

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan pelanggan berdasarkan perilaku pembelian dan frekuensi kunjungan toko menggunakan algoritma Agglomerative Hierarchical Clustering (AHC). Dataset yang digunakan berasal dari Kaggle, terdiri atas 2.240 pelanggan dengan 29 atribut. Tahapan penelitian meliputi seleksi atribut, normalisasi, perhitungan jarak menggunakan Euclidean distance, serta pengelompokan dengan tiga metode linkage: single, complete, dan average. Evaluasi hasil clustering dilakukan menggunakan indeks Silhouette dan metode elbow, baik untuk lima atribut terpilih maupun seluruh atribut numerik. Hasil menunjukkan bahwa single linkage dan average linkage memberikan hasil terbaik dengan nilai indeks Silhouette 0,6817 untuk dua cluster. Analisis karakteristik menunjukkan bahwa cluster pertama terdiri dari pelanggan aktif, sedangkan cluster kedua berisi pelanggan yang kurang aktif. Penelitian ini memberikan kontribusi terhadap penerapan data mining dalam segmentasi pelanggan dan pengambilan keputusan strategis di sektor ritel, sehingga perusahaan dapat merancang strategi pemasaran yang lebih tepat sasaran serta meningkatkan kepuasan dan loyalitas pelanggan.

Kata kunci: segmentasi pelanggan, *Agglomerative hierarchical clustering*, perilaku pembelian, frekuensi kunjungan, *Silhouette index*, data mining.

ABSTRACT

This study aims to segment customers based on purchasing behavior and store visit frequency using the Agglomerative Hierarchical Clustering (AHC) algorithm. The dataset used is obtained from Kaggle, consisting of 2,240 customer records with 29 attributes. The research stages include attribute selection, normalization, distance calculation using Euclidean distance, and clustering using three linkage methods: single, complete, and average. The clustering results were evaluated using the Silhouette index and the elbow method, both for five selected attributes and for all numerical attributes. The results show that single linkage and average linkage methods provide the best clustering performance with a Silhouette index score of 0.6817 for two clusters. Cluster analysis indicates that the first cluster consists of active customers, while the second includes less active ones. This study contributes to the application of data mining in customer segmentation and strategic decision-making in the retail sector, enabling businesses to develop more targeted marketing strategies and enhance customer satisfaction and loyalty.

Keywords: customer segmentation, Agglomerative Hierarchical Clustering, purchasing behavior, visit frequency, Silhouette index, data mining.