

Data Mining Approach to Predict Climate Changes in Sri Lanka

K.H.S. Kumarasinghe, H.M.S.N. Ariyadasa

Department of Computer Science and Technology, Uva Wellassa University, Badulla, Sri Lanka.

Knowledge of climate data or weather data in a country is essential for business, society, agriculture and energy applications. Therefore, extracting some hidden knowledge relevant to business, society, agriculture and energy by using different techniques are essential to enrich the knowledge base. The technique, data mining can answer questions that have not traditionally been solved or very time consumed to resolve. Therefore, this technique can be used to predict climate changes by using available data. Climate change prediction is a vital application in meteorology. It is one of the most scientifically, technically challenging problem across the world in the last century. Many climate predictions like rainfall prediction, thunderstorm prediction, predicting cloud conditions are major challenges for atmospheric research. Predicting the climate changes is essential to help to be cautious for the climate. This paper presents the data mining technique for predict precipitation changing patterns and temperature changing patterns using classification technique. The classification is done using decision tree algorithm with 50 years average data. The data were collected from WorldClim website, which includes climate data of many countries in the world. Data was collected from 10,000 random places in Sri Lanka. The data were in a satellite image format. Around 10,000 of random data samples were extracted using ArcGIS application including attributes Temperature, Precipitation, Altitude, Bioclim, etc. The Standard Knowledge Discovery in Databases (KDD) process was applied to the dataset to discover the hidden pattern in climate. After removing data inconsistencies in the pre-processing stage, smoothing, generalization and aggregation were applied in the transformation stage. A data model for the climate data was developed and trained by using J48 decision tree classifier algorithm. WEKA was selected as the data mining tool and it was produced a decision tree relevant to the data set. The finding of this research deep-rooted again, the different altitude levels and precipitation levels affect the temperature norm.

Keywords: Climate change, Data mining, J48, Classification, Weka, Worldclim