Abstract

One to many & many to many data linkage are necessary in data mining. OCCT Implementation for one to many & Many to many Data Linkage is to identify different entities across different Data sources. Data Linkage is linking data between two different database. One to many data linkage is associated an entity from first data set with a group matching from the other data set. In many to Many Data Linkage method the entities of same type and different nature should be arrange with Map Reduce method. In the OCCT was evaluated after using data sets from three different domains: , recommender system, data leakage prevention and fraud detection. data leakage prevention domain, the goal is to detect abnormal access. Recommender system, the method is used for matching new users of the system with the items. In fraud detection legitimate transactions performed by users.

References


2. Dr. Anju Singh, Dr. Divakar Singh, Gopal Patidar "Document Clustering approach using Hebbian-type Neural Network and Agglomerative Clustering," vol. 75, issue 9, 2013.


8. Dr. Divakar Singh" Intrusion Detection based System on Probabilistic Neural Network and Fuzzy C Means Clustering, D Singh – 2013 vol. 74, issue 2, pp. 30-33


10. Flach P., Ferri C., M. Guha and Hernández-Orallo J., "DecisionTrees Model showing the Use of Area under curve ROC. Machine Learning, pp. 139-146, 2002.


Index Terms

Computer Science

Information Systems

Keywords

Clustering, classification, data matching, decision tree induction keywords, Map Reduce, Data Linkage Matching