Abstract

In recent years, much attention has been given to the problem of outlier detection, where the aim is to identify data which behave in an unexpected way or have abnormal properties. This paper focuses on the development of non-parametric association based algorithm to effectively identify classes and outliers simultaneously. The proposed algorithm identifies outlier transaction by enhancing an association classification approach using FP-Growth. The algorithm is enhanced by using an automatic procedure for calculating the minimum support and minimum confidence automatically and introduces two new measures called collective support and confidence measure. Using these thresholds, frequent itemsets and association rules are generated. Pruning algorithms and redundant rule identification and removal procedures are used as speed optimizers. An outlier degree using the threshold is called for each association rules, using which high confidence rules are identified. The various experimental results prove that the proposed model are scalable and efficient in terms of outlier detection and classification accuracy and can be used by data mining techniques for accurate and fast knowledge discovery.

References


Index Terms

Computer Science Pattern Recognition

Keywords

Outlier Associative Rule Itemsets Confidence Support Fp-growth