

REVIEW PEMANFAATAN TEKNIK DATA MINING DALAM SEGMENTASI KONSUMEN

Armadyah Amborowati^{1,2}
*Edi Winarko*³

¹Mahasiswa Program Doktor Program Studi Ilmu Komputer MIPA
Universitas Gadjah Mada

²Program Studi Teknik Informatika STMIK AMIKOM Yogyakarta

³Program Studi Ilmu Komputer Fakultas MIPA Universitas Gadjah Mada
^{1,2}armadyah.a@amikom.ac.id, ³ewinarko@ugm.ac.id

Abstrak

Dukungan data mining dalam strategi pemasaran sangat besar. Segmentasi konsumen merupakan salah satu proses yang dilakukan dalam strategi pemasaran. Untuk mendukung hasil yang proses pengelompokan konsumen atau segmentasi konsumen ini maka dukungan data mining sangat berperan penting. Berdasarkan pemetaan penelitian mengenai dukungan data mining pada segmentasi konsumen didapat ada dua algoritma yang sering digunakan untuk segmentasi konsumen antara lain K-Means Clustering dan Fuzzy C-Means clustering. Adapun atribut-atribut yang digunakan untuk proses mining pada segmentasi konsumen adalah data konsumen, produk, demografi, perilaku konsumen, transaksi, RFMDC, RFM (Rerency, Frequency Monetary), dan LTV (Life Time Value). Dan penting untuk menggabungkan algoritma clustering dengan algoritma Classification, Association, dan CPV untuk mendapatkan nilai potensial dari tiap cluster.

Kata kunci: *Data mining, strategi pemasaran, segmentasi konsumen*

PENDAHULUAN

Data mining adalah suatu teori yang mempelajari bagaimana memanfaatkan data-data historis dengan menemukan suatu pola-pola yang digunakan untuk membantu dalam pengambilan keputusan. Hasil dari proses mining adalah pola-pola atau *pattern*. Pola-pola ini dianalisis sehingga ditemukan suatu pengetahuan baru yang berguna bagi manager dalam mengambil suatu keputusan. Dukungan keputusan dari proses mining ini biasanya digunakan untuk membantu dalam penyelesaian masalah-masalah strategis karena hasil

yang ditemukan biasa berupa suatu prediksi untuk masa depan.

Bisnis pada masa sekarang sangat ketat. Persaingan ada dimana-mana. Apalagi dengan adanya pasar terbuka maka suatu Perusahaan harus melakukan suatu proses untuk meningkatkan keunggulan kompetitifnya. Beberapa cara yang dilakukan oleh Perusahaan salah satunya adalah dengan membangun suatu sistem yang disebut *Business Intelligence* (BI). *Business Intelligence* (BI) adalah suatu sistem yang terintegrasi yang dikoordinasi penuh oleh Perusahaan. BI dibangun untuk mendukung dalam pengambilan

keputusan di semua area dalam perusahaan termasuk oleh top manager. BI mengkombinasikan antara penyimpanan data dan pengetahuan manajemen untuk menganalisisnya. Teknik yang dilakukan dalam BI antara lain ETL (*Extract Transformation Loading*), data warehouse, data mining dan data visualisasi seperti OLAP (*Online Analytical Processing*). Paper ini memfokuskan teknik data mining pada area strategi pemasaran khususnya segmentasi konsumen.

METODE PENELITIAN

Teknik Data Mining untuk Strategi Pemasaran

Pemasaran adalah suatu proses tentang pengembangan produk, periklanan, distribusi dan penjualan (Zhang, 2008). Dukungan data mining pada pemasaran adalah pada *marketing reaserach* dan *Business Intelligence*. Dalam mengoptimisasi proses pemasaran diperlukan suatu strategi sehingga dapat digunakan untuk meningkatkan keunggulan kompetitif.

Data mining dalam strategi pemasaran menggunakan salah satunya menggunakan database marketing untuk melakukan proses pencarian pengetahuan baru guna mendukung pengambilan keputusan (Guaxiang, 2013). Proses mining pada database marketing dilakukan untuk menganalisa prediksi konsumen potensial, menganalisis optimasi *sales-channel*, menganalisis produk yang habis bersamaan, menganalisis karakteristik konsumen, dan lain sebagainya.

Dukungan data mining untuk strategi pemasaran lainnya yaitu sistem BI berdasarkan CRM (Hong, 2009). Penelitian ini memfokuskan menganalisis aplikasi CRM dan data warehouse. Lihe, 2008, menerangkan

bahwa pada suatu sistem CRM berbasis BI harus mempunyai fungsi utama, antara lain:

1. Terdiri dari *marketing automation, sales force automation, customer service automation*.
2. Menggunakan data warehouse untuk penyimpanan data-data historis CRM dan dianalisis menggunakan OLAP, data mining, dan teknik lainnya untuk membantu dalam pengambilan keputusan.
3. Memodelkan dan mengoptimisasi *business process* dengan menghubungkan antara bisnis dan konsumen.

Menurut Hsien, 2009, Peranan data mining sangat penting khususnya pada area pemasaran. Data mining digunakan untuk mendukung kesuksesan CRM (*Customer Relationship Management*). Pan, 2010, melakukan penelitian untuk membuat model proses data mining untuk pemasaran. Model yang dibangun untuk memprediksi konsumen berdasarkan karakteristik demografi, *life-style*, dan perilaku konsumen.

Proses data mining untuk menghasilkan suatu pola sangat tergantung atribut-atribut yang digunakan. Atribut ini berisi data-data yang digunakan untuk melakukan proses mining. Atribut yang sering digunakan untuk melakukan segmentasi konsumen adalah data konsumen, produk, demografi, perilaku konsumen, transaksi, RFMDC, RFM (*Rerency, Frequency Monetary*), dan LTV (*Life Time Value*).

Pada tabel 1 adalah pemetaan algoritma atau teknik data mining yang sering digunakan untuk melakukan segmentasi konsumen antara lain algoritma clustering, classification, dan association.

HASIL DAN PEMBAHASAN

Pemanfaatan algoritma *Association rule* pada segmentasi konsumen

Pemanfaatan algoritma *association rule* pada segmentasi konsumen dilakukan oleh Lin (2012). Dalam papernya dilakukan proses segmentasi konsumen dengan menggunakan tiga langkah, pertama menggunakan CPV (*Customer Pontential Value*) untuk mengukur *customer pontential life value*, kedua membuat matriks CPV, ketiga menggunakan *association rule* mining untuk menjelaskan perilaku konsumen di tiap segmen.

CPV (*Customer Pontential Value*) digunakan untuk mengukur konsumen yang mempunyai nilai potensial untuk menggunakan suatu produk tertentu. Dari hasil pengukuran CPV didapat kriteria seorang customer yang dalam matriks CPV terdapat empat bagian, antara lain kelompok yang mempunyai keinginan kurang yang mempunyai level 6 sampai 7, kelompok yang mempunyai persepsi yang kurang yang mempunyai level 3 sampai 5, kelompok yang terbuka yang mempunyai level 8 sampai 10, dan kelompok yang tertutup yang mempunyai level 1 sampai 2 (Lin, 2012). Penggunaan *association rule* seperti diterangkan diatas untuk menjelaskan karakteristik konsumen pada masing-masing bagian atau kelompok pada matriks dan hasil yang didapat pada kelompok terbuka konsumennya mempunyai karakteristik laki-laki, single, umur antara 20 tahun sampai 29 tahun dan lebih dari 60 tahun. Untuk kelompok tertutup mempunyai karakteristik laki-laki, menikah, umur antara 40 tahun sampai 59 tahun. Pada kelompok yang mempunyai tingkat persepsi yang kurang mempunyai karakteristik single, umur antara 20 tahun sampai 29 tahun dan lebih besar

60 tahun. Dan yang terakhir kelompok yang mempunyai tingkat keinginan yang kurang mempunyai karakteristik perempuan, single, umur antara 20 tahun sampai 39 tahun.

Pemanfaatan algoritma *K-Means Clustering* pada segmentasi konsumen

Proses Segmentasi konsumen terbanyak menggunakan algoritma *K-Means Clustering*, dan hasil menyatakan bahwa segmentasi sukses dan efektif (Ye, 2012). Tidak hanya menggunakan *K-Means Clustering* saja, dalam penelitian lainnya segmentasi konsumen menggabungkan algoritma *K-Means Clustering* dengan algoritma lainnya seperti dengan neural network, C&R Tree, Bayesian network, Decision Tree, C5.0, dan CHAID. Algoritma *K-Means Clustering* mempunyai beberapa keuntungan, antara lain memberikan solusi yang baik untuk masalah clustering untuk objek data yang mempunyai atribut numerik, relatif terukur dan efisien dalam memproses data set yang besar, algoritma ini tidak sensitif terhadap input data, algoritma ini termasuk cepat dalam pemodelan dan mudah dimengerti (Han, 2001).

Hajiha, 2011, melakukan segmentasi konsumen menggunakan algoritma *K-Means Clustering* dengan atribut RFMDC.

RFMDC ini adalah pengembangan dari model RFM (*recency, frecuency, monetary*) yang digunakan untuk melakukan segmentasi konsumen dimana pengembangannya adalah *deversity* dan *continuousness*.

Dalam (Ren, 2011), melakukan segmentasi konsumen dengan menggunakan dua algoritma data mining. Yang pertama dilakukan adalah melakukan segmentasi dengan Neural Netwok untuk mendapatkan kesamaan tipe pada saat sample data digunakan

pada kategori yang berbeda. Hasil dari langkah pertama digunakan sebagai inputan dalam algoritma K-Means Clustering. Penggunaan K-Means lainnya oleh (Khajvand, 2011), dimana untuk segmentasi konsumen menggunakan K-Means Clustering dan Decision Tree untuk melakukan klasifikasi. Langkah pertama yang dilakukan adalah menentukan RFM parameters yang dengan cara membuat skala berdasarkan transaksi, demografi, dan produk. Langkah kedua melakukan proses segmentasi menggunakan K-Means Clustering berdasarkan hasil FRM. Lain halnya dengan (Namvar, 2010), dalam papernya melakukan segmentasi konsumen dengan menggunakan dua langkah. Langkah pertama menggunakan K-Means Clustering untuk menghasilkan cluster berdasarkan data demografi dan langkah kedua menggunakan Neural Network untuk menghitung nilai potensial dari setiap cluster. Sedangkan Li, 2010, melakukan segmentasi konsumen dengan K-Means Clustering untuk mendapatkan empat kelompok konsumen, dan memprediksinya menggunakan C5.0, Neural Network, Chi-squared automation interaction detector (CHAID). Hasilnya adalah konsumen baru dan nilai potensialnya.

Atribut-atribut yang digunakan untuk segmentasi konsumen

Pada saat melakukan clustering akan digunakan atribut-atribut untuk proses

miningnya. Atribut-atribut yang sering digunakan untuk melakukan segmentasi konsumen antara lain data konsumen, faktor demografi, jenis produknya, perilaku konsumen, transaksi konsumen terhadap produk, RFM (*recency, frequency, monetary*), (RFMDC (*recency, frequency, monetary, diversity dan continuousness*), dan LTV (*life time value*), *value information*, dan *behavior information*). Adapun masing-masing atribut yang digunakan dalam melakukan segmentasi konsumen dapat dilihat pada tabel 2.

KESIMPULAN DAN SARAN

Segmentasi konsumen merupakan fundamental pada strategi pemasaran. Untuk mendukung hasil proses pengelompokkan konsumen atau segmentasi konsumen ini maka dukungan data mining sangat berperan penting. Algoritma data mining yang sering digunakan untuk segmentasi adalah K-Means Clustering dan atribut-atribut yang digunakan untuk proses mining pada segmentasi konsumen adalah data konsumen, produk, demografi, perilaku konsumen, transaksi, RFMDC, RFM (*Recency, Frequency Monetary*), dan LTV (*Life Time Value*). Dan penting untuk menggabungkan algoritma clustering dengan beberapa algoritma data mining lainnya seperti Classification, Association, dan CPV matriks untuk mendapatkan nilai potensial dari tiap cluster.

Tabel 1. Matrik Algoritma Data Mining dalam Segmentasi Konsumen

No	Nama Penulis/ Peneliti	Tahun	Clustering				Classification						Association		Lainnya			
			k-Means Clustering	Clustering lainnya	Fuzzy C-Means clustering	SOM-Ward Clustering	Support Vector Machine	Neural Network	Boosted Decision Tree (C4.5)	Decision Tree (C4.5) with boosting	C&R Tree	Bayesian Network	Decision Tree	C 5.0		CHAID	Association Rule	(CASA Algorithm)
1	Lin, Jian-Bang, Liang, Te-Hsin, Lee, Yong-Goo	2012																
2	Ye, Luo, Chiu, Cai, Shi, Z	2012																
3	Li, Ke, Furukawa, Teisuija	2012																
4	Hajira, Ail, Rizki, Reza, Makaveri, Samira S.	2011																
5	Ren, Shoufa, Sun, Qiming, Shi, Yuguang	2011																
6	Chiang-shun, Yan, Yu-chiang	2011																
7	Shi, Yuanjuan, Sun, Zhou, Shuwen, Le, Guanghong	2011																
8	Bonsnjak, Z., Gjelic, O.	2011																
9	Yao, Zhiyuan, Edouard, Thomas, Beck, Baroro	2010																
10	Krijward, Mardoukhi, Taroun, Mohammad Jaber, Naimar, Morbeza, Goharian, M.R.	2011																
11	Naimar, Morbeza, Goharian, M.R.	2010																
12	Yu, Gu, Jianli, Li, Bi, Jianon	2010																
13	Wang, Wei, Fan, Shidong	2010																
14	Yihua, Zhang	2010																
15	Li, Wu, Wu, Xuemai, Xiaobin, Zhang, Feng, Gao, Hu, Huang	2010																
16	Li, Xih-an	2009																
17	Gong, Huiqing, Xia, Chong	2009																
18	Hong, Xu, Gangyi, Qian, Hsieh, Nan-Chien, Chin, Kuo-Chung	2009																
19																		
20																		
21																		

No	Nama Penulis/ Peneliti	Tahun	Atribut
1	Lin, Jian-Bang., Liang, Te-Hsin., Lee, Yong-Goo, 2012	2012	Demografi, Product
2	Ye, Luo., Qiu-ru, Cai., Etl., 2012	2012	Customer: identifikasi konsumen, , metode kontak, waktu akses
			Value Information: Biaya bulanan individu atau perusahaan
			Behavior Information: durasi telepon, waktu telepon
3	Lu, Ke., Furukawa, Tetsuya, 2012	2012	Customer, Transaksi
4	Hajiha, Ali., Radjar, Reza., Malayeri., Samira S., 2011	2011	RFMDC
5	Ren, Shuxia., Sun, Qiming., Shi, Yuguang., 2010	2011	Customer, Customer Account, Pinjaman
6	Chang-shun, Yan, Yu-liang, Shi., Yuan-yuan, Sun, 2011	2011	Customer, Penyakit
7	Zhou, Shuwen., Lei, Guanghong, 2011	2011	Customer
			Data CRM : umur, jenis kelamin, pendidikan, pekerjaan, status pernikahan, pendapatan
			Informasi pembelian konsumen
8	Bonsnjak, Z., Grljevic, O., 2011	2011	Customer, Pinjaman
9	Yao, Zhiyuan., Eklund, Thomas., Back, Barbro, 2010	2010	Demografi, perilaku konsumen
10	Namvar, Morteza., Glolamian, M.R., 2010	2010	RFM (Ferency, frequebcym Monetary), Demografi, LTV (Life Time Value)
11	Yu, Gu., Jiahui, Li, 2010	2010	Customer Value
12	Bi., Jianxin, 2010	2010	Customer Value, product, pangsa pasar
13	Wang, Wei., Fan, Shidong., 2010	2010	Customer, Transaksi
14	Yihua, Zhang, 2010	2010	Customer, Product, Transaksi
15	Li, We., Wu, Xuemei., 2010	2010	Application Source, Number Credit Card
			Frekuensi penggunaan, demografi konsumen
16	Xiao-bin, Zhang., Feng, Gao., Hui, Huang., 2009	2009	Customer
17	Lai, Xin-an, 2009	2009	Customer, Transaksi
18	Gong, Huaping., Xia, Qiong., 2009	2009	Customer, Transaksi
19	Hong, Xu., Gangyi, Qian., 2009	2009	Customer, Transaksi
20	Hsieh, Nan-Chen., Chu, Kuo-Chung., 2009	2009	Customer, Transaksi

DAFTAR PUSTAKA

Bi., Jianxin, 2010, *Research for Customers Segmentation of Medical Insurance Based on K-Means and C&R Tree Algorithms*, Sixth

International Conference on Semantics, Knowledge and Grids.
 Bonsnjak, Z., Grljevic, O., 2011, *Credit Users Segmentation for Improved Customer Relationship Management in Banking*, 6th IEEE International

- Symposium on Applied Computational Intelligence and Informatics.
- Chang-shun, Yan, Yu-liang, Shi., Yuan-yuan, Sun, 2011, *Customer Segmentation Methods Analysis Based on the Support-significant Structure*, Power and Energy Engineering Conference (APPEEC).
- Gong, Huaping., Xia, Qiong., 2009, *Study on Application of Customers Segmentation Based on Data Mining Technology*, 2009 ETP International Conference on Future Computer and Communication.
- Guoxiang, Liu., Zhiheng, Qi., 2013, *Data Mining Applications in Marketing Strategy*, 2013 Third International conference on Intelligenece System Design and Engineering Application Digital Object Identifier: 10.1109/ISDEA.2012.397 Publication Year: 2013 , Page(s): C4.
- Hajiha, Ali., Radjar, Reza., Malayeri., Samira S., 2011, *Data Mining Application for Customer Segmentation Based on Loyalty: An Iranian Food Industry Case Study*, Industrian Engginering and Management Engineering Management IEEM
- Han, Jia-wei, Micheline Kamber, 2001, *Data Mining: Concepts and Techniques*, Higher Education Press.
- Hong, Xu., Gangyi, Qian., 2009, *Data Mining in Market Segmentation and Tariff Policy Design: a Telecommunication Case*, 2009 Asia-Pacific Conference on Information Processing, IEEE DOI 10.1109/APCIP.2009.90.
- Hsieh, Nan-Chen., Chu, Kuo-Chung., 2009, *Enhancing Consumer Behavior Analysis by Data Mining Techniques*, International Journal of Information and Management Sciences vol.20 (2009), 39-53.
- Khajvand, Mahboubeh., Tarokh, Mohammad Jafar., 2011, *Analyzing Customer Seqmentation Based on Customer Value Componens (Case Study: a Private Bank)*, Journal of Industrial Engineering, University of Tehran, Spesial Issue, 2011, PP.79-93.
- Lai, Xin-an, 2009, *Segmentation Study on Enterprise Customers Based on Data Mining Technology*, 2009 First International Workshop on Database Technology and Applications.
- Li, We., Wu, Xuemei., 2010, *Credit Card Customer Segementation and Target Marketing Based on data Mining*, 2010 International Conference on Computational Intelligence and Security, DOI 10.1109/CIS.2010.23.
- Lin, Jian-Bang., Liang, Te-Hsin., Lee, Yong-Goo, 2012, *Mining Important Association Rules on Diferent Customer Potential Value Segments for life Insurance Database*, 2012 IEEE International Conference on Granular Computing.
- Lu, Ke., Furukawa, Tetsuya, 2012, *A Framework for Segmentation Customers based on Probability Density of Transaction Data*, IIAI International Conference on Advanced Applied Informatics.
- Namvar, Morteza., Glolamian, M.R., 2010, *A Two Phase Clustering Method for Intelligent Customer Segmentation*, 2010 International Conference on Intelligence Systems, Modelling and Simulation.
- Pan, Ding , 2010, *Data Mining Process Model for Marketing and CRM*, Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010.
- Ren, Shuxia., Sun, Qiming., Shi, Yuguang., 2010, *Customer Segmentation of Bank Based on Data*

- Warehouse and Data Mining, Information Management and Engineering (ICIME).*
- Wang, Wei., Fan, Shidong., 2010, *Application of Data Mining Technique in Customers Segmentation of Shipping Enterprises*, Database Technology and Application (DBTA) 2010 International Workshop.
- Xiao-bin, Zhang., Feng, Gao., Hui, Huang., 2009, *Customer-churn Research Based on Customer Segmentation*, 2009 International Conference on Electronic Commerce and Business Intelligence.
- Yao, Zhiyuan., Eklund, Thomas., Back, Barbro, 2010, *Using SOM-Ward Clustering and Predictive Analytics for Conducting Customer Segmentation*, 2010 IEEE International Conference on Data Mining Workshops.
- Ye, Luo., Qiu-ru, Cai., Etl., 2012, *Telecom Customer Segmentation with K-means Clustering*, the 7th International Conference on Computer Science & Education (ICCSE 2012).
- Yihua, Zhang, 2010, *Customer Segmentation Based on Data Mining in Mobile-Communications Industry*, The 5th International Conference on Computer Science and Education.
- Yu, Gu., Jiahui, Li, 2010, *Hierarchy Analysis and Optimization of Customer Value Based on Data Mining Technology*, 2010 International Conference on Intelligent Computation Technology and Automation.
- Zhang, Guazhen., Zhou, faming, etl., 2008, *Knowledge creation in marketing based on data mining*, Intelligent Computation Technology and Automation (ICICTA), 2008 International Conference on **Page(s):** 782 – 786.
- Zhou, Shuwen., Lei, Guanghong, 2011, *Application of Data Mining Technology in Membership supermarket's Customer Segmentation*, 2011 International Conference on Business Computing and Global Informations, IEEE DOI 10.1109/BCGIN.2011.53