

Application of Data Mining Techniques for Tourism Knowledge Discovery

Authors : Teklu Urgessa, Wookjae Maeng, Joong Seek Lee

Abstract : Application of five implementations of three data mining classification techniques was experimented for extracting important insights from tourism data. The aim was to find out the best performing algorithm among the compared ones for tourism knowledge discovery. Knowledge discovery process from data was used as a process model. 10-fold cross validation method is used for testing purpose. Various data preprocessing activities were performed to get the final dataset for model building. Classification models of the selected algorithms were built with different scenarios on the preprocessed dataset. The outperformed algorithm tourism dataset was Random Forest (76%) before applying information gain based attribute selection and J48 (C4.5) (75%) after selection of top relevant attributes to the class (target) attribute. In terms of time for model building, attribute selection improves the efficiency of all algorithms. Artificial Neural Network (multilayer perceptron) showed the highest improvement (90%). The rules extracted from the decision tree model are presented, which showed intricate, non-trivial knowledge/insight that would otherwise not be discovered by simple statistical analysis with mediocre accuracy of the machine using classification algorithms.

Keywords : classification algorithms, data mining, knowledge discovery, tourism

Conference Title : ICADMA 2017 : International Conference on Advanced Data Mining and Applications

Conference Location : Sydney, Australia

Conference Dates : January 26-27, 2017