An extreme value support measure machine for group anomaly detection

Lixuan An[1], **Bernard De Baets**[2], **Stijn Luca**[3]

[1] *Ghent University, Department of Data Analysis and Mathematical Modelling, Belgium, lixuan.an@ugent.be*
[2] *Ghent University, Department of Data Analysis and Mathematical Modelling, Belgium, bernard.debaets@ugent.be*
[3] *Ghent University, Department of Data Analysis and Mathematical Modelling, Belgium, stijn.luca@ugent.be*

Abstract

Group anomaly detection is a subfield of pattern recognition that aims at detecting anomalous groups, or anomalous realizations of point process patterns rather than individual anomalous points. Existing approaches, however, are mainly focusing on modelling unusual aggregates of points in the bulk of the data, *i.e.*, in high-density regions. In this way, unusual group behaviour with a number of points located in low-density regions is not fully detected. In this study, we introduce a probabilistic model for group anomaly detection, which comprehensively detects group anomalous behaviour at both the point level and the distribution level. At the point level, we prove an analytical result where extreme value theory, a field of statistics that is well suited to model tails of distributions, is combined with point process models for generating a probabilistic point-based anomaly score. This effectively addresses the multiple hypothesis testing issue that often arises in conventional group anomaly detection methods. At the distribution level, we extend a stable discriminative model one-class support measure machine with a sigmoid probabilistic calibration technique to define a probabilistic distribution-based anomaly score. We employ a well-known uninorm aggregation function to aggregate the point-based and distribution-based anomaly scores, leading to an overall score that is more sensitive and achieves a higher accuracy compared to existing group anomaly detection models across synthetic as well as real-world datasets.

Keywords

Group anomaly detection, Extreme value theory, Point processes, One-class support measure machine, Uninorm