Abstrak

Beberapa perusahaan tidak menyadari bahwa dengan banyaknya data yang berada di database dapat diolah untuk mengembangkan bisnisnya. Untuk itu diperlukan suatu cara untuk mengolah data untuk mendapatkan manfaat dari data itu, salah satunya dengan *data mining*. Dalam *data mining*, terdapat salah satu metode yang sering digunakan yaitu *clustering*. *Clustering* merupakan pengolompokan objek berdasarkan karakteristiknya. Perusahaan dapat menggunakan metode *clustering* dengan tujuan untuk mengenali pola data perusahaan sehingga akan ditemukan kecenderungan tertentu dari data tersebut.

Dalam tugas akhir ini mengimplementasikan suatu metode *clustering*, yaitu *k-harmonic means*, pada data tagihan pelanggan sebuah perusahaan. *K-harmonic means* merupakan variansi dari sebuah metode *clustering* yang terkenal yaitu *k-means*. *K-harmonic means* dapat mengatasi kelemahan yang terjadi pada *k-means* yaitu dalam hal inisialisasi titik pusat. K-means tidak dapat bekerja secara optimal pada inisialisasi titik pusat yang buruk. Pengujian dilakukan untuk melihat kualitas hasil *clustering* berdasarkan metode validasi yaitu *silhouette coefficient*.

Berdasarkan pengujian yang sudah dilakukan, K-harmonic means dapat bekerja baik dalam inisialisasi titik pusat secara random maupun inisialisasi yang buruk dibandingkan dengan k-means. Selain itu k-harmonic means dapat diterapkan pada segmentasi data tagihan pelanggan dimana *clustering* yang dihasilkan memiliki karakteristik yang berbeda antar satu sama lain.

Kata Kunci : *data mining*, *clustering*, *k-means*, *k-harmonic means*, *silhouette coefficient*