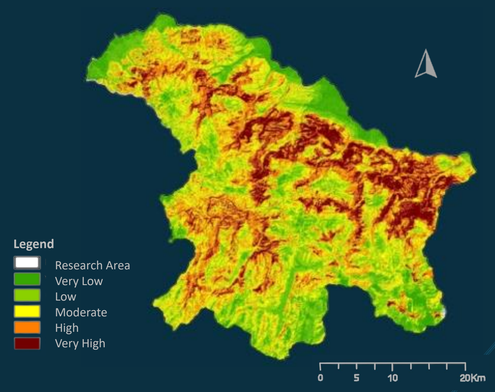
Remote Sensing & Hazards / Integrating Digital Technologies and eXplainable AI (XAI) in Natural Hazard and Disaster Management

Integrating XBoost and SHAP for Enhanced Interpretability in Landslide Susceptibility Assessment: A Case Study in North-western Peloponnese, Greece.

Maria Sotiria Frousiou[1], Ioanna Ilia[2], Dimitrios Kasmas[1], and Ioanna Petropoulou[1]

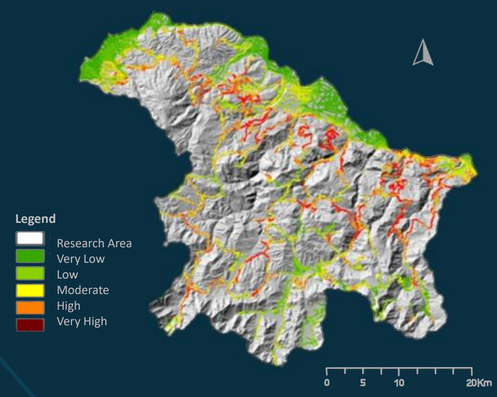
[1]Undergraduate Students at National Technical University of Athens, School of Mining and Metallurgical Engineering, marisafrousiou@gmail.com , ivanapetropoulou@gmail.com, dimkasmas@gmail.com
[2]Laboratory of Engineering Geology and Hydrogeology, Department of Geological Sciences, School of Mining and Metallurgical Engineering, National Technical University of Athens, 15780 Zografou, Greece; gilia@metal.ntua.gr



Landslide Susceptibility Map for North-western Peloponnese

Landslide Susceptibillity Parameters

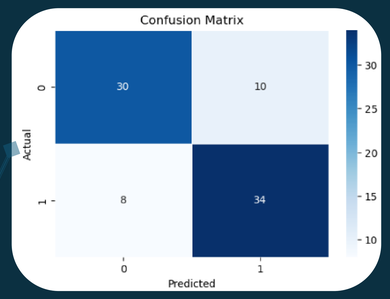
Landslide related parameters have incorporated a wide range of attributes into landslide susceptibility assessment, such as the research area's elevation, slope, plan (Plan Curvature) and prof (Profile Curvature), while also having a high SHAP value for each of them. Some other just as important parameters are the research area's geological and hydrogeological properties including, but differentiating the importance value of each, existing near-by faults and rivers that provide the model with a more detailed background of the influence of each parameter on each location of the database.

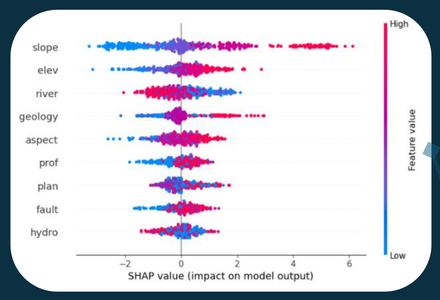


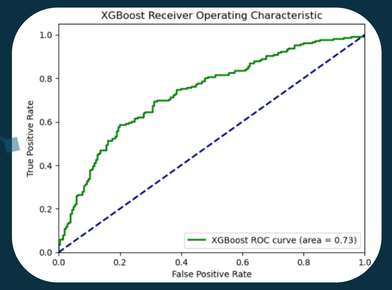
Landslide Susceptibility Map for North-western Peloponnese road network

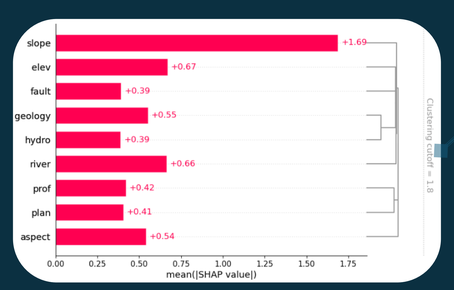
XGBoost

Extreme Gradient Boost (XGBoost) is a gradient-boosted decision tree machine learning library and is used for regression, classification and ranking problems while providing parallel tree boosting. XGBoost builds upon Gradient Boosting, which is a powerful technique for building predictive models and ensemble learning that combines multiple algorithms for the model's optimization. Code-wise, XGBoost has been integrated with a variety of tools and packages such as scikit-learn for python.





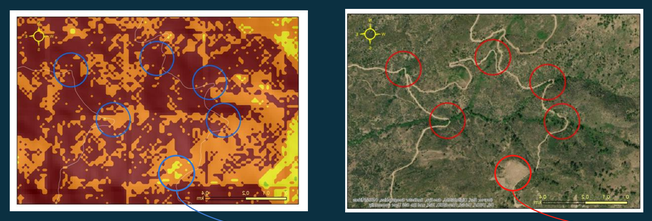


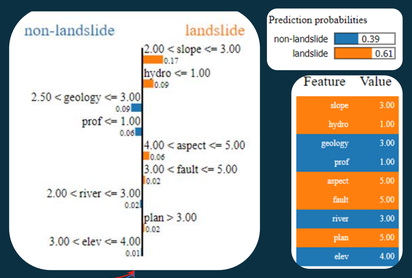


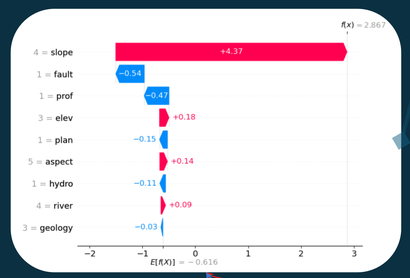
SHAP

SHapley Additive exPlanations (SHAP) is a Python Package, based on Shapley values, a game theoretic approach that calculates each player's contribution to the final outcome. SHAP provides a consistent and objective explanation of the impact of each feature to the model's prediction by assigning an importance value to each feature that represents its contribution.

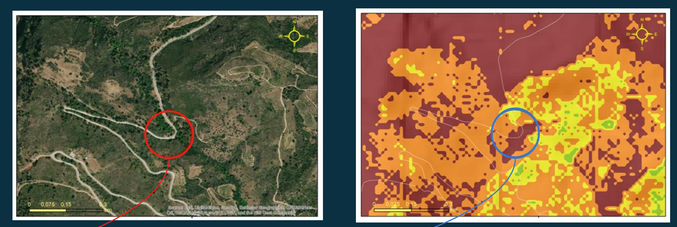
Landslide susceptibility assessment of specific location no. 104 of database







Assessment of landslide susceptibility parameters' importance value for specific location no. 131 of database



Conclusions

Landslide susceptibility assessment is based on classification and ranking, making the use of XGBoost an excellent choice for such problems. It is open-source, constantly improving and increasing its reliability while it can handle large datasets with many features. SHAP is used to examine the validity and visualise the given result from XGBoost's predictions, while SHAP's additivity enables the independent computation of each feature's contribution, providing analytical results and unique insights. The integrated approach overcomes the obstacles of higher error margin and slow return of results, while SHAP minimizes the black-box characteristics of non-ML approaches by interpreting the data that would otherwise be unknown. More specifically for the given research material, the confusion matrix XGBoost analysis gave back values of Accuracy: 0.78, Sensitivity: 0.81 and Specificity: 0.75, which proves that the combinational method for XAI has high predictive power whilst keeping the explainability at high values as well. The outcome maps on GIS provide the visualization of the model's results and so indicate the locations that are most likely to form landslide phenomena and thus makes a great predictability and prevention tool.

References. Zhang, Q; et.al Hybrid Integration of Bagging and Decision Tree Algorithms for Landslide Susceptibility Mapping || XGBoost_NVIDIA || A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning || An Introduction to SHAP Values and Machine Learning Interpretability