

SEGMENTASI PELANGGAN MENGGUNAKAN ALGORITMA FUZZY C-MEANS (FCM) DAN ANALISIS RFM (RECENCY, FREQUENCY, AND MONETERY) PADA DATA PELANGGAN KEDAI

UUS RUSDIANA

ABSTRAK

Segmentasi pelanggan bertujuan untuk mengelompokkan pelanggan berdasarkan kesamaan karakteristik berupa ketertarikan atau permintaan pelanggan yang serupa. Untuk melakukan segmentasi pelanggan dapat dicapai melalui *data mining* dengan mengimplementasikan teknik *clustering*. Algoritma yang sering digunakan untuk melakukan *clustering* yaitu algoritma *fuzzy c-means* (berbasis *fuzzy*) dan algoritma *k-means* (berbasis klasik). Pemilihan metode pengukuran jarak data yang optimal untuk algoritma *clustering* menjadi masalah utama pada penelitian ini. Karena penggunaan metode pengukuran jarak data berpengaruh signifikan terhadap kualitas *cluster* yang dihasilkan. Algoritma yang akan digunakan pada penelitian ini yaitu *fuzzy c-means clustering* dengan *k-means clustering* sebagai pembanding yang kemudian akan coba diterapkan metode pengukuran jarak data *Euclidean distance*, *manhattan distance*, *Chebyshev distance*, dan *minkowski distance*. *Cluster* yang dihasilkan menggunakan metode pengukuran jarak data yang berbeda-beda akan dievaluasi menggunakan indeks validitas diantaranya *partition coefficient index* (PC), *modified partition coefficient index* (MPC), dan RMSE. Hasil penelitian menunjukkan algoritma *fuzzy c-means* lebih unggul daripada *k-means* dengan metrik jarak yang optimal yaitu *manhattan distance* (PC = 0.95, MPC = 0.9, dan RMSE = 0.7745) untuk pengujian pada *cluster* sebesar 2 dan *minkowski distance* (PC = 0.9338, MPC = 0.9007, dan RMSE = 0.8366) untuk pengujian pada *cluster* sebesar 3. Analisis RFM terhadap hasil segmentasi pelanggan menggunakan algoritma *fuzzy c-means* menunjukkan jumlah *cluster* yang optimal yaitu sebanyak 3 *cluster* yang memisahkan pelanggan ke dalam tiga karakteristik yaitu retensi pelanggan tinggi (nilai RFM = 3 hingga 5), retensi pelanggan sedang (nilai RFM = 2 hingga 3) dan retensi pelanggan rendah (nilai RFM=1).

Kata kunci: Segmentasi pelanggan, Analisis RFM, *Fuzzy C-Means*, *K-Means*, Metrik jarak.

**CUSTOMER SEGMENTATION USING FUZZY C-MEANS (FCM)
ALGORITHM AND RFM (RECENCY, FREQUENCY, AND MONETERY)
ANALYSIS ON CUSTOMER DATA**

UUS RUSDIANA

ABSTRACT

Customer segmentation aims to classify customers based on the similarity of characteristics in the form of similar customer interests or demands. To perform customer segmentation, it can be achieved through data mining by implementing clustering techniques. The algorithms that are often used for clustering are the fuzzy c-means algorithm (fuzzy based) and the k-means algorithm (classical based). The selection of the optimal distance metric measurement method for the clustering algorithm is the main problem in this study. Because the use of the distance metric measurement method has a significant effect on the quality of the resulting clusters. The algorithm that will be used in this study is fuzzy c-means clustering with k-means clustering as a comparison which will then try to apply the Euclidean distance, Mahnattan distance, Chebyshev distance, and Minkowski distance measurement methods. The resulting clusters using different distance metrics measurement methods will be evaluated using validity indices including Partition Coefficient Index (PC), Modified Partition Coefficient Index (MPC), and RMSE. The results show that the fuzzy c-means algorithm is superior to k-means with the optimal distance metric, namely Manhattan distance (PC = 0.95, MPC = 0.9, and RMSE = 0.7745) for testing on clusters of 2 and minkowski distance (PC = 0.9338, MPC = 0.9007, and RMSE = 0.8366) for testing on clusters of 3. RFM analysis of the results of customer segmentation using the fuzzy c-means algorithm shows the optimal number of clusters, namely 3 clusters which separate customers into three characteristics, namely high customer retention (value RFM = 3 to 5), moderate customer retention (RFM value = 2 to 3) and low customer retention (RFM value = 1).

Keywords: Customer Segementation, RFM Analysis, Fuzzy C-Means, K-Means, Distance Metric.