

Comparative Analysis of Classification Methods in Determining Non-Active Student Characteristics in Indonesia Open University

Authors : Dewi Juliah Ratnaningsih, Imas Sukaesih Sitanggang

Abstract : Classification is one of data mining techniques that aims to discover a model from training data that distinguishes records into the appropriate category or class. Data mining classification methods can be applied in education, for example, to determine the classification of non-active students in Indonesia Open University. This paper presents a comparison of three methods of classification: Naïve Bayes, Bagging, and C.45. The criteria used to evaluate the performance of three methods of classification are stratified cross-validation, confusion matrix, the value of the area under the ROC Curve (AUC), Recall, Precision, and F-measure. The data used for this paper are from the non-active Indonesia Open University students in registration period of 2004.1 to 2012.2. Target analysis requires that non-active students were divided into 3 groups: C1, C2, and C3. Data analyzed are as many as 4173 students. Results of the study show: (1) Bagging method gave a high degree of classification accuracy than Naïve Bayes and C.45, (2) the Bagging classification accuracy rate is 82.99 %, while the Naïve Bayes and C.45 are 80.04 % and 82.74 % respectively, (3) the result of Bagging classification tree method has a large number of nodes, so it is quite difficult in decision making, (4) classification of non-active Indonesia Open University student characteristics uses algorithms C.45, (5) based on the algorithm C.45, there are 5 interesting rules which can describe the characteristics of non-active Indonesia Open University students.

Keywords : comparative analysis, data mining, classification, Bagging, Naïve Bayes, C.45, non-active students, Indonesia Open University

Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development

Conference Location : Chicago, United States

Conference Dates : December 12-13, 2020