

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Proses data mining pada dataset memperoleh hasil berupa gambaran tentang karakteristik donatur pada dataset. Setelah mendapatkan hasil karakteristik donatur berdasarkan hasil RFM dilakukan metode clustering untuk mengelompokkan dataset tersebut. Pengelompokan ini dilakukan dengan membagi 2 dataset yang sudah dioptimalkan. Optimasi ini dilakukan untuk mengurangi dataset yang ada dan untuk tetap membatasi penelitian ini agar sesuai dengan tujuan penelitian. Kedua dataset tersebut dibagi berdasarkan distribusi datanya. Dataset pertama berisi data yang merupakan outlier dari data keseluruhan dan dataset kedua berisi data distribusi normal. Dengan kesimpulan tahap ini sudah menjawab proses data mining untuk menghasilkan output yang diinginkan.

Masing-masing dataset memiliki nilai evaluasi yang berbeda-beda. Dataset yang merupakan kumpulan outlier dari keseluruhan data memiliki nilai silhouette lebih tinggi dari pada data distribusi normal. Dataset outlier memiliki nilai silhouette 0,591 pada pembagian 4 cluster, 0,936 pada pembagian 2 cluster, 0,605 pada pembagian 6 cluster dan 0,609 pada pembagian 7 cluster. Dataset yang merupakan distribusi normal memiliki nilai silhouette 0,345 pada pembagian cluster 3, 0,316 pada pembagian cluster 2, 0,319 pada pembagian cluster 4, 0,328 pada pembagian cluster 5, 0,297 pada pembagian cluster 6, 0,250 pada pembagian cluster 7, 0,285 pada pembagian cluster 8. Penulis menyimpulkan bahwa adanya pengaruh dataset pada pembagian cluster pada metode *Agglomerative Ward's Method*. Data outlier mendapatkan nilai silhouette lebih besar daripada data normal.

Berdasarkan hasil nilai silhouette score penulis dapat membagi dataset normal menjadi 3 cluster sedangkan untuk dataset outlier penulis dapat membagi dataset menjadi 2, 4, 6 dan 7 cluster. Penulis membagi dataset normal hanya menjadi 3 cluster karena memiliki nilai skor silhouette di bawah 0,5 sedangkan data outlier memiliki nilai skor silhouette di atas 0,5. Kelebihan HAC adalah tidak perlu mendefinisikan cluster di awal. Ini sesuai dengan tujuan penelitian cluster yang dapat dibuat dalam dataset.

Karakteristik dari setiap cluster dapat dilihat dari nilai rata-rata RFM. Setelah pembagian cluster penulis melakukan *grouping cluster* dengan melakukan analisis nilai RFM dengan rata-rata nilai pada setiap feature pada kedua dataset. Dalam dataset outlier dengan pembagian 2 cluster, cluster yang memiliki nilai *Monetary* tertinggi adalah cluster 2 tetapi memiliki nilai *Recency* yang tinggi. Pada pembagian 4 cluster, cluster yang memiliki nilai *Monetary* dan *Frequency* yang tertinggi adalah cluster 1. Pada pembagian 6 cluster, cluster 3 memiliki keseimbangan pada setiap nilai RFM. Pada pembagian 7 cluster, cluster 3 juga memiliki nilai keseimbangan pada setiap *feature* RFM. Pada dataset normal pembagian 3 cluster, cluster 1 dan 2 memiliki nilai yang seimbang pada datasetnya tetapi cluster 3 memiliki nilai *Recency* terbilang cukup tinggi. Penulis mendapati karakteristik dari setiap data berdasarkan cluster yang terbentuk dan dapat dilihat pada tabel 4.15.

5.2 Saran

Pada penelitian ini penulis masih banyak mengalami kekurangan. Dataset yang terlalu besar tidak cocok dalam melakukan clustering menggunakan metode hirarki. Penyebaran data yang tidak normal akibat nilai *Monetary* donatur baru terpaut jauh dengan donatur lama. Akibat dari temuan ini donatur lama menjadi outlier. Penulis menyarankan melakukan pembobotan pada *feature* RFM agar meingkatkan hasil clustering. Pembobotan ini dilakukan agar machine learning lebih memperhatikan *feature* yang ingin digunakan pada tujuan analisis.

Metode yang mungkin cocok untuk melakukan clustering analisis segmentasi berdasarkan analisis RFM adalah metode DBSCAN. DBSCAN (Algoritma Density-based Spatial Clustering of Application with Noise) dasar metode ini adalah kepadatan dari posisi kedekatan data. Pada penelitian ini banyak data yang dianggap noise sehingga algoritma ini cocok karena tidak memasukan data yang dianggap noise ke dalam cluster manapun.

