Abstract

Clustering is one of the most important techniques of data mining. Clustering technique in data mining is an unsupervised machine learning algorithm that finds the groups of object such that objects in one group will be similar to one another and are dissimilar to the objects belonging to other clusters. Clustering is called unsupervised machine learning algorithm as groups are not predefined but defined by the data. So the most similar data are grouped into the clusters. In this paper, we compare five clustering algorithm namely Farthest first, MakeDensityBasedClusterer, Simple K-means, EM, Hierarchical clustering algorithm for recommending the course to the student based on student course selection & present the result. According to our simulation, we find that Simple K-means works better than other algorithms.

References

- D. A. Clausi "K-means Iterative Fisher (KIF) unsupervised clustering algorithm applied to image texture segmentation" 0031-3203/02/$22. 00 ? 2002 Pattern Recognition Society. Published by Elsevier Science Ltd
- Xuezhi Zheng, Zhihua Cai, Qu Li"An Experimental Comparison of Three Kinds of Clustering Algorithms" in Neural Networks and Brain, 2005. ICNN&B '05.
- Yushi Jing, Michele Covell, Henry A. Rowley "Comparison of Clustering Approaches for Summarizing Large Populations of Images" in Multimedia and Expo (ICME), 2010 IEEE International Conference, 978-1-4244-7493-6/10/$26.00 c2010
- Yu Guan1, Ali A. Ghorbani1, and Nabil Belacel "An Unsupervised Clustering Algorithm for Intrusion Detection" in Advances in Artificial Intelligence, The 16th Conference of the Canadian Society for Computational Studies of Intelligence (AI 2003)

Index Terms

Computer Science Data Mining

Keywords

Farthest First Makedesitybasedclusterer Simple K-means Em Hierarchical Clustering Algorithm Weka