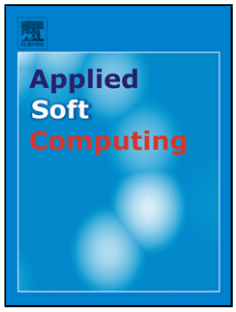
Accepted Manuscript



Title: Hesitant Fuzzy Decision Tree Approach for Highly Imbalanced Data Classification

Authors: Sahar Sardari, Mahdi Eftekhari, Fatemeh Afsari

PII: S1568-4946(17)30534-3
DOI: http://dx.doi.org/10.1016/j.asoc.2017.08.052
Reference: ASOC 4442

To appear in: *Applied Soft Computing*

Received date: 24-12-2016
Revised date: 18-8-2017
Accepted date: 23-8-2017

Please cite this article as: Sahar Sardari, Mahdi Eftekhari, Fatemeh Afsari, Hesitant Fuzzy Decision Tree Approach for Highly Imbalanced Data Classification, Applied Soft Computing Journalhttp://dx.doi.org/10.1016/j.asoc.2017.08.052

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.

ACCEPTED MANUSCRIPT

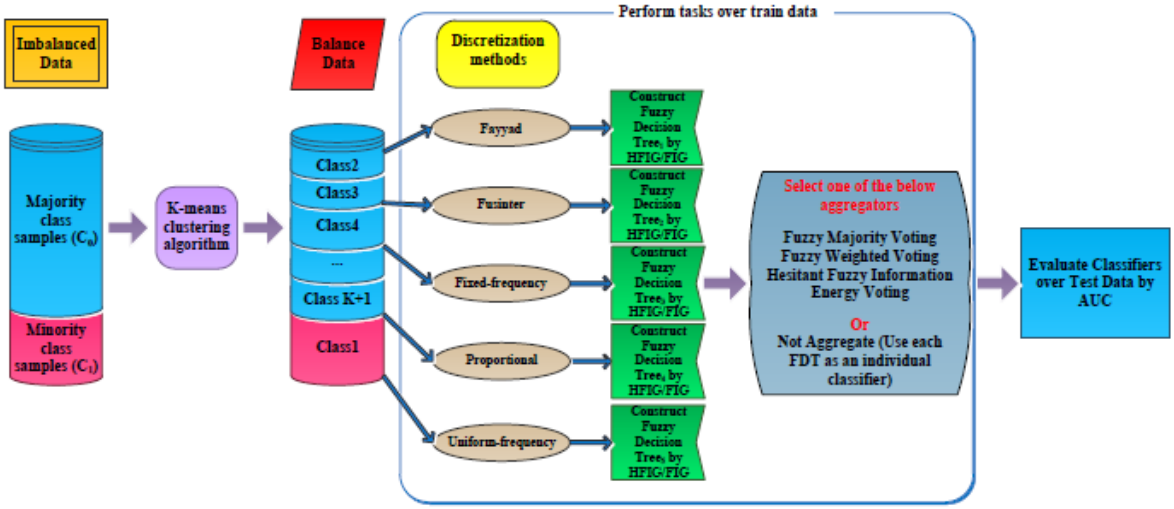
Hesitant Fuzzy Decision Tree Approach for Highly Imbalanced Data Classification

Sahar Sardari[a], Mahdi Eftekhari[a,*], Fatemeh Afsari[a]

[a]Department of computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran.

[*] - Corresponding Author, Email: m.eftekhari@uk.ac.ir

Graphical abstract



Highlights

- Using K-means clustering algorithm to divide majority class samples of imbalanced data sets into several clusters and label each cluster sample by a new synthetic class label for balancing data sets.
- Using hesitant fuzzy sets to construct five different fuzzy decision trees to aim of perform imbalanced data classification task.
- Defining a new selecting attribute criterion for expanding nodes in construction of fuzzy decision trees based on the information energy five fuzzy information gains (obtained by employing five discretization methods).
- Aggregating the results of five different fuzzy decision trees via defining of three different methods for predicting the class label of new data.
- Not aggregating the results of five fuzzy decision trees and considering the results of each fuzzy decision tree separately for predicting the class label of new data.

Abstract

Fuzzy decision tree algorithms provide one of the most powerful classifiers applied to any kind of data. In this paper, some new Fuzzy Decision Tree (FDT) approaches based on Hesitant Fuzzy Sets (HFSs) have been introduced to classify highly imbalanced data sets. Our proposed classifiers employ k-means clustering algorithm

Download English Version:

https://daneshyari.com/en/article/4962948

Download Persian Version:

https://daneshyari.com/article/4962948

Daneshyari.com