

Abstrak

Data mining adalah proses untuk mendapatkan informasi maupun pengetahuan yang terdapat pada sekumpulan data dalam jumlah besar. Klasterisasi merupakan salah satu teknik dalam *data mining* dimana prosesnya yaitu dengan mengelompokkan sekumpulan data ke dalam klaster – klaster berdasarkan banyaknya kemiripan yang dimiliki oleh masing – masing objek.

Pada tugas akhir ini, digunakan pendekatan teori graf untuk mengklasterkan data penjualan. Data penjualan direpresentasikan dahulu ke dalam sebuah graf, dimana setiap item yang berbeda dipetakan menjadi sebuah simpul, dan setiap pasangan item akan dihubungkan oleh sebuah sisi. Pasangan item akan dihubungkan oleh sebuah sisi jika nilai korelasi dari pasangan item tersebut lebih besar atau sama dengan nilai *threshold*. Proses klasterisasi dilakukan dengan memecah graf yang terbentuk menjadi beberapa sub-graf yang lengkap (*clique partition*), dimana sub-graf tersebut sering disebut dengan *clique*. Dari setiap *clique* hasil klasterisasi dapat diekuivalenkan dengan sebuah klaster.

Perubahan nilai *threshold* berpengaruh terhadap hasil klasterisasi, karena nilai *threshold* inilah yang menentukan ada atau tidaknya korelasi yang terjadi dari setiap pasangan item. Semakin kecil nilai *threshold* maka akan semakin kecil pula jumlah klaster yang diperoleh dan semakin banyak jumlah item yang terdapat di dalam sebuah klaster.

Dari proses klasterisasi dapat diperoleh informasi mengenai item – item yang rata - rata fluktuasi penjualannya tidak jauh berbeda. Informasi yang diperoleh tersebut dapat digunakan sebagai acuan untuk proses *take place* item – item tersebut.

Kata kunci: *clique, clique partition, data mining, klasterisasi, threshold*