Abstract

Generation of a raw data set incorporating co-related attributes, providing an insight into a student’s personality and academic performance will be our primary agenda. Subsequently, the records in the data set will be grouped into different clusters. Post clustering, each cluster will be assigned a class label considering the overall student performance in that cluster. At this stage, the raw data set is segregated into training and testing data sets. A data model can now be developed as a result of a learning algorithm which will be implemented on the training data set. Succeeding, the developed data model will be evaluated based on accuracy using the testing data set. Finally, the data model would be invoked from MATLAB for predicting a student’s performance (given all the attributes).

References


**Index Terms**

Computer Science Information Systems

**Keywords**

Educational Data mining, EM luster, Filtered Clustered, SimpleKMeans, classification