, determine the suitable land-use planning areas, and establish early warning systems.

Based on the results of mentioned models, XGBoost and Random Forest showed the highest predictive power for the study area, whether in AUCPRC or AUCROC.

The study showed the importance of choosing other metrics for validation such as PRC, instead of only relying on ROC

assessment, which gives more insights into model performance. The study resolved the problem of severe imbalance by undersampling the majority class (non-landslides), which considerably improved the performance of models as confirmed by the PRC metric and the resulting maps.

In any landslide susceptibility analysis, a level of susceptibility is assumed that an active landslide will occur. If only areas of high to very high susceptibility were at risk, and almost the majority of used models agree on these spots especially XGBoost and Random Forest

The results from this study demonstrate the benefit of applying the optimal MLT with proper accuracy and lesser learning time in landslide susceptibility assessment.

REFERENCES

- Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65, 15–31. <https://doi.org/10.1016/j.geomorph.2004.06.010>
- Benchelha, S., Chennaoui Aoudjehane, H., Hakdaoui, M., El Hamdouni, R., Mansouri, H., Benchelha, T., Layelmam, M., Alaoui, M., 2019. LANDSLIDE SUSCEPTIBILITY MAPPING IN THE MUNICIPALITY OF OUDKA, NORTHERN MOROCCO: A COMPARISON BETWEEN LOGISTIC REGRESSION AND ARTIFICIAL NEURAL NETWORKS MODELS. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-4/W12, 41–49. <https://doi.org/10.5194/isprs-archives-XLII-4-W12-41-2019>
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Byou, T., Obda, K., Taous, A., Obda, I., 2020. Susceptibilité aux glissements de terrain dans la ville d'Al Hoceima et sa périphérie : application de la méthode de la théorie de l'évidence. *Geomatica*. <https://doi.org/10.1139/geomat-2019-0025>
- Can, R., Kocaman, S., Gokceoglu, C., 2021. A Comprehensive Assessment of XGBoost Algorithm for Landslide Susceptibility Mapping in the Upper Basin of Ataturk Dam, Turkey. *Applied Sciences* 11, 4993. <https://doi.org/10.3390/app11114993>
- Clark, J.C., Brabb, E.E., Greene, H.G., Ross, D.C., 1984. *Geology of Point Reyes Peninsula and Implications for San Gregorio Fault History* 67–85.
- Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* 31, 181–216. [https://doi.org/10.1016/S0169-555X\(99\)00078-1](https://doi.org/10.1016/S0169-555X(99)00078-1)
- Karakas, G., Can, R., Kocaman, S., Nefeslioglu, H.A., Gokceoglu, C., 2020. LANDSLIDE SUSCEPTIBILITY MAPPING WITH RANDOM FOREST MODEL FOR ORDU, TURKEY. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLIII-B3-2020, 1229–1236. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1229-2020>
- Lee, S., Hong, S.-M., Jung, H.-S., 2017. A Support Vector Machine for Landslide Susceptibility Mapping in Gangwon Province, Korea. *Sustainability* 9, 48. <https://doi.org/10.3390/su9010048>
- Muñoz, E.N., Sánchez, O.D., Hernandez, F.L., 2020. METHODOLOGY FOR ESTIMATING LANDSLIDES SUSCEPTIBILITY USING ARTIFICIAL NEURAL NETWORKS, in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Presented at the ISPRS WG III/10-2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS 2020) (Volume XLII-3/W12-2020) - 22–26 March 2020, Santiago, Chile, Copernicus GmbH, pp. 533–538. <https://doi.org/10.5194/isprs-archives-XLII-3-W12-2020-533-2020>
- One, M., Kumra, V.K., Rai, P.K., 2012. APPLICATION OF MULTIPLE LINEAR REGRESSION MODEL IN LANDSLIDE SUSCEPTIBILITY ZONATION MAPPING THE CASE STUDY NARMAB BASIN IRAN. *International Journal of Geology* 2, 15.
- Panahi, M., Sadhasivam, N., Pourghasemi, H.R., Rezaie, F., Lee, S., 2020. Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR). *Journal of Hydrology* 588, 125033. <https://doi.org/10.1016/j.jhydrol.2020.125033>
- Pourghasemi, H.R., Kerle, N., 2016. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environ Earth Sci* 75, 185. <https://doi.org/10.1007/s12665-015-4950-1>
- Sevgen, E., Kocaman, S., Nefeslioglu, H.A., Gokceoglu, C., 2019. A Novel Performance Assessment Approach Using Photogrammetric Techniques for Landslide Susceptibility Mapping with Logistic Regression, ANN and Random Forest. *Sensors* 19, 3940. <https://doi.org/10.3390/s19183940>
- Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., 2012. Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, and Naïve Bayes Models. *Mathematical Problems in Engineering* 2012, e974638. <https://doi.org/10.1155/2012/974638>
- Vapnik, V.N., 1995. Conclusion: What is Important in Learning Theory?, in: Vapnik, V.N. (Ed.), *The Nature of Statistical Learning Theory*. Springer, New York, NY, pp. 167–175. https://doi.org/10.1007/978-1-4757-2440-0_7
- Yordanov, V., Brovelli, M.A., 2020. COMPARING MODEL PERFORMANCE METRICS FOR LANDSLIDE SUSCEPTIBILITY MAPPING, in: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Presented at the XXIV ISPRS Congress, Commission III (Volume XLIII-B3-2020) - 2020 edition, Copernicus GmbH, pp. 1277–1284. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1277-2020>
- Youssef, A.M., Pourghasemi, H.R., 2021. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geoscience Frontiers* 12, 639–655. <https://doi.org/10.1016/j.gsf.2020.05.010>