

ABSTRAK

Metode penambangan aturan asosiasi langka muncul karena adanya perhatian khusus yang diberikan terhadap himpunan item langka yang terjadi dalam basis data. Item langka adalah item yang jarang muncul dalam basis data (Koh & Pears, 2010). Item langka menjadi penting karena adanya informasi penting yang terkandung di dalamnya.

Saat ini belum ada perangkat lunak generik yang bisa digunakan sebagai alat bantu untuk mendeteksi aturan langka. Kebanyakan perangkat lunak untuk proses penambangan data masih menerapkan aturan atau teknik untuk mencari aturan yang sering muncul dalam *dataset*.

Skripsi ini dimaksudkan untuk membangun sebuah perangkat lunak yang dapat digunakan sebagai alat bantu untuk mendeteksi aturan asosiasi langka dalam basis data. Algoritma untuk proses pendeteksian yang dilakukan adalah algoritma APIC (*Apriori Inverse with Clustering*).

Perangkat lunak yang dibangun telah diuji dengan menggunakan 3 *dataset* yang berbeda yaitu *congressional votes*, *zoo*, dan *heart Cleveland* yang diambil dari *website archive.ics.uci.edu/ml/machine-learning-databases*. Pengujian ini dimaksudkan untuk menguji validitas dari perangkat lunak yang dibuat dalam skripsi ini. Hasil pengujian menunjukkan bahwa perangkat lunak yang dibuat dapat menemukan aturan asosiasi langka yang valid dari dataset tersebut.

ABSTRACT

Rare association rule mining method appears because of particular attentions to the set of rare items that occur in the database. Rare items are items that appears in a few transactions (Koh & Pears, 2010). Rare items become important because of the important of information contained.

Up to know, there is no generic software that can be used as a tool to detect rare rules. Most softwares for data mining process applying rules or techniques to search for rules that frequently appear in the dataset.

This thesis is intended to build a software that can be used as a tool to detect rare association rules in the database. Algorithm for the detection process is APIC (Apriori Inverse with Clustering) algorithm.

The software has been tested using three different datasets, namely congressional votes, zoo, and heart Cleveland taken from archive.ics.uci.edu website/ml/machine-learning-databases. The testing is intended to test the validity of the software. The results show that the software could find valid rare association rules from the datasets.