

Title: Flood mapping based on novel ensemble modeling involving the deep learning, Harris Hawk optimization algorithm and stacking based machine learning

Author: Data Science, Flooding, Learning algorithms, Machine Learning, Predictive markers, Statistical Learning

Abstract: Among the various natural disasters that take place around the world, flood is considered to be the most extensive. There have been several floods in Buzău river basin, and as a result of this, the area has been chosen as the study area. For the purpose of this research, we applied deep learning and machine learning benchmarks in order to prepare flood potential maps at the basin scale. In this regard 12 flood predictors, 205 flood and 205 non-flood locations were used as input data into the following 3 complex models: Deep Learning Neural Network-Harris Hawk Optimization-Index of Entropy (DLNN-HHO-IOE), Multilayer Perceptron-Harris Hawk Optimization-Index of Entropy (MLP-HHO-IOE) and Stacking ensemble-Harris Hawk Optimization-Index of Entropy (Stacking-HHO-IOE). The flood sample was divided into training (70%) and validating (30%) sample, meanwhile the prediction ability of flood conditioning factors was tested through the Correlation-based Feature Selection method. ROC Curve and statistical metrics were involved in the results validation. The modeling process through the stated algorithms showed that the most important flood predictors are represented by: slope (importance $\approx 20\%$), distance from river (importance $\approx 17.5\%$), land use (importance $\approx 12\%$) and TPI (importance $\approx 10\%$). The importance values were used to compute the flood susceptibility, while Natural Breaks method was used to classify the results. The high and very high flood susceptibility is spread on approximately 35–40% of the study zone. The ROC Curve, in terms of Success, Rate shows that the highest performance was achieved FPIDLNN-HHO-IOE (AUC = 0.97), followed by FPIStacking-HHO-IOE (AUC = 0.966) and FPIMLP-HHO-IOE (AUC = 0.953), while the Prediction Rate indicates the FPIStacking-HHO-IOE as being the most performant model with an AUC of 0.977, followed by FPIDLNN-HHO-IOE (AUC = 0.97) and FPIMLP-HHO-IOE (AUC = 0.924).

Keywords: Flood susceptibility, Flood zonation mapping, Trotuș River basin, Geographic Information System (GIS), Machine learning, Decision tree models, Forest-PA-WOE

DOI: <https://doi.org/10.1007/s13201-024-02131-4>

<https://link.springer.com/article/10.1007/s13201-024-02131-4>