

Customer Segmentation Dengan Metode Self Organizing Map (Studi Kasus: UD. Fenny)

A. A. Gde Bagus Ariana
STIKI Denpasar
email: gungariana@yahoo.com

Abstrak

Saat ini persaingan bisnis pada perusahaan retail tidak hanya dengan menggunakan perangkat sistem informasi namun sudah dilengkapi dengan sistem pendukung keputusan. Salah satu metode sistem pendukung keputusan yang digunakan adalah data mining. Data mining digunakan untuk menemukan pola-pola yang tersembunyi pada database. UD. Fenny sebagai perusahaan retail ingin menemukan pola segmentasi pelanggan dengan menggunakan model RFM (Recency, Frequency, Monetary). Metode data mining untuk melakukan proses segmentasi adalah metode clustering. Clustering merupakan proses penggugusan data menjadi kelompok-kelompok yang memiliki kemiripan secara tidak terawasi (*unsupervised*). Sebelum melakukan proses clustering, dilakukan proses persiapan data dengan membuat datawarehouse menggunakan skema bintang (*star scema*). Selanjutnya dilakukan proses clustering dengan menggunakan metode Self Organizing Map (SOM/Kohonen). Metode ini merupakan salah satu model jaringan saraf tiruan yang menggunakan metode *unsupervised*. Dari hasil percobaan metode SOM melakukan proses clustering dan menggambarkan hasil clustering pada SOM plot. Dengan melakukan proses clustering, pihak pengambil keputusan dapat memahami segmentasi customer dan melakukan upaya peningkatan pelayanan customer.

Kata kunci: Segmentasi, SOM, Recency, Frequency, Monetary

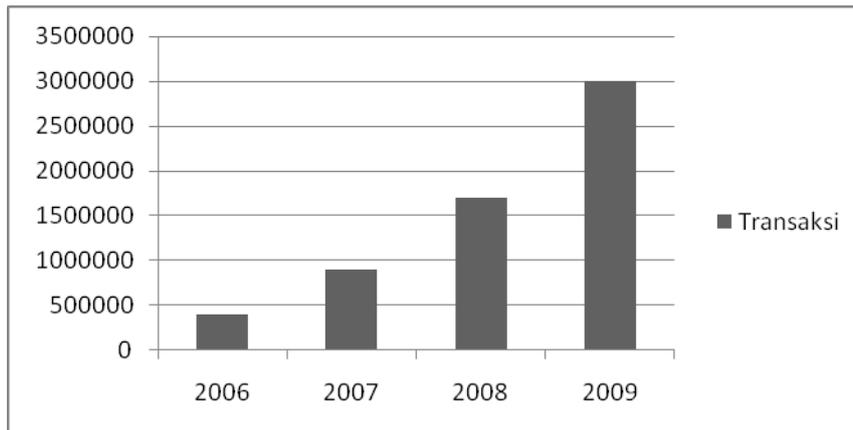
Abstract

Currently the company's retail business competition is not only using the information system but also is equipped with a decision support system. One method is to use decision support system is data mining. Data mining is used to find patterns hidden in databases. UD. Fenny as a retail company would like to find a pattern by using a customer segmentation model RFM (Recency, Frequency, Monetary). Data mining methods to make the process of segmentation is a method of clustering. Clustering is an *unsupervised* process of segmentation data into groups that have similarities. Before performing the clustering process, carried out the data preparation process by creating datawarehouse using star schema. Clustering process is then performed using the Self Organizing Map (SOM / Kohonen). This method is one of the artificial neural network models that use *unsupervised* methods. From the experimental results of the SOM method to describe the process of clustering and clustering on the SOM plot. By performing the process of clustering, decision makers can understand customer segmentation and improve efforts in customer service.

Key words: Segmentation, SOM, recency, Frequency, Monetary

1. PENDAHULUAN

Perusahaan ritel adalah perusahaan yang bergerak dalam bidang transaksi jual beli dalam jumlah kecil, satuan, atau eceran. Usaha ritel ini memiliki peranan penting dalam dunia perekonomian, karena tanpa usaha ritel, suatu barang dari produsen tidak akan sampai di tangan konsumen. UD. Fenny merupakan perusahaan ritel yang menjual bahan kue, berlokasi di Jalan Nakula, Denpasar. Sejak tahun 2006 UD. Fenny telah menggunakan sistem informasi untuk menangani *point of sales* (POS) dan sistem *inventory*. Database UD. Fenny bertambah dengan cepat dari tahun ke tahun. Ukuran database yang bertambah dengan cepat disebabkan oleh volume transaksi yang tinggi, jumlah customer yang banyak dan item barang yang beragam. Saat ini terdapat hampir tiga juta transaksi penjualan, 1.200 customer, dan 10.000 item barang.



Gambar 1 Pertumbuhan *record* transaksi penjualan (UD. Fenny, 2010)

Jumlah *record* transaksi yang besar ternyata menyebabkan pihak manajemen sulit untuk menentukan keputusan-keputusan bisnis yang terkait dengan masalah strategi pemasaran. Akibatnya perusahaan sering kehilangan *customer* yang potensial menjadi *customer* loyal dan biaya *marketing* terkadang tidak tepat sasaran. Kondisi seperti ini disebut "*data rich, poor information*", atau kebanjiran data, namun miskin informasi.

Untuk mengatasi permasalahan tersebut, dilakukan proses *customer identification* atau pengenalan *customer*. Salah satu cara untuk melakukan hal tersebut adalah *customer segmentation*. Tujuan utama melakukan *customer segmentation* adalah untuk mengetahui tingkah laku *customer*. *Customer* dapat disegmentasi dengan mengacu pada beberapa variabel. Salah satu model yang sering digunakan adalah model RFM (*Recency, Frequency, Monetary*). *Recency* menggambarkan waktu pembelian akhir dari *customer*, *frequency* menggambarkan frekuensi kedatangan *customer* dan *monetary* menggambarkan jumlah uang yang dibelanjakan oleh *customer*.

Untuk melakukan segmentasi pada *database* besar, dilakukan proses *data mining*. Menurut Berry dan Linoff (2004) *data mining* adalah proses untuk mengekstraksi atau mendeteksi pola-pola yang tersembunyi dari sebuah *database* besar. Salah satu teknik *data mining* yang digunakan untuk menyelesaikan permasalahan segmentasi adalah *clustering*. *Clustering* merupakan penggugusan data ke dalam kelompok-kelompok secara alami. Dengan melakukan *clustering*, maka data yang memiliki kemiripan pola akan berada pada *cluster* yang sama dan data yang berbeda pola akan berada pada *cluster* yang berbeda.

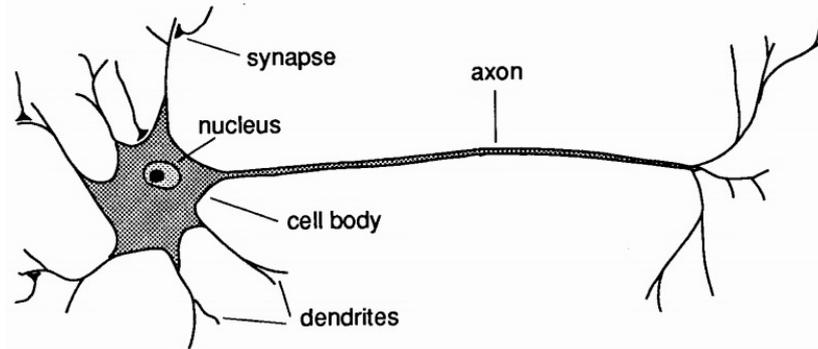
Terdapat bermacam-macam metode untuk melakukan *clustering*, mulai dari metode sederhana dengan menggunakan statistik hingga metode kompleks dengan menggunakan kecerdasan buatan. Salah satu metode *clustering* adalah dengan menggunakan jaringan saraf tiruan atau *neural network*. Metode jaringan saraf tiruan yang digunakan untuk melakukan *clustering* adalah metode jaringan saraf yang menggunakan paradigma pembelajaran tidak terawasi (*unsupervised*). Penelitian ini akan dilakukan proses *data mining* untuk menyelesaikan masalah *customer* segmentasi pada perusahaan *retail* UD. Fenny dengan menggunakan metode SOM/Kohonen.

2. TEORI PENUNJANG

Jaringan saraf tiruan memegang peranan penting pada kemajuan teknologi terutama digunakan untuk membaca pola atau *pattern recognizing*. Berikut akan dijabarkan perbandingan jaringan saraf tiruan dengan jaringan biologis, arsitektur jaringan saraf, dan metode pembelajaran pada jaringan saraf.

2.1. Perbandingan Jaringan Saraf Biologi dan Tiruan

Pembuatan struktur jaringan saraf tiruan diilhami oleh struktur jaringan biologi, khususnya jaringan otak manusia. Neuron adalah satuan unit pemroses terkecil pada otak, bentuk sederhana sebuah neuron yang para ahli dianggap sebagai satuan unit pemroses tersebut di gambarkan sebagai berikut:



Gambar 2 Struktur Jaringan Saraf Biologis dan Tiruan

Jaringan saraf biologi merupakan kumpulan dari sel-sel saraf (neuron). Neuron memiliki tugas mengolah informasi. Komponen-komponen utama dari sebuah neuron dapat dikelompokkan menjadi tiga bagian yaitu:

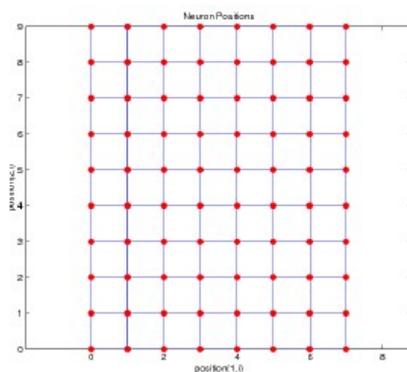
- 2 Dendrit, bertugas untuk menerima informasi
- 3 Badan Sel (Soma), badan sel berfungsi sebagai tempat pengolah informasi.
- 4 Akson, mengirim impuls-impuls ke sel saraf lainnya.

Tabel 1 Perbandingan Jaringan Saraf Biologis dan Tiruan

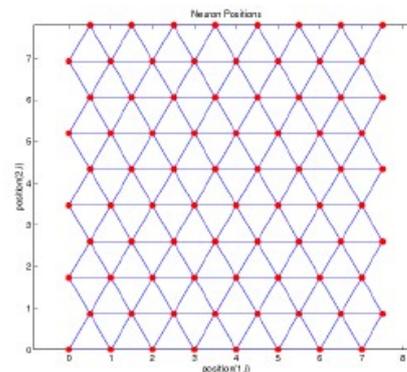
Jaringan Saraf Tiruan	Jaringan Saraf Biologis
Node atau <i>input</i>	Badan sel (soma)
<i>Input</i>	Dendrit
<i>Output</i>	Akson
Bobot	Sinapsis

2.2. Topologi Jaringan SOM

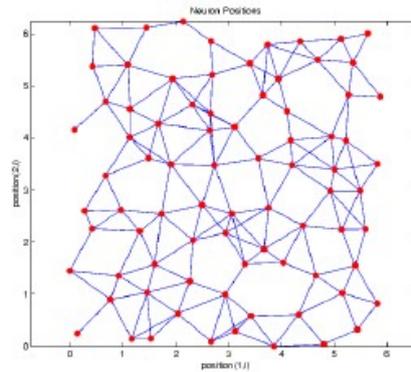
Neuron target Jaringan SOM tidak diletakkan dalam sebuah baris seperti layaknya model jaringan syaraf tiruan yang lain. *Neuron* target diletakkan dalam 2 dimensi yang bentuk/topologinya dapat diatur. Fasilitas untuk mendefinisikan topologi jaringan yang berbeda-beda disediakan oleh Matlab. Topologi yang berbeda akan menghasilkan *neuron* sekitar *neuron* pemenang yang berbeda sehingga bobot yang dihasilkan juga akan berbeda. Ada 3 macam topologi yang dapat dibentuk yaitu:



(a) Gridtop



(b) Hextop



(c) Randtop

Gambar 3 Topologi jaringan SOM.

2.3. Algoritma Jaringan Kohonen

Algoritma pengelompokan pola jaringan Kohonen adalah sebagai berikut:

0. Inisialisasi
 - a. Bobot w_{ji} (acak)
 - b. Laju pemahaman awal dan faktor penurunannya
 - c. Bentuk dan jari-jari ($=R$) topologi sekitarnya.
1. Selama kondisi penghentian bernilai salah, lakukan langkah 2-7
2. Untuk setiap vektor masukan x , lakukan langkah 3-5
3. Hitung $D_j = \sum_i (w_{ji} - x_i)^2$ untuk semua j
4. Tentukan indeks J sedemikian hingga $D_{(j)}$ minimum
5. Untuk setiap unit j di sekitar J modifikasi bobot:
$$w_{ji}^{baru} = w_{ji}^{lama} + \alpha(x_i - w_{ji}^{lama})$$
6. Modifikasi laju pemahaman
7. Uji kondisi penghentian

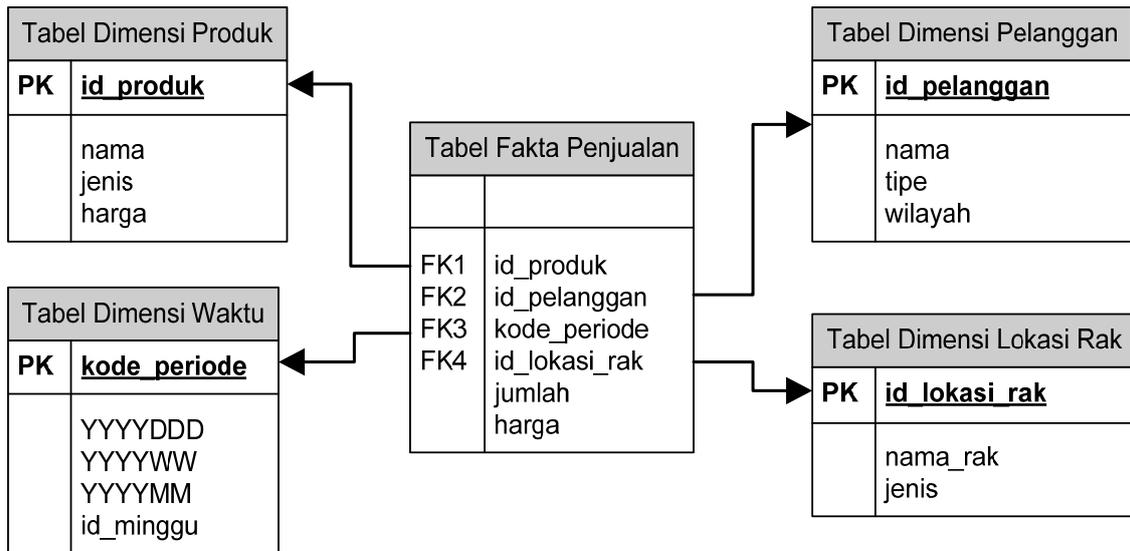
Kondisi penghentian iterasi adalah selisih antara w_{ji} saat itu dengan w_{ji} pada iterasi sebelumnya. Apabila semua w_{ji} hanya berubah sedikit saja, berarti iterasi sudah mencapai konvergensi sehingga dapat dihentikan.

3. METODE PERANCANGAN

4.1. Tahap perancangan datawarehouse.

Pada tahap ini, akan dirancang *data warehouse* untuk menampung data dari *database* transaksi. Data dari *database* tersebut akan melawati proses ETL (*Extracts, Transfer, Loading*). Model konseptual *data warehouse* yang digunakan adalah *star schema*.

Star schema akan menggambarkan *fact table*, yaitu tabel yang merepresentasikan *measure*, sebagai pusat data. Tabel ini nantinya akan terkoneksi dengan tabel-tabel yang mendeskripsikan dimensi untuk *measure* tersebut (*dimension table*). Adapun tabel fakta adalah tabel fakta penjualan, sedangkan tabel dimensi terdiri dari tabel dimensi produk, pelanggan, lokasi rak, dan waktu. Berikut adalah skema *data warehouse*.



Gambar 4 Rancangan datawarehouse

4.2. Tahap Persiapan Data.

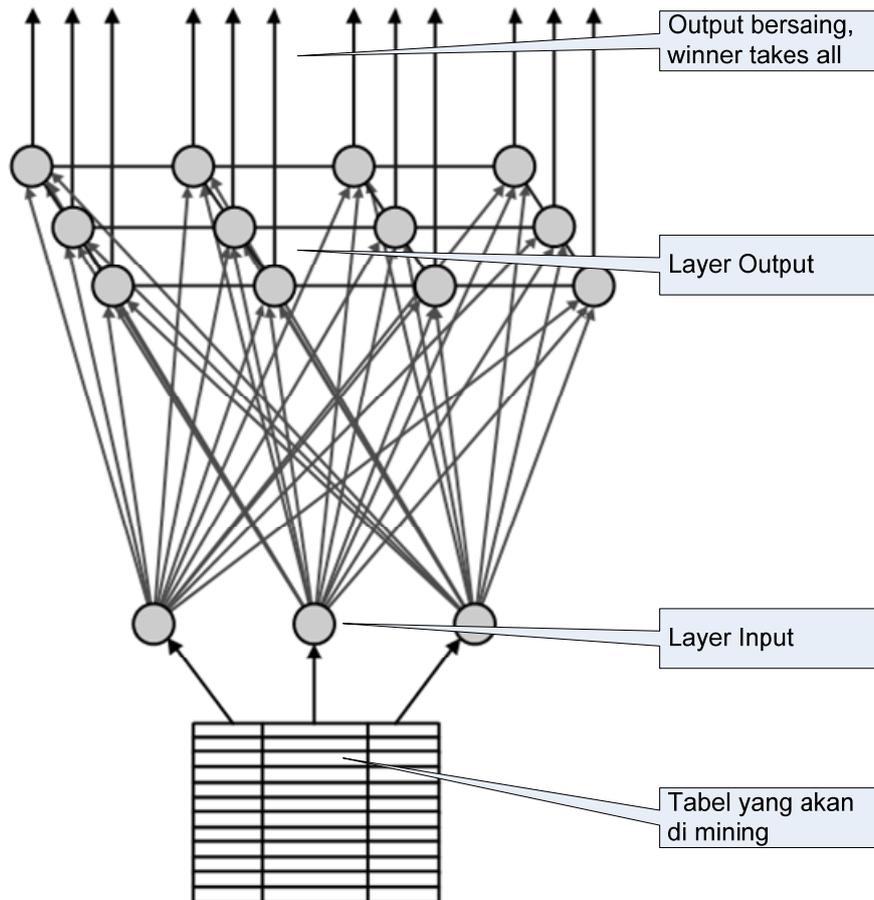
Setelah proses pembuatan data *warehouse*, maka sistem akan melakukan *query* data untuk mendapatkan data yang siap untuk di-*mining*. Sesuai dengan tujuan penelitian, maka akan dilakukan ekstraksi data terhadap data yang mewakili atribut *recency*, *frequency*, dan *monetary*.

Tabel 2 Data yang akan di-mining

Field	Keterangan
KodeCustomer	Kode <i>customer</i>
TglBeliAkhir	Menandakan <i>recency</i> , merupakan tanggal transaksi pembelian terakhir yang dilakukan oleh <i>customer</i>
FrekuensiBeli	Menandakan <i>frequency</i> , merupakan jumlah transaksi selama periode yang ditentukan
TotalBeli	Menandakan <i>monetary</i> , merupakan jumlah uang selama periode yang ditentukan

4.3. Tahap Mining dengan Metode SOM

Berikut adalah gambaran topologi metode SOM yang akan melakukan proses *cluster* terhadap data yang sudah disiapkan.



Gambar 5 Arsitektur *cluster* dengan metode SOM

4. HASIL DAN PEMBAHASAN

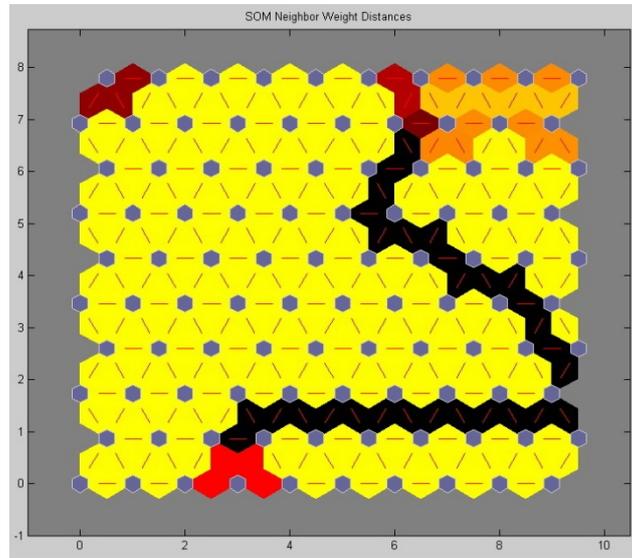
Database transaksi yang digunakan di UD. Fenny menggunakan SQL Server 2000. Pembuatan *data warehouse* juga menggunakan SQL Server 2000. Selanjutnya untuk mendapatkan data yang mewakili *recency*, *frequency*, dan *monetary* digunakan bahasa pemrograman Microsoft Visual Basic. 6.0. Berikut adalah tampilan data sebelum dilakukan proses *mining*.

Customer ↓	Total	Frekwensi	AVG	Recent
00000	8,259,950.00	12.00	688,329.17	212.00
0005	6,463,350.00	12.00	538,612.50	2.00
0006	4,500,000.00	1.00	4,500,000.00	32.00
7000	92,158,550.00	36.00	2,559,959.72	17.00
70001	56,998,150.00	41.00	1,390,198.78	3.00
700010	19,093,000.00	23.00	830,130.43	3.00
700011	19,779,300.00	10.00	1,977,930.00	5.00
700012	20,542,750.00	9.00	2,282,527.78	24.00
700013	14,162,280.00	9.00	1,573,586.67	10.00
700014	58,476,200.00	19.00	3,077,694.74	9.00

Gambar 6 Query Data yang akan di-clustering.

Setelah dilakukan proses *mining* dengan menggunakan *software* Matlab 2010a, berikut adalah *output* yang dihasilkan dari proses *clustering* dengan menggunakan metode SOM. Pada

gambar terlihat terbentuk tiga *cluster*. Jika diperbesar maka akan tampak label kode *customer* pada masing-masing neuron yang akan menunjukkan *cluster* dari masing-masing *customer*.



Gambar 7 Hasil cluster dengan SOM

5. PENUTUP

5.1. Kesimpulan

Berdasarkan penelitian yang telah dilakukan maka dapat disimpulkan bahwa metode *self organizing map* dapat digunakan untuk mengetahui pola-pola *cluster* pada sebuah basis data. Metode SOM bersifat *black box*, jadi kita tidak mengetahui bagaimana metode mengelompokkan data menjadi sebuah *output*. Diperlukan penelitian lebih lanjut untuk mengetahui perbandingan ukuran map pada *layer* dua dari topologi SOM, sehingga didapat hasil *cluster* yang lebih valid dan akurat.

6. DAFTAR PUSTAKA

- [1] Berry, M. J. A., & Linoff, G. S. (2004). *Data mining techniques second edition-for marketing, sales, and customer relationship management*. Wiley.
- [2] Carpenter, G.A., & S. Grossberg. (1987b). "ART2: Self-organization of Stable Category Recognition Codes for Analog Input Patterns." *Applied Optics*, 26:4919-4930. a reprint from *Applied Optics* volume 26, number 23, December, 1987
- [3] Carrier, C. G., & Povel, O. (2003). *Characterising data mining software*. *Intelligent Data Analysis*, 7, 181-192.
- [4] Cheng, C. H. & Chen Y. S. (2009). *Classifying the segmentation of customer value via RFM model and RS Theory*. *Expert System with Application*, 36, 4176-4184.
- [5] Fausset, Laurene,(1994) *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*, Prentice-Hall, New Jersey, USA.
- [6] H Lu, R Setiono,(1996) H Liu. *Effective Data Mining Using Neural Network*. *IEEE Transactions on Knowledge and Data Engineering*, 8(6): 957-961.
- [7] Hughes, A. M. (1994). *Strategic database marketing*. Chicago: Probus Publishing Company.
- [8] Kincaid, J. W. (2003). *Customer relationship management: Getting it right*. Upper saddle river. N. J.: Prentice Hall PTR.
- [9] Kohonen, T. (1990). *The Self-Organizing Map, Invited Paper*. *Proceedings of the IEEE*, Vol. 78, No. 9, September 1990.
- [10] Krieger, C. (1996). *Neural Networks in Data Mining*
- [11] Tsipstis, K., Choriantopoulos, A. 2009. "Data Mining Techniques in CRM". A John Wiley and Sons, Ltd., Publication. United Kingdom.