Abstract

Granular computing (GrC) is an emerging computing paradigm of information processing. It concerns the processing of complex information entities called information granules, which arise in the process of data abstraction and derivation of knowledge from information. Granular computing is more a theoretical perspective, it encourages an approach to data that recognizes and exploits the knowledge present in data at various levels of resolution or scales. Granular
computing provides a rich variety of algorithms including methods derived from interval mathematics, fuzzy and rough sets and others. Within this framework granular box regression was proposed recently. The core idea of granular box regression is to determine a fuzzy graph by embedding a given dataset into a predefined number of "boxes". Granular box regression utilizes intervals a challenge is the detection of outliers. In this paper, we propose borderline method and residual method to detect outliers in granular box regression. We also apply these methods to artificial as well as to real data of motor insurance.

References

- E. M. Knorr, R. T. Ng, "Finding intentional knowledge of distance-based outliers", 25th International Conference on Very Large DataBases (VLDB 1999).
- J. Kaufman, P. J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster
  - P. J. Rousseeuw, M. Hubert, "Robust statistics for outlier detection"; WIREs Data Mining and Knowledge Discovery 1 (2011) 73–79.
  - Fang Yu Ke, Fu Yan, Zhou Jun Lin, "Research of outlier mining based adaptive intrusion detection techniques"; third international conference on knowledge discovery and data mining 2010.
  - V. Arunkumar, Dr. A. Saradha, "An efficient data retrieval clustering based anomaly intrusion detection system in mining process with time prediction"; International journal of communication and engineering, volume 04- no. 4, Issue: 02 March2012.
  - Sherish Johri, "Novel method for intrusion detection using data mining"; International journal of advanced research in computer science and software engineering, volume2, issue4, April 2012.

Index Terms

Computer Science  Data Mining

Keywords

Granular Computing  Granular Box Regression  Outliers.