

ABSTRACT (Indonesian)

Hal yang dijadikan acuan pada association rule mining adalah nilai support dan nilai confidence. Semakin tinggi nilai support dan nilai confidence maka semakin baik rules yang dihasilkan. Algoritma association rule mining menerapkan unsupervised learning dimana rule yang dihasilkan tidak ditentukan untuk menjadi class tertentu. Performansi algoritma association rule mining sangat bergantung pada ukuran dataset /dimensi yang digunakan. Performansi dapat di ukur dari waktu pemrosesan yang dihasilkan. Semakin besar dataset maka dimensinya akan semakin besar dan waktu pemrosesan akan semakin lama. Jika dimensi dapat disusutkan maka waktu proses akan semakin cepat sehingga performansi semakin baik, dengan catatan nilai confidence relatif tidak berubah.

Intersection merupakan jenis dari teori himpunan yang dapat mengurangi jumlah atribut pada 2 atau lebih himpunan yang berelasi. Oracle merupakan salah satu RDBMS, himpunan yang berelasi dapat diterapkan pada Oracle RDBMS sebagai tabel-tabel yang berelasi. IST-EFP algorithm adalah proposed algoritma yang mengkombinasikan algoritma EFP (Expand FP-Growth) dengan teori himpunan.

Pada penelitian ini algoritma IST-EFP dapat menyusutkan dimensi dari dataset sampai 87.5% dengan perbaikan waktu proses 26.6%. Nilai confidence yang dihasilkan relatif tidak berubah, contoh rule FP-18 \Rightarrow FKUE72 sebelumnya memiliki nilai confidence 23.4%, dengan algoritma IST-EFP menjadi 23.41%.

Hasil-hasil yang didapatkan dapat digunakan untuk menentukan business actions.

Keywords: association rule mining, unsupervised learning, confidence value, IST-EFP algorithm, dataset, set theory, pruned, intersections, Oracle RDBMS, business actions.