

ABSTRACT

Intense competition in business areas, motivates a industry to manage customer optimally. Knowing the characteristic of customers will help companies to determine the right and effective marketing strategies, so that give more benefits for the company. This study discussing about how the process of data mining of customer data in UD Gemilang Kencana, as food industry which sell product of “manisan carica”.

The method used is CRISP-DM (Cross Industry Process for Data Mining) through business understanding, data understanding, data preparation, modeling, evaluation and deployment. Data mining process begins with clustering process using K-Means algorithm. The clustering result is used for the segmentation process using Fuzzy RFM model to obtain customer class. Fuzzy RFM model combines the fuzzy set theory with RFM model, the segmentation model based on recency, frequency dan monetary attribute.

The results of the test, obtained the 3 cluster ($k=3$) is the most optimal number of cluster according to the calculation with Davies Bouldin Index (DBI). Customer segmentation results obtained customers by 26 customers Superstar D (20,47%) in cluster 1, cluster 2 as much as 80 customers (62,99%) as Dormant E class and cluster 3 as much as 21 customers (16,54%) as Dormant F. The testing data result obtained the value of Sum of Squared Error decreased, shows clustering result is good enough with increasing number of cluster (k). Meanwhile for Silhouette Coefficient test obtained the quality of the cluster in medium structure category. It can be concluded the process of segmentation using the K-Means algorithm has been successfully identified potential customers with 3 clusters. The determination of customer label with fuzzy RFM obtained more accurate and the customer class is described more detail for each attribute.

Keywords -- customer segmentation, data mining, clustering, K-Means algorithm, fuzzy RFM

INTISARI

Persaingan di bidang bisnis memotivasi sebuah industri untuk mengelola pelayanan terhadap konsumen tetap (pelanggan) secara maksimal. Mengetahui karakteristik setiap pelanggan akan membantu pelaku bisnis dalam menentukan strategi pemasaran yang tepat dan efektif sehingga mendatangkan keuntungan bagi pihak perusahaan. Penelitian ini membahas tentang bagaimana proses *data mining* dari data pelanggan di UD Gemilang Kencana, sebuah industri makanan berskala UKM yang menjual produk manisan carica.

Metode yang digunakan dalam penelitian ini adalah CRISP-DM (*Cross Industry Process for Data Mining*) dengan melalui beberapa proses, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment*. Proses *data mining* dimulai dengan melakukan proses *clustering* menggunakan algoritme K-Means. Hasil *clustering* dari masing-masing metode tersebut digunakan untuk melakukan segmentasi menggunakan model *Fuzzy RFM* untuk mendapatkan kelas pelanggan. *Fuzzy RFM* menggabungkan teori himpunan *fuzzy* dengan model RFM, yaitu model segmentasi berdasarkan atribut *recency*, *frequency* dan *monetary*.

Hasil uji coba yang dilakukan, diperoleh jumlah *cluster* 3 merupakan jumlah paling optimal menurut perhitungan *Davies Bouldin Index* (DBI). Hasil segmentasi pelanggan diperoleh pelanggan pada *cluster* 1 sebanyak 26 pelanggan (20,47%) sebagai *Superstar D Customers*, *cluster* 2 sebanyak 80 pelanggan (62,99%) sebagai *Dormant E Customers* dan pada *cluster* 3 sebanyak 21 pelanggan (16,54%) sebagai *Dormant F Customer*. Hasil pengujian diperoleh nilai *Sum of Squared Error* menurun, menunjukkan hasil *clustering* cukup baik seiring dengan meningkatkan jumlah *cluster* (*k*). Sedangkan untuk uji *Silhouette Coefficient* diperoleh kualitas *cluster* dalam kategori *medium structure*. Jadi dapat disimpulkan proses segmentasi menggunakan algoritme K-Means telah berhasil mengidentifikasi pelanggan potensial dengan 3 *cluster*. Penentuan label pelanggan dengan *fuzzy RFM* diperoleh hasil yang lebih akurat dan kelas pelanggan dijelaskan lebih detail untuk setiap atribut yang digunakan.

Kata kunci -- segmentasi pelanggan, *data mining*, *clustering*, CRISP-DM, algoritme K-Means, *fuzzy RFM*.