

## PEMBENTUKAN TARGET PASAR BERDASARKAN *DATA STREAM* TRANSAKSI KARTU KREDIT (*CLUSTERING DAN ASSOCIATION RULE*) PADA PT BANK BUKOPIN

*ESTABLISHMENT OF TARGET MARKET BASED ON THE DATA STREAM OF CREDIT CARD TRANSACTIONS (CLUSTERING AND ASSOCIATION RULE) AT PT BANK BUKOPIN*

Muhammad Riza<sup>\*1</sup>, Kudang Boro Seminar<sup>\*\*</sup>, dan Agus Maulana<sup>\*\*\*</sup>

<sup>\*</sup>) Sekolah Bisnis, Institut Pertanian Bogor  
Jl. Raya Pajajaran, Bogor 16151

<sup>\*\*</sup>) Departemen Teknik Mesin dan Biosistem, Fakultas Teknologi Pertanian, Institut Pertanian Bogor  
Kampus IPB Darmaga, Jl. Lingkar Akademik, Jawa Barat 16680

<sup>\*\*\*</sup>) Universitas Dr Sutomo  
Jl. Semolowaru No.84, Surabaya 60118

**Abstract:** *The purpose of research is to analyze the formation of the target market customer segmentation based on job characteristics, income, education, age, region of origin, and patterns of credit card merchant. The data were analyzed using two data mining techniques of clustering with K-Means and Association Rule Mining (ARM) and supported by Apriori and Random Sampling technique with the Slovin formula and Principal Component Analysis (PCA). The clustering tests in 10 replications on the sampling of 350 clients supported by PCA produced the three best clusters that had the silhouette value close to 1 i.e. 0.39 to 0.40. Meanwhile, the ARM Testing with Apriori using a minimum support of 1% and a minimum confidence of 40% produced two patterns of credit card merchant transactions. In the first pattern, when the Hotel merchant type (hhl) was transacted, the Restaurant merchant type (RRT) was also transacted, and in the second pattern, if the Service Station merchant type (RSS) was transacted, the Restaurant merchant type (RRT) was also transacted. The three clusters and two types of merchant patterns obtained can generate inputs for the company to identify its potential customers based on the characteristics of the target customers by connecting them to the merchant type pattern frequently used.*

**Keywords:** *credit card, data mining, clustering, ARM, data stream*

**Abstrak:** Tujuan penelitian adalah untuk menganalisis pembentukan target pasar segmentasi nasabah dengan karakteristik pekerjaan, penghasilan, pendidikan, usia, asal wilayah dan pola merchant kartu kredit. Teknik analisis data menggunakan dua teknik data mining, yaitu *clustering* dengan K-Means dan *Association Rule Mining* (ARM) dengan Apriori didukung juga dengan teknik Random Sampling dengan rumus Slovin serta *Principal Component Analysis* (PCA). Hasil pengujian clustering 10 kali percobaan pada sampling 350 nasabah dengan dukungan PCA menghasilkan tiga buah cluster terbaik yang memiliki nilai silhouette mendekati 1 sebesar 0,39–0,40. Sementara hasil pengujian ARM dengan Apriori menggunakan minimum support 1% dan minimum confidence 40% menghasilkan dua buah pola transaksi merchant kartu kredit, yaitu pola pertama merchant jenis Hotel (HHL) ditransaksikan maka merchant jenis Restaurant (RRT) ikut ditransaksikan dan pola kedua jika merchant jenis Service Station (RSS) ditransaksikan maka merchant jenis Restaurant (RRT) akan ikut ditransaksikan. Tiga cluster dan dua pola jenis merchant yang terbentuk menghasilkan masukan bagi perusahaan untuk mengidentifikasi target nasabah potensial berdasarkan karakteristik nasabah dengan menghubungkannya pada pola jenis merchant yang sering ditransaksikan.

**Kata kunci:** *kartu kredit, data mining, klasterisasi, ARM, data stream*

---

<sup>1</sup> Corresponding author:  
Email: [mriiss.cold@gmail.com](mailto:mriiss.cold@gmail.com)

## PENDAHULUAN

Bidang pemasaran memegang peranan yang cukup vital dalam setiap usaha bisnis. Pengembangan pemasaran dalam mencapai tujuan bisnis seperti: bertahan hidup, memperoleh keuntungan dan berkembang (Kotler dan Armstrong, 2008) merupakan bagian terpenting dalam pengembangan bisnis perusahaan. Memahami bagaimana konsumen menanggapi tawaran perusahaan, memilih dan memutuskan produk (Bayer, 2010) yang digunakan merupakan kunci sukses dalam mencapai tujuan perusahaan. Oleh karena itu, perlu memahami bagaimana segmentasi dan pola bertransaksi konsumen agar menjadi bahan pertimbangan manajemen perusahaan dalam merencanakan strategi pemasaran (Mabrur dan Lubis, 2012).

Produk bisnis kartu kredit pada perbankan merupakan salah satu bidang usaha bisnis konsumen atau ritel yang terus mengalami pertumbuhan data. Kartu kredit memiliki peranan besar bagi perseroan sebagai produk yang memberikan kemudahan dalam bertransaksi dan menjadi perhatian besar dalam pertumbuhan perusahaan. Bank Bukopin mulai meluncurkan kartu kredit sejak tahun 2009 terdiri atas Visa dan Master Card. Di tahun 2014 realisasi kinerja bisnis produk konsumen kartu kredit meningkat lebih besar dari produk lainnya sebesar 41,8% dan telah memberikan kontribusi besar pada *fee based income* perusahaan. Peningkatan kinerja bisnis produk *consumer* pada pertumbuhan jumlah pemegang kartu tidak diimbangi dengan meningkatnya pertumbuhan transaksi dari 5,6 kali rata-rata jumlah transaksi setiap orangnya pada tahun 2013 menjadi 4,8 kali rata-rata transaksi per-orangnya pada tahun 2014. Penurunan pertumbuhan dibawah 20% juga terjadi pada penjualan total rata-rata *merchant* di setiap nasabah. Peningkatan terjadi pada biaya pemasaran dari tahun 2012 sampai dengan tahun 2014. Namun, rasio pertumbuhan transaksi dari banyaknya pemegang kartu kredit mengalami penurunan. Oleh sebab itu, menurut Annie dan Kumar (2012) tantangan terbesar perusahaan dalam pengumpulan atau penggalan informasi penting (*data mining*) pada historical data stream transaksi kartu kredit berdasarkan segmen nasabah diharapkan lebih tepat sasaran dalam menentukan target pasar (Brachman, 1996), menekan biaya pemasaran, meningkatkan penggunaan kartu dan meningkatkan *fee based income* perusahaan.

Tiga kelompok variabel yang digunakan secara luas dalam melakukan segmentasi pasar konsumen menurut Jaman (2012), yaitu a) segmentasi geografis, yaitu karakteristik nasabah berdasarkan unit-unit geografis; b) segmentasi demografis, menurut Krishnamurthi (2007) dibagi menjadi kelompok-kelompok seperti: pekerjaan, penghasilan, pendidikan dan usia; c) segmentasi psikografis dibagi menjadi kelompok berbeda berdasarkan gaya hidup atau kepribadian dan nilai (Kotler dan Lane, 2008).

Penelitian ini bertujuan menganalisis karakteristik konsumen dalam menggunakan kartu kredit berdasarkan faktor demografis (Garver *et al.* 2012; Davies *et al.* 2009) dengan karakteristik pekerjaan, penghasilan, pendidikan, usia, asal wilayah, geografis nasabah dan identifikasi jenis *merchant* kartu kredit berdasarkan pola transaksi. *Data mining* dapat diterapkan dalam berbagai data stream yang pada penerapannya akan menjadi dua kategori (Liu *et al.* 2009), yaitu *transactional data stream* yang mencatat kegiatan interaksi antara dua entitas dan *measurement data stream*, mencatat hasil pengukuran sebuah entitas. *Data stream* terdiri atas data berseri pada waktu tertentu dan memiliki kolom yang berisi data-data berupa transaksi yang telah diidentifikasi memiliki nilai untuk dapat diolah.

Teknik *clustering mining* dengan *k-Means* algorithm pada *data mining* dapat digunakan sebagai solusi untuk analisis segmentasi berdasarkan faktor demografis dan geografis nasabah dalam mengelompokkan jumlah data stream transaksi berukuran besar dengan waktu komputasi yang relatif cepat dan efisien (Alfina *et al.* 2012). *Association Rules Mining* (ARM) dengan Apriori Algorithm digunakan sebagai identifikasi frekuensi pola *merchant* yang sering digunakan sebagai tempat bertransaksi. Penelitian dengan menggabungkan hasil *clustering* dan ARM ini diharapkan dapat membantu dalam pengambilan keputusan manajemen dan dijadikan sebagai rekomendasi strategi menetapkan target nasabah dalam bidang pemasaran sehingga dapat meningkatkan jumlah transaksi kartu kredit.

Merujuk pada penelitian Alfi Wijaya tahun 2006 dengan rumusan alternatif strategi pemasaran sesuai segmentasi pasar, faktor memilih dan preferensi nasabah terhadap bank syariah. Persamaannya membuat rumusan alternatif strategi pemasaran sesuai dengan

segmentasi pasar, perbedaannya pada penelitian ini tidak menggunakan teknik data *mining clustering* dan ARM untuk pengolahan data.

Rujukan penelitian berikutnya oleh Cahyo (2002) bertujuan menganalisis keterkaitan profil demografis dan psikografis dengan pola penggunaan kartu menggunakan analisis *cross tabulation* dengan alat bantu SPSS dan uji Chi-Kuadrat. Persamaan pada penelitian ini menganalisis konsumen berdasarkan faktor demografis dan pola penggunaan kartu sedangkan perbedaannya pada penggunaan metode data *mining* dan menggunakan alat bantu *software orange mining*.

Penelitian lainnya terkait dengan segmentasi pasar pada penelitian Robson *et al.* (2013) dan penelitian terkait kartu kredit dengan menggunakan teknik data *mining* oleh Zheng (2009). Perbedaannya adalah variabel yang digunakan dalam identifikasi konsumen menggunakan faktor psikografis dan menggali data nasabah berdasarkan metode-metode pada data *mining* menggunakan metode *decision tree algorithm*, *neural networks*, *X-Mean* dan *Apriori*.

Proses pembentukan target pasar kartu kredit menggunakan dua teknik data *mining* dalam menentukan segmentasi nasabah dan identifikasi pola *merchant* yang sering digunakan. Identifikasi segmentasi nasabah berdasarkan karakteristik pekerjaan, penghasilan, pendidikan, asal wilayah, usia menggunakan teknik *clustering* didukung dengan reduksi dimensi variabel menggunakan *Principal Component Analysis* (PCA) dan penilaian *Silhouette Coefficient* (SC) yang mendekati nilai 1. Hasil seleksi nasabah berdasarkan *random sampling* rumus slovin digunakan sebagai *pre-processing* sebelum dilakukan proses penentuan pola transaksi jenis *merchant* dengan ARM. Penggabungan dua teknik data *mining* ini bertujuan dalam memberikan rekomendasi bagi perusahaan menentukan target pasar dan identifikasi keterkaitan pola antar jenis *merchant*.

## METODE PENELITIAN

Jenis data yang digunakan dalam penelitian ini adalah data sekunder harian berupa data time series yang diambil langsung dari *Data Ware House* (DWH) transaksi kartu kredit sejak 1 Januari 2010 sampai dengan 31 Desember 2014. Data-data yang digunakan dalam penelitian ini adalah data *profile* nasabah, data *merchant* dan *historical* data transaksi yang telah tercatat

pada transaksi kartu kredit. Jumlah data nasabah yang tercatat sebesar 272.067 yang tersebar diseluruh kota di Indonesia berdasarkan karakteristik demografis dan geografis yang terdiri dari nomor nasabah, pekerjaan, penghasilan, pendidikan, umur, kota asal. Data profil nasabah terdiri atas 26 jenis pekerjaan, 7 rentang penghasilan, 7 jenis pendidikan, 6 daftar kisaran usia, dan 34 daftar wilayah. *Historical* data transaksi yang diperoleh selama 5 tahun berkisar 13.765.577 data berdasarkan 16 group *merchant* transaksi. Alat yang digunakan sebagai pengolahan data adalah *Orange Mining*, dan database SQL Server.

Penelitian menggunakan metode kuantitatif dengan analisis pada data-data numerik. Hal ini, bertujuan mendapatkan gambaran yang jelas terhadap data yang diperoleh. Dengan demikian, dapat menjadi informasi baru yang bisa digunakan terhadap permasalahan yang sedang diteliti.

Penelitian menggunakan gabungan dua teknik data *mining* dengan teknik *clustering* dan ARM. ARM dapat digunakan pada *market basket transaction* (Sim, 2009). ARM mencari hubungan antar item dalam suatu data set yang ditentukan (Han, 2001), meliputi dua tahap menurut Ulmer dan David (2002) yaitu pertama mencari kombinasi yang paling sering terjadi dari suatu itemset atau himpunan dari item-item yang muncul bersama-sama (Kotsiantis dan Kanellopoulos, 2006), kedua adalah mendefinisikan *condition* dan *result* untuk *conditional association rule*. *Clustering* merupakan proses pengelompokan sekumpulan objek fisik atau abstrak kedalam kelas atau objek yang serupa (Vaishali, 2014; Han & Kamber, 2006). *Cluster* dapat disebut juga sebagai koleksi objek data yang memiliki kesamaan satu sama lain dalam kelompok yang sama dan berbeda dengan objek dikelompok lain.

Dalam data *mining* terdapat permasalahan yaitu salah satunya adalah banyaknya data, sehingga terdapat kemungkinan terjadinya duplikasi data (Tan *et al.* 2006). Salah satu solusinya adalah menggunakan anti-*monotonic*. Anti-*monotonic* mempunyai tujuan mengurangi masalah kompleksitas komputasi dengan cara mengurangi jumlah kandidat menggunakan teknik pemangkasan, seleksi, mengurangi jumlah perbandingan atau disebut juga *pre-processing*. Anti-*monotonic* dapat digunakan dalam *apriori principle*, dengan nilai *support* yang tidak pernah melebihi dari *minimum support* yang telah ditentukan sehingga diasumsikan bahwa item dalam

setiap transaksi disimpan dan diurutkan dalam urutan leksikografis, hal ini digunakan untuk mengadaptasi algoritma jika terjadi kasus di mana *database* disimpan secara normalisasi (Srikant, 1996; Gürel, 2008).

Proses pembentukan segmentasi nasabah berdasarkan transaksi pada suatu *merchant* menggunakan teknik *clustering k-Means* dan ARM dengan apriori, dimana kedua teknik berperan besar sebagai pengolahan data berbasis sistem informasi. Dalam menyeleksi data nasabah dan transaksi dibutuhkan beberapa tahapan pre-processing sebelum menjalankan data *mining*, yaitu sebagai berikut:

- 1) Data yang diperoleh dari data nasabah dan DWH Historical transaksi perusahaan dalam bentuk tabel dikumpulkan dan dimasukkan kedalam tabel baru SQL Server 2000. Tabel baru nasabah tersebut terdiri atas dua tabel, yaitu *t\_nasabah* terdiri dari kode nasabah, kode penghasilan, kod epekerjaan, kode pendidikan, kode usia, dan kode wilayah. *t\_transaksi* terdiri dari tanggal, kode nasabah, kode kartu, kode *merchant*
- 2) Proses seleksi dilakukan dengan pemangkasan data menggunakan fungsi IF dan ELSE untuk mengetahui apakah data-data nasabah tersebut memiliki jumlah transaksi atau tidak. Jika data nasabah tidak memiliki data transaksi maka data nasabah akan dieliminasi sehingga data hanya terdiri dari data yang memiliki transaksi.

Pemangkasan data nasabah dibutuhkan dengan mengeliminasi data pasif atau tidak pernah bertransaksi. Dengan menggunakan kondisi fungsi IF dan ELSE, pemangkasan bertujuan menseleksi data sehingga dari 272.067 akan dipangkas menjadi 95.906 data. Selanjutnya pemangkasan data nasabah kembali dilakukan dengan menggunakan kriteria pengecekan transaksi *merchant* yang terdaftar dalam *group merchant*. Setelah itu dilakukan pengecekan valid-nya kriteria nasabah berupa kode pekerjaan, kode penghasilan, kode pendidikan, kode wilayah dan kode usia. Dari hasil pemangkasan data berdasarkan *merchant* dan kriteria nasabah akan memangkas menjadi 1938 data nasabah.

- 3) Seleksi dilanjutkan dengan menyeleksi data populasi dengan menggunakan metode random sampling rumus slovin. Pemangkasan data menggunakan *random sampling* bertujuan mengurangi kompleksitas dan besarnya jumlah data populasi. Populasi penelitian dari 1 Januari 2010

sampai dengan 31 Desember 2014 memiliki profil pekerjaan, penghasilan, pendidikan, usia, wilayah dan berjumlah 1938 data nasabah aktif yang telah bertransaksi pada mesin EDC dalam negeri. Besarnya sampel data nasabah pada penelitian ini berdasarkan rumus Slovin:

$$n = \frac{N}{(N(d)^2 + 1)} \rightarrow = \frac{1938}{(1938(0.05)^2 + 1)} = 331,56$$

Keterangan: n (Sampel); N (Populasi); d (Nilai presisi 95% atau sigma = 0,05 (tingkat kesalahan yang dikehendaki adalah 5%))

Selanjutnya, *random sampling* dilakukan dengan cara mengurutkan data dari nilai terbesar sampai dengan yang terkecil sebanyak 350 data nasabah. *Preprocessing* juga berfungsi dalam mengkonversi data sehingga bisa diolah pada proses klasterisasi *k-Means* dan proses ARM. Pertama adalah pengubahan data variabel pekerjaan, penghasilan, pendidikan, umur dan wilayah dengan pengkodean ordinal atau skala. Data *merchant* di groupkan menjadi sebuah kode *group merchant* dengan mengubah nilai return menjadi "1= transaksi dan "?" = tidak bertransaksi. Tabel hasil konversi variabel menjadi pengkodean skala dapat dilihat pada Tabel 1.

*Pre-processing* data menggunakan *Principal Componen Analysis* (PCA) bertujuan menyederhanakan variabel yang diamati dengan cara menyusutkan/mereduksi dimensi variabel (Helbich *et al.* 2013) dengan menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. PCA memproyeksikan citra kedalam *subspace* dalam menentukan sistem koordinat baru dari *dataset* nasabah. Reduksi data menggunakan PCA dapat mengurangi dimensi dan menyederhanakan informasi. Konversi hasil reduksi menjadi PC1, PC2 dapat dilihat pada Tabel 2.

## HASIL

### Segmentasi Nasabah dengan *Clustering K-Means*

Pembentukan segmentasi nasabah berdasarkan pengolahan data *pre-processing* menghasilkan 350 data nasabah dan variabel yang lebih sederhana PC1 dan PC2. *Clustering k-Means* dengan menggunakan

alat bantu *software orange tools* membutuhkan pengaturan identifikasi awal jumlah *cluster* yang akan dibentuk mulai dari 2 sampai dengan 8 *cluster* dan pengaturan maksimal iterasi (penghitungan jarak antar item) yang akan dijalankan maksimal sebanyak 10 perulangan. Proses iterasi dilakukan secara berulang sampai tidak adanya perubahan atau perpindahan item pada *cluster* yang terbentuk. Hasil dari *silhouette* terbaik menunjukkan nilai yang paling mendekati nilai 1. Selama 10 kali percobaan menghasilkan k=3 dan rentang score yang didapatkan adalah 0,39 dan 0,40.

Hasil score *k-Means* dan *scatter plot* menghasilkan *score* yang tertinggi adalah k=3 yang berarti *cluster*

berdasarkan *silhouette coefficient* terbaik adalah 3 dengan *score* = 0,40, atau paling mendekati 1. Namun, terdapat kekurangan pada penggunaan alat bantu *software orange*, yaitu tidak diketahuinya *variable* acak *cluster* yang dipilih pertama kali sebagai *candidate* awal *cluster* dan tidak ditampilkannya berapa banyak iterasi yang telah dilakukan untuk mendapatkan masing-masing nilai k terbaik tersebut. Kelebihannya berdasarkan 10 kali percobaan, hasil *cluster* tetap stabil berdasarkan hasil *silhouette* yang memiliki rentang tidak terlalu berbeda antara 0,39 sampai dengan 0,40. Hasil dari pembentukan *scatter plot* dan percobaan *k-Means* dapat dilihat pada Gambar 1.

Tabel 1. Tabel Konversi variabel pekerjaan, penghasilan, pendidikan, usia, wilayah dan merchant

Variabel	Kode	Variabel	Kode	Variabel	Kode
<b>Nama Wilayah</b>		Sumatera Barat	13	Polisi	22
Aceh	11	Sumatera Selatan	16	Paramedis	23
Bali	51	Sumatera Utara	12	Petani/Nelayan	24
Banten	36	Surabaya	37	Wartawan	25
Bengkulu	17	Ujung Pandang	73	Pegawai Yayasan	26
Dki Jakarta	31	Yogyakarta	34	<b>Penghasilan</b>	
Jambi	15	<b>Pekerjaan</b>		< 1 juta	1
Jawa Barat	32	Pegawai Negeri/PNS	1	1–5 jt	2
Jawa Timur	35	TNI/ABRI/Militer	2	5–10 jt	3
Jawatengah	33	Pegawai Swasta/Karyawan-Swasta	3	10–25 jt	4
Kalimantan Barat	61	Profesional	4	25–50 jt	5
Kalimantan Selatan	63	Wiraswasta/Wirausaha	5	50–100 jt	6
Kalimantan Tengah	62	Pensiun/Pensiunan	6	> 100 jt	7
Kalimantan Timur	64	Akuntan	7	<b>Pendidikan</b>	
Kalimantan Utara	65	Dokter	8	SD	1
Kep Bangka Belitung	19	Engineer/Teknisi	9	SMP	2
Kep Riau	14	Guru/Dosen	10	SMA	3
Lampung	18	Ibu Rumah Tangga	11	D3	4
Maluku	81	Konsultan	12	S1	5
Nusa Tenggara Barat	52	Pegawai Koperasi	13	S2	6
Nusa Tenggara Timur	53	Mahasiswa/Pelajar	14	S3	7
Palembang	14	Notaris	15	<b>Usia</b>	
Papua	92	Pegawai BUMN/BUMD	16	Dibawah < 18 Tahun	1
Riau	14	Pengacara	17	18–24 Tahun	2
Sulawesi Barat	76	Pekerja Informal	18	25–29 Tahun	3
Sulawesi Selatan	73	Pekerja Seni	19	30–39 Tahun	4
Sulawesi Tengah	72	Pekerja Sosial/Lsm	20	40–55 Tahun	5
Sulawesi Tenggara	74	Pejabat Negara	21	> 55 Tahun	6
Sulawesi Utara	71				

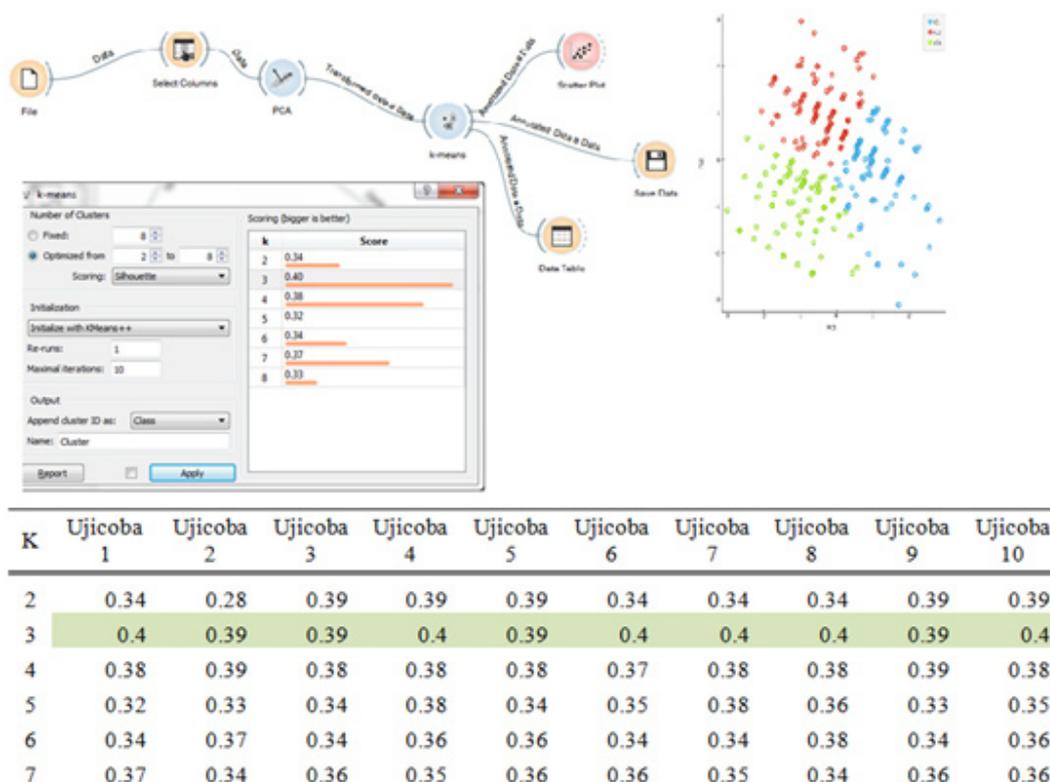
**Pola Merchant dengan Association Rule Mining (ARM)**

Identifikasi pola jenis merchant kartu kredit yang diteliti harus berdasarkan aturan *minimum support* dan *minimum confidence*. *Minimum support* digunakan untuk menguji pembentukan hasil data yang diperoleh berdasarkan 1%, 5%, 20%, 40% dan 50% sehingga dapat menjadikannya sebagai aturan dalam penentuan *minimum support frequently itemsets* (Tabel 3).

Hasilnya menunjukkan jumlah data yang semakin besar tidak mendukung semakin besarnya *minimum support* kemunculan merchant. Semakin besar jumlah data maka akan semakin kecil nilai *minimum support* yang akan dihasilkan (Tan *et al.* 2006). Pada *minimum support* diatas 20% menghasilkan hanya 1 jenis merchant saja yang dominan dan sering digunakan. Oleh karena itu, frekuensi digunakannya merchant setiap tahunnya lebih menunjukkan hubungan antar merchant menggunakan *minimum support* 1%.

Tabel 2. Tabel konversi hasil reduksi PCA data nasabah

Nasabah	Pekerjaan	Penghasilan	Pendidikan	Wilayah	Usia	PC1	PC2
9989042	1	2	5	31	6	0.3197	0.830032
914111000	1	3	7	31	6	0.640042	2.231384
225140000	1	3	3	33	5	-0.11337	-0.88818
509557001	1	3	5	35	4	1.294559	-0.34784
28392011	1	4	5	35	6	-0.5743	1.056101
76868008	1	2	3	73	5	-0.6512	-2.43649
869263000	1	3	5	31	6	-0.0804	1.011413
6955010	1	2	4	63	5	-0.05649	-1.48477
7179037	1	6	3	52	5	-1.7592	-0.99333
22011022	1	2	5	14	4	2.187078	0.188419
1410495001	1	3	5	35	5	0.560181	0.263442
7090037	1	6	3	52	5	-1.7592	-0.99333
1706043	1	6	5	31	5	-0.54634	0.944278
1197168001	1	3	5	31	5	0.653974	0.400135
3211037	1	5	5	52	4	0.095729	-0.56602



Gambar 1. Hasil percobaan k-Means menggunakan tools orange

Penelitian ini menggunakan pengujian dengan nilai *minimum confidence* sebesar 40% yang didapatkan berdasarkan asumsi tiga segmentasi konsumen melalui *clustering k-Means* (Tabel 4). Apabila tiga segmentasi berdasarkan karakteristik konsumen maka memberikan kesempatan sebesar 33,33% dan dibulatkan menjadi 40% terhadap kemungkinan terjadinya transaksi kartu kredit pada keseluruhan nasabah yang diteliti. Oleh karena itu penelitian ini dapat di tetapkan *minimum confidence* ARM sebesar 40%, berdasarkan tiga segmentasi nasabah yang terbentuk.

Hasil identifikasi *role* menggambarkan keterkaitan antar *merchant* yang memenuhi kriteria *minimum support* 1% dan *minimum confidence* 40% adalah sebanyak 2 *role*, yaitu (1) HHL → RRT. Artinya, jika (*antecedent merchant* HHL ditransaksikan maka (*consequent merchant* RRT akan ikut ditransaksikan yang memiliki nilai kepercayaan sebesar 54% dan didukung oleh 1% dari seluruh data yang merupakan kombinasi keduanya. (2) RSS → RRT. Artinya, jika (*antecedent merchant* RSS ditransaksikan maka (*consequent merchant* RRT akan ikut ditransaksikan yang memiliki nilai kepercayaan sebesar 60% dan didukung oleh 1% dari seluruh data yang merupakan kombinasi keduanya. Dari hasil pengujian juga menunjukkan 37 *role* tingkat *confidence* 100%, 75% dan 50% pada suatu *merchant* namun hanya memiliki tingkat kepercayaan rentang 0,1–1%. Sehingga dapat disimpulkan *role* terbaik yang yang diambil adalah 2 *role* dengan menghasilkan 2 itemset diatas *minimum support* 1% dan *minimum confidence* sebesar 40%

### Pembentukan Target Pasar berdasarkan Segmentasi Nasabah (*Clustering*) dan Pola *Merchant* (ARM)

Hasil dari segmentasi nasabah yang telah diolah menggunakan teknik *Clustering* dan identifikasi frekuensi pola *merchant* yang terbentuk menggunakan teknik ARM menjadikannya sebagai target pasar berdasarkan pengelompokkan segmen nasabah potensial terhadap kecenderungan jenis *merchant* kartu kredit yang akan digunakan. Hasil ini ditujukan sebagai kebutuhan pengembangan kegiatan promosi dan target pemasaran.

Hubungan pada pola *merchant* pertama dan 3 segmen nasabah menghasilkan jenis *merchant* hotel (HHL) dan restaurant (RRT) memberikan informasi bahwa *cluster* pertama memiliki kriteria pekerjaan PNS, penghasilan Rp10–25 juta, pendidikan SMA, umur 40–55 tahun dan asal Sulawesi Selatan. Pada *cluster* kedua memiliki segmen pekerjaan PNS, penghasilan Rp25–50 juta, pendidikan S1, umur 40–55 tahun, asal DKI Jakarta, Jawa Tengah dan Banten. Pada *cluster* ketiga memiliki karakteristik pekerjaan PNS, berpenghasilan Rp5–10 juta, pendidikan S1, umur 30–39 tahun dan asal dari Riau, Jawa Barat, Jawa Tengah, Banten.

Tabel 3. Hasil data uji minimum support dari *frequently itemset*

Tahun	Jumlah	Min Sup > 1%	Min Sup >5%	Min Sup >20%	Min Sup >40%	Min Sup >50%
2010	48	3 <i>merchant</i> (tel,rrt,rbs)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	0 <i>merchant</i>
2011	74	4 <i>merchant</i> (tel,rss,rfc,rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)
2012	140	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)
2013	266	2 <i>merchant</i> (cdm,rrt)	1 <i>merchant</i> ser (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)
2014	177	8 <i>merchant</i> (tel,rss,hhl,rrt, ros, rps,rbs,roe)	3 <i>merchant</i> (rtt,rss,hhl)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)
ALL	705	4 <i>merchant</i> (hhl, tel, rrt, rss)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)	1 <i>merchant</i> (rrt)

Sementara itu, hubungan pada pola *merchant* kedua dan pada 3 segmen nasbaah menghasilkan jenis *merchant service station* (RSS) dan restaurant (RRT) memberikan informasi bahwa *cluster* pertama berpenghasilan pekerjaan PNS, Rp1–25 juta, pendidikan SMA umur 40–55 tahun dan asal Sulawesi selatan. Pada *cluster*

kedua berpenghasilan pekerjaan PNS, diatas Rp10–50 juta, pendidikan S1–S2, umur 40–55 tahun dan asal DKI Jakarta. Pada *cluster* ketiga pekerjaan PNS, berpenghasilan Rp1–5 juta, pendidikan SMA & S1, umur 25–29 tahun asal Riau, Jawa Barat, Palembang, DKI Jakarta, Yogyakarta.

Tabel 4. Hasil *minimum support* 0,1% dan *minimum confidence* 40%

<i>Itemset</i>	<i>Antecedent</i>	<i>Consequent</i>	<i>Minimum Support (%)</i>	<i>Minimum Confidense (%)</i>	<i>Count</i>
2	HHL	RRT	1	54	14
2	RSS	RRT	1	60	9
4	RPS,RBS, ROE	RRT	0,1	100	1
4	RRT,HHL, RBS	RSS	0,1	100	1
4	RRT, RBS, ROE	RPS	0,1	100	1
4	RRT,RPS, RBS	ROE	0,1	100	1
4	RRT, RPS, ROE	RBS	0,1	100	1
4	RSS, HHL, RBS	RRT	0,1	100	1
4	RSS, RRT, HHL	RBS	0,1	100	1
4	RSS, RRT,RBS	HHL	0,1	100	1
3	HHL,RBS	RSS	0,1	100	1
3	HHL,RBS	RRT	0,1	100	1
3	RBS,ROE	RPS	0,1	100	1
3	RBS,ROE	RRT	0,1	100	1
3	RPS,RBS	RRT	0,1	100	1
3	RPS,RBS	ROE	0,1	100	1
3	RPS,ROE	RBS	0,1	100	1
3	RPS,ROE	RRT	0,1	100	1
3	RRT,ROE	RPS	0,1	100	1
3	RRT,ROE	RBS	0,1	100	1
3	RSS,HHL	RRT	0,1	100	1
3	RSS,HHL	RBS	0,1	100	1
3	RSS,RBS	HHL	0,1	100	1
3	RSS,RBS	RRT	0,1	100	1
3	TEL,HHL	RRT	0,1	100	1
3	TEL,RBS	RRT	0,1	100	1
2	HHL,RBS	RSS,RRT	0,1	100	1
2	RBS,ROE	RRT,RPS	0,1	100	1
2	RPS,RBS	RRT,ROE	0,1	100	1
2	RPS,ROE	RRT,RBS	0,1	100	1
2	RRT,ROE	RPS,RBS	0,1	100	1
2	RSS,HHL	RRT,RBS	0,1	100	1
2	RSS,RBS	RRT,HHL	0,1	100	1
2	RBS	RRT	0,1	50	3
2	RCP	RRT	0,1	50	1
2	RFC	TEL	0,1	50	1
2	RFC	RHA	0,1	50	1
2	RHA	RFC	0,1	50	1
2	RPS	RRT	0,1	75	3

## Implikasi Manajerial

Hasil penelitian pembentukan target pasar kartu kredit berdasarkan data stream transaksi kartu kredit di PT Bank Bukopin dibutuhkan beberapa hal yang menjadi prioritas manajemen dalam menetapkan strategi peningkatan transaksi kartu kredit, yaitu (1) meningkatkan kerjasama bisnis pada jenis *merchant* hotel, *service station*, dan *restaurant*; (2) meningkatkan kegiatan promosi dengan menghubungkan diskon atau *cashback* pada jenis *merchant* hotel-restaurant dan *merchant service station-restaurant* agar tepat sasaran dan dapat mengefisienkan biaya promosi; (3) identifikasi perubahan segmen nasabah dan pola persebaran transaksi pada *merchant* secara periodik tahunan; (4) penggunaan database Data Warehouse transaksi kartu kredit dapat terus digali manfaatnya dengan teknik data *mining* dalam meningkatkan efektivitas perencanaan promosi kartu kredit.

## KESIMPULAN DAN SARAN

### Kesimpulan

Penggunaan *Simple Random Sampling* dengan rumus Slovin dan PCA berguna dalam menyederhanakan populasi dan mereduksi variabel yang digunakan pada awal proses pengolahan data dengan menggunakan data *mining*. Hasil penelitian menggunakan *cluster mining* menghasilkan tiga *cluster* segmentasi atau  $k=3$  terbaik berdasarkan 5 karakteristik nasabah, yaitu pekerjaan, pendidikan, penghasilan, usia dan asal wilayah nasabah. Hasil nilai *silhouette* koefisien yang dihasilkan dalam 10 kali percobaan memiliki rentang 0,39–0,40 atau paling mendekati nilai satu, artinya hasil dominan dari percobaan dapat dipilih sebagai tiga *cluster* segmentasi terbaik. Penggunaan ARM untuk mendapatkan pola transaksi pada *merchant* berdasarkan nilai *minimum support* 1% dan *minimum confidence* 40% menghasilkan dua pola jenis *merchant*. Pola pertama jika hotel (HHL) ditransaksikan maka restaurant (RRT) akan ikut ditransaksikan. Pola kedua jika bertransaksi pada *service station* (RSS) maka restaurant (RRT) akan ikut serta ditransaksikan. Dari hasil penelitian dapat disimpulkan *merchant* jenis restaurant memiliki frekuensi tersering digunakan oleh nasabah pemegang kartu kredit. Pembentukan target pasar pada penelitian

ini menghasilkan 3 kelompok konsumen pada 2 pola jenis *merchant* yang sering ditransaksikan oleh nasabah berdasarkan pekerjaan, penghasilan, pendidikan usia dan asal wilayah.

Penelitian juga menyimpulkan bahwa penggunaan data *mining* dengan *Clustering* dan ARM dapat bermanfaat dalam pembentukan segmentasi berdasarkan karakteristik nasabah dan identifikasi pola jenis *merchant* yang sering digunakan dalam menetapkan target pasar kartu kredit.

### Saran

Saran yang dapat diberikan pada penelitian ini yaitu perlunya analisis terhadap persebaran transaksi selain transaksi (*on us*) atau transaksi yang dilakukan tanpa menggunakan kartu kredit penerbit atau EDC (*Electronic Data Capture*) penerbit sehingga tidak ditemukan lost data saat dilakukan pengolahan dengan data *mining*. Perlunya identifikasi segmen pada faktor-faktor demografi lainnya dan psikografis nasabah seperti: jenis kelamin, tempat tinggal domisili, perilaku, dan ketertarikan konsumen pada jenis *merchant*. Identifikasi terhadap keluhan konsumen pada call center kartu kredit dapat dijadikan faktor menarik dalam menggali informasi *merchant* dan kepuasan konsumen kartu kredit. Kegiatan promosi seperti potongan harga (diskon) menghubungkan jenis *merchant* restaurant dengan jenis *merchant* lainnya, *cashback*, *loyalty* program, *blasting* sms/email, penggunaan billboard, banner dapat diefektifkan secara berkala pada target segmen nasabah yang potensial.

Evaluasi tindak lanjut hasil penelitian ini sebaiknya dilakukan secara berkala atau tahunan agar dapat mengetahui perkembangan upaya peningkatan transaksi kartu kredit terhadap nasabah dan kerja sama *business to business* dengan *merchant*. Jumlah sampling nasabah pada penelitian ini terbatas sebesar *error* 5% dikarenakan *tools* yang digunakan maksimal 10.000 data. Oleh karena itu, perlu penelitian lebih lanjut untuk mengetahui pembentukan pasar kartu kredit dengan mengidentifikasi segmentasi konsumen dan pola *merchant* kartu kredit dengan menggunakan teknik data *mining* lainnya serta menggunakan alat bantu *software* lainnya.

## DAFTAR PUSTAKA

- Alfina T, Santosa B, Barakbah A. 2012. Analisis perbandingan metode hierarChical clustering, K-Means dan gabungan keduanya dalam cluster data (studi kasus: problem kerja praktek Jurusan Teknik Industri ITS). *Jurnal Teknik ITS* 1:521–525.
- Annie CL, Kumar A. 2012. Market basket analysis for a supermarket based on frequent itemset mining. *International Journal of Computer Science* 9(3): 257–264.
- Bayer J. 2010. Customer segmentation in the telecommunications industry. *Journal of Database Marketing & Customer Strategy Management* 17(13): 247–256. <https://doi.org/10.1057/dbm.2010.21>.
- Brachman RJ, Khabasa T, Kloesgen W, Piatetsky Shapiro G, Simoudis E. 1996. Mining bussiness databases. *Communications of the ACM* 39(11):42–45. <https://doi.org/10.1145/240455.240468>.
- Cahyo A. 2002. Identifikasi segmentasi dan analisis perilaku penggunaan kartu kredit citibank, Bank “X” dan BCA [tesis]. Bogor: Management Business IPB.
- Davies CG, Blayney DP, Yen ST, Cooper J. 2009. An analysis of at home demand for ice cream in the united states. *Journal of Dairy Science* 92(12):6210–6216. <https://doi.org/10.3168/jds.2009-2536>.
- Garver MS, Spralls SA, Divine RL. 2012 Need based segmentation analysis of university career services: implication for increasing student participation. *Research In Higher Education Journal* 11(11): 611–628.
- Gürel G. 2008. Mining XML documents with association rule algorithms [thesis]. Izmir: Izmir Institute of Technology.
- Helbich M, Brunauer W, HAgebauer J, Leitner M 2013. Data-driven regionalization of housing markets. *Annals of the Association of American Geographers* 104(3): 460–484.
- Jaman M. 2012. Critical analysis of segmentation strategy for potential product launch – mapping the customers. *International Journal of Scientific & Technology Research* 1(11):62–65.
- Kotler, Amstrong. 2008. *Prinsip-prinsip Pemasaran. Ed ke 12*. Jakarta: Erlangga.
- Kotler P, Lane K. 2008 . *Manajemen Pemasaran. Ed ke 12*. Jakarta: PT. Macanan Jaya Cemerlang.
- Kotsiantis S, Kanellopoulos D. 2006. Association rules mining: a recent overview. *GESTS International Transactions on Computer Science and Engineering* 32(1):71–82.
- Krishnamurthi M. 2007. Improving credit card operations with data mining techniques. *Journal of International Technology and Information Management* 16(2):43–60.
- Liu H, Lin Y, Han J. 2009. Methods for mining frequent items in data streams: an overview. *Knowledge Information System* 26: 1–30. <https://doi.org/10.1007/s10115-009-0267-2>.
- Mabrur GA, Lubis R. 2012. Penerapan data mining untuk memprediksi kriteria nasabah kredit. *Jurnal Komputer dan Informatika (KOMPUTA)* 1(1):53–57.
- Robson K, Pitt L, Wallstrom A. 2013. Creative market segmentation: understanding the bugs in consumer behavior. *Journal of Public Affairs* 13(2):218–223. <https://doi.org/10.1002/pa.1477>.
- Sim THA. 2009. Discovering of association rules without a minimum support threshold – coherent rules discovery [dissertation]. Victoria: Monash University Australia.
- Srikant R. 1996. Fast algorithms for mining association rules and sequential patterns [dissertation]. Wisconsin: University of Wisconsin.
- Tan PN, Steinbach M, Kumar V. 2006. *Introduction to Data Mining. Ed ke-1*. Boston: Pearson.
- Vaishali. 2014. Fraud detection in credit card by clustering approach. *International Journal of Computer Applications* 98(3):29–32. <https://doi.org/10.5120/17164-7225>.
- Zheng Y. 2009. Analysis of credit card data based on data mining technique [thesis]. Ottawa: Quebec University.