Abstract

Clustering techniques have a wide use and importance nowadays and this importance tends to increase as the amount of data grows. K-means is a simple technique for clustering analysis. Its aim is to find the best division of n entities into k groups (called clusters), so that total distance between the group's members and corresponding centroid, irrespective of the group is minimized. Each entity belongs to the cluster with the nearest mean. It results into a
partitioning of the data space into Voronoi Cells. This paper is about the implementation of k-means clustering using crop yield records by Weka Interface. The data has been taken from the website "Agricultural Statistics of India". This paper also includes detailed result analysis of rice data after demonstration via Weka Interface.

References

- Privacy-Preserving K-Means clustering over vertically Partitioned Data- By Jaideep Vaidya and Chris Clifton, Deptt. Of Computer Sciences, Purdue University, 250 N University St, West Lafayette, IN 47907-2066
- K-means clustering Tutorial- By Kardi Teknomo, Ph. D
- K-means encyclopedia

Index Terms

Computer Science
Data Mining
Keywords
K-means Clustering    Euclidean Distance    Spatial Data Mining    Weka Interface