Abstract

Selection of classifier for use in prediction is a challenge. To select the best classifier comparisons can be made on various aspects of the classifiers. The key objective of this paper was to compare performance of nearest neighbor (ibk), regression by discretization and isotonic regression classifiers for predicting predefined precipitation classes over Voi, Kenya. We sought to train, test and evaluate the performance of nearest neighbor (ibk), regression by discretization and isotonic regression classification algorithms in predicting precipitation classes. A period of 1979 to 2008 daily Kenya Meteorological Department historical dataset on minimum/maximum temperatures and precipitations for Voi station was obtained. Knowledge discovery and data mining method was applied. A preprocessing module was designed to produce training and testing sets for use with classifiers. Isotonic Regression, K-nearest neighbours classifier, and RegressionByDiscretization classifiers were used for training training and testing of the data sets. The error of the predicted values, root relative squared error and the time taken to train/build each classifier model were computed. Each classifier predicted output classes 12 months in advance. Classifiers performances were compared in terms of error of the predicted values, root relative squared error and the time taken to train/build each
classifier model. The predicted output classes were also compared to actual year classes. Classifier performances to actual precipitation classes were compared. The study revealed that the nearest neighbor classifier is a suitable for training rainfall data for precipitation classes prediction.

References

- Lin Zhu, Jing M. Chen, Qiming Qin, Mei Huang, Lianxi Wang, Jianping Li, Bao Cao: Assimilating Remote Sensing based Soil Moisture in an Ecosystem Model (BEPS) for Agricultural Drought Assessment. IGARSS (5) 2008: 437-440

Index Terms

Computer Science Algorithms
Keywords

Regression by discretization  isotonic regression  nearest neighbor(ibk)
precipitation prediction

classification algorithms

classifier performance