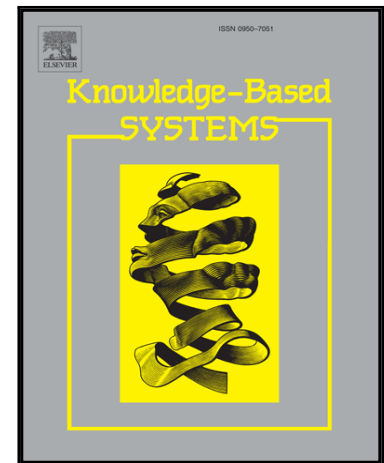


Accepted Manuscript

Natural Neighbor-based Clustering Algorithm with Local Representatives

Dongdong Cheng, Qingsheng Zhu, Jinlong Huang, Lijun Yang, Quanwang Wu

PII: S0950-7051(17)30111-9
DOI: [10.1016/j.knosys.2017.02.027](https://doi.org/10.1016/j.knosys.2017.02.027)
Reference: KNOSYS 3841



To appear in: *Knowledge-Based Systems*

Received date: 18 November 2016
Revised date: 21 February 2017
Accepted date: 22 February 2017

Please cite this article as: Dongdong Cheng, Qingsheng Zhu, Jinlong Huang, Lijun Yang, Quanwang Wu, Natural Neighbor-based Clustering Algorithm with Local Representatives, *Knowledge-Based Systems* (2017), doi: [10.1016/j.knosys.2017.02.027](https://doi.org/10.1016/j.knosys.2017.02.027)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Natural Neighbor-based Clustering Algorithm with Local Representatives

Dongdong Cheng^a, Qingsheng Zhu^{a,*}, Jinlong Huang^a, Lijun Yang^a,
Quanwang Wu^a

^aChongqing Key Lab. of Software Theory and Technology, College of Computer Science,
Chongqing University, Chongqing 400044, China

Abstract

Clustering by identifying cluster centers is important for detecting patterns in a data set. However, many center-based clustering algorithms cannot process data sets containing non-spherical clusters. In this paper, we propose a novel clustering algorithm called NaNLORE based on natural neighbor and local representatives. Natural neighbor is a new neighbor concept and introduced to compute local density and find local representatives which are points with local maximum density. We first find local representatives and then select cluster centers from the local representatives. The density-adaptive distance is introduced to measure the distance between local representatives, which helps to solve the problem of clustering data sets with complex manifold structure. Cluster centers are characterized by higher density than their neighbors and a relatively large density-adaptive distance from any local representatives with higher density. In experiments, we compare the proposed algorithm NaNLORE with existing algorithms on synthetic and real data sets. Results show that NaNLORE performs better than existing algorithm, especially on clustering non-spherical data and manifold data.

Keywords: clustering, natural neighbor, local representatives

*Corresponding author: Qingsheng Zhu, qszhu@cqu.edu.cn

Download English Version:

<https://daneshyari.com/en/article/4946315>

Download Persian Version:

<https://daneshyari.com/article/4946315>

[Daneshyari.com](https://daneshyari.com)