

Using machine learning methods in order to classify watershed land use type

Semih Ediş¹, Efehan Ulaş^{2,*}

¹ Cankiri Karatekin University, Faculty of Forestry, Department of Forest Engineering, Çankırı, Turkey

² Cankiri Karatekin University, Faculty of Science, Department of Statistics, Çankırı, Turkey

* Corresponding author: ef_ulas@hotmail.com

Abstract: Machine learning is a system that studies the construction of algorithms that can be learned as structural functions and can be predicted from data. Such algorithms work by constructing a model to make database estimates and decisions from sample entries rather than alphabetically tracking static program instructions. There are many research has gone into the application of neural network classifiers to land use classification. However, none of them focus on Bayesian view. In our study, we included the research area (Tatlıçay and Acıçay) is located in the Kizilirmak basin of the north of the central district of Çankırı. The central district includes some of the Yapraklı and Korgun districts. The location of the area is between 40 ° 52 '11 " - 40 32' 58" northern latitudes and 33 ° 18 '19 " - 34 ° 3' 43" east longitudes. 15 different parameters were determined that can affect the land use type by taking expert opinions. The data collection has been made based on those parameters. Some parameters were collected as continuous variables but they were modified as categorical variable at the end. In statistical analyses part, Watershed land use type was classified with selected machine learning algorithms (k-nearest neighbor algorithm, naive bayes algorithm and c4.5 classifier algorithm) and artificial neural networks. For this purpose, hydro-morphological observations and water quality parameters such as Rosgen type, salinity, sinuosity, ph, salt, oxygen etc. were considered in order to predict the watershed land use type. In addition, the *k*-nearest neighbor algorithm (kNN), Naive Bayes classifier, c4.5 classifier and artificial neural networks (ANN) were used as classifier machine learning methods. Analyzes were performed with R statistical software. In conclusion, the classification results for each algorithm are underlined and the classifier which has revealed the best performance was determined.

Keywords: Rosgen, Machine learning, Watershed land, Classification