

Abstrak

Data mining merupakan sebuah proses penemuan pola-pola yang menarik dari sekumpulan data berukuran besar. Dalam data mining, banyak fungsi yang bisa dilakukan. Mulai dari klasifikasi, klastering dan aturan asosiasi. Dalam kehidupan nyata, implementasi dari data mining sangat banyak. Contohnya adalah prediksi churn, prediksi kecurangan, segmentasi pelanggan dan lain-lain. Pada tugas akhir kali ini akan dibahas mengenai segmentasi pelanggan dengan menggunakan teknik klastering yang disebut BIRCH (*Balanced Iterative Reducing and Clustering using Hierarchies*).

“Klastering merupakan cara untuk mengelompokkan objek-objek yang ada dalam basis data ke dalam kelas-kelas [8]”. Jadi, ini mampu menjawab kebutuhan untuk segmentasi pelanggan. Dengan segmentasi, pelanggan-pelanggan dibagi kedalam beberapa segmen, sehingga perusahaan dapat memberikan perlakuan kepada pelanggannya dengan cara yang berbeda berdasarkan segmennya. Dengan segmentasi pelanggan, basis data pemasaran yang sangat besar dapat diekstrak dengan tujuan untuk memperoleh informasi mengenai tingkah laku pelanggan dan akhirnya perusahaan dapat mengupayakan inisiatif dan strategi untuk retensi.

Metode clustering BIRCH menggunakan struktur pohon dengan tinggi tertentu untuk mengklasterkan dataset yang sangat besar. Kluster BIRCH menggunakan konsep *Cluster feature* (CF) yang meringkas informasi mengenai *subcluster-subcluster* pada pohon CF yang kemudian dikelompokkan kedalam sejumlah *k*-kelompok menggunakan prosedur klastering hierarki. Dari hasil analisa, didapatkan bahwa nilai *threshold* sebagai salah satu parameter yang diinputkan user dapat mempengaruhi kinerja perangkat lunak. Jika semakin kecil nilai *threshold*, jumlah *subcluster* yang dihasilkan semakin banyak dan waktu eksekusi meningkat. Selain itu, peningkatan jumlah data dan dimensi juga turut mempengaruhi waktu untuk *preclustering* dan *clustering* global, semakin banyak jumlah data dan jumlah dimensi maka semakin lama waktu yang dibutuhkan karena semakin banyaknya *subcluster* yang terbentuk.

Kata kunci: data mining, segmentasi pelanggan, klastering, BIRCH