Abstract

Optimization is an important and critical step in the data mining process and it has a huge impact on the success of a data mining process. Selecting a set of feature which is optimal for a given task is a problem which plays an important role in a wide variety of context including pattern recognition, adaptive control and machine learning.

Clusters are formed of the reduced dataset using Swarm Intelligence Technique algorithms i.e. Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO). Cluster Hypothesis is verified which is the intra cluster distance should be minimum and inter cluster distance should be maximum. Most relevant documents are stored in the clusters.

An Information Retrieval System is used for retrieval of data from the clusters. When user enters a query from a Graphical User Interface, using Information Retrieval algorithm the document is searched and retrieved from the clusters. It is then given as an output to the user.
References

1. Sandeep U. Mane Assistance Professor, Dept. of CSE, Pankaj G. Gaikwad M. Tech
   Student, Dept. of CSE, "Nature Inspired Techniques for Data Clustering" 2014
2. C. Ahn, J. An, J. Yoo, "Estimation of particle swarm distribution algorithms: combining the
   benefits of PSO and EDAs", 2012
3. D. Karaboga, B. Akay, "A modified artificial bee colony (ABC) algorithm for constrained
   optimization problems", 2011.
   real-world vehicle routing problems", 2007
7. X. Zhang, L. Tang, "A new hybrid ant colony optimization algorithm for the vehicle routing
   problem", 2009
8. S. Mazzeo, I. Loiseau, "An ant colony algorithm for the capacitated vehicle routing
   problem", 2004

Index Terms

Computer Science  Information Sciences

Keywords

Optimization, Swarm Intelligence Technique, Clusters.