ANALISIS PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK MENERAPKAN SEGMENTASI PELANGGAN

TESIS

RACHMAWATI DARMA ASTUTI (14002140)

PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2019

ANALISIS PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK MENERAPKAN SEGMENTASI PELANGGAN

TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar

Magister Ilmu Komputer

RACHMAWATI DARMA ASTUTI (14002140)

PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2019

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini :

Nama	:	Rachmawati Darma Astuti
NIM	:	14002140
Program Studi	:	Ilmu Komputer (S2)
Jenjang	:	Strata Dua (S2)
Konsentrasi	:	*Software Engineering*

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: "ANALISIS PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK MENERAPKAN SEGMENTASI PELANGGAN" adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan.

Jakarta, 08 Januari 2020
Yang menyatakan,



Rachmawati Darma Astuti

Dipindai dengan CamScanner

PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh:

Nama	: Rachmawati Durma Astuti
NIM	: 14002140
Program Studi	: Ilmu Komputer
Jenjang	: Program Strata Dua (S2)
Konsentrasi	: *Data Mining*
Judul Tesis	: Analisis Perbandingan Algoritma K-means Dan K-medoids Untuk Menerapkan Segmentasi Pelanggan

Telah dipertahankan pada periode 2019-2 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Magister Ilmu Komputer (M.Kom) pada Program Strata Dua (S2) Program Studi Ilmu Komputer di Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 22 Januari 2020

PEMBIMBING TESIS

Dosen Pembimbing : Dr. Dwiza Riana, S.Si, MM, M.Kom .....

DEWAN PENGUJI

Penguji I : Dr. Hilman Ferdinandus Pardede, ST, MEICT .....

Penguji II : Dr. Didi Rosiyadi, M.Kom .....

Penguji III / Dosen Pembimbing : Dr. Dwiza Riana, S.Si, MM, M.Kom .....

KATA PENGANTAR

Alhamdulillah, Segela puji syukur penulis panjatkan kepada Allah SWT atas rahmat dan seluruh nikmat yang telah diberikan sehingga penulis dapat menyelesaikan pembuatan tesis yang berjudul "ANALISIS PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK MENERAPKAN SEGMENTASI PELANGGAN". Adapun tujuan pembuatan tesis ini adalah sebagai persyaratan untuk memperoleh gelar (S2) Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.

Tesis ini diambil berdasarkan data transaksi yang ada pada distributor pulsa elektronik PT. Pulsantara. Penulis juga mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, internet, dll yang terkait dengan pembahasan pada tesis ini

. Selama proses penyusunan tesis ini penulis banyak mendapatkan bantuan dari berbagai pihak. Penulis menyampaikan ucapan terima kasih yang sebesar besarnya kepada:

1. Ibu Dr. Dwiza Riana, S.Si, M.M, M.Kom selaku ketua STMIK Nusa Mandiri dan selaku pembimbing tesis yang telah meluangkan
   waktu, tenaga dan fikiran guna memberikan bimbingan bagi penulis dalam menyelesaikan tesis ini.
2. Orang tua dan keluarga yang telah memberikan dukungan moral, materil dan doa.
3. Seluruh staf dan karyawan Universitas Bina Sarana Informatika Dan STMIK Nusa Mandiri Jakarta.

4. Teman-teman seperjuangan yang telah memberikan dukungan moril kepada penulis serta semua pihak yang telah membantu penulis dalam proses penyelesaian tesis ini.

Dalam penulisan tesis ini penulis menyadari bahwasannya tesis ini masih jauh dari sempurna dan masih banyak kekurangan. Untuk itu penulis mengharapkan kritik dan saran yang bersifat membangun dari semua pihak untuk dapat lebih menyempurnakan tesis ini di masa yang akan datang.

Semoga tesis ini dapat memberikan manfaat bagi penulis dan pembaca serta sedikit memiliki peran bagi kemajuan ilmu pengetahuan.

Jakarta, 08 Januari 2020

Rachmawati Darma Astuti

Penulis

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya :

Nama	: Rachmawati Darma Astuti
NIM	: 14002140
Program Studi	: Ilmu Komputer (S2)
Jenjang	: Strata Dua (S2)
Konsentrasi	: *Software Engineering*
Jenis Karya	: Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) **Hak Bebas Royalti Non-Eksklusif** ***(Non-exclusive Royalti-Free Right)*** atas karya ilmiah kami yang berjudul : "ANALISIS PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK MENERAPKAN SEGMENTASI PELANGGAN" beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau *bentuk*-kan, mengelolaannya dalam pangkalan data *(database)*, mendistribusikannya dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 08 Januari 2020
Yang menyatakan,



Rachmawati Darma Astuti

Dipindai dengan CamScanner

ABSTRAK

Nama	:	Rachmawati Darma Astuti
NIM	:	14002140
Program Studi	:	Magister Ilmu Komputer
Jenjang	:	Strata Dua (S2)
Konsentrasi	:	*Software Engineering*
Judul Tesis	:	Analisis Perbandingan Algoritma K-Means Dan K-Medoids Untuk Menerapkan Segmentasi Pelanggan

Dalam persaingan bisnis di era modern saat ini, pelanggan merupakan hal yang perlu menjadi fokus utama perusahaan untuk mempertahankan eksitensinya. Perusahaan harus merencanakan dan menggunakan strategi yang jelas dalam memperlakukan pelanggan . Untuk mendapatkan pelanggan baru membutuhkan biaya lebih besar dibandingkan dengan mempertahankan pelanggan lama. Pada penelitan ini dilakukan segmentasi dengan menggunakan metode teknik clustering agar proses dan hasil pengelompokannya lebih baik. Algoritma yang digunakan dalam pengelompokan pelanggan server Pulsantara menggunakan K-Means dan K-medoids yang nantinya akan dibandingkan algoritma yang terbaik antara K-Means dan K-medoids. Pada hasil akhir penelitian dapat disimpulkan bahawa lagoritma terbaik untuk segmentasi pelanggan Pulsantara adalah algoritma K-medoids dengan jumlah cluster berjumlah 6 memiliki DBI sebesar 0.660%. *Cluster* yang terbentuk menggunakan algoritma *K*-Medoids memiliki 6 *cluster*/segmen pelanggan dari keseluruhan jumlah transaksi 85010 transaksi dengan jumlah agen sebanyak 133. Jumlah anggota tiap cluster antara lain: 22 anggota berada pada *cluster 0* disebut *Typical Customer*, 14 anggota berada pada *cluster 1 disebut Occational customer*, 13 anggota berada pada *cluster 2* disebut *Golden Customer* , 4 anggota berada pada *cluster 3* disebut *Superstar*, pada *cluster 4* memiliki jumlah anggota 52 penggan disebut *Dormant Customer* sedangkan untuk *cluster 5* memiliki 28 jumlah anggota yang disebut dengan *Everydar Shopper.*

Kata Kunci : RFM, Segmentasi Pelanggan, *Clustering*, K-Medoids, K-Means

ABSTRACT

Name	:	Rachmawati Darma Astuti
NIM	:	14002140
Study of Program	:	Magister Ilmu Komputer
Levels	:	Strata Dua (S2)
Concentration	:	*Software Engineering*
Title	:	*Comparative Analysis of K-Means and K-Medoids Algorithms for Implementing Customer Segmentation*

In business competition in the modern era, customers are the main focus of the company to maintain its existence. Companies must plan and use clear strategies in treating customers. Getting new customers costs more than keeping old customers. In this research segmentation is done by using clustering techniques so that the process and the results of grouping are better. The algorithm used in grouping Pulsantara server customers uses K-Means and K-medoids which will be compared to the best algorithm between K-Means and K-medoids. In the final results of the study it can be concluded that the best lagorithm for Pulsantara customer segmentation is the K-medoids algorithm with a cluster number of 6 having a DBI of 0.660%. Clusters formed using the K-Medoids algorithm have 6 clusters / customer segments of the total number of transactions of 85010 transactions with the number of agents as many as 133. The number of members of each cluster include: 22 members are in cluster 0 called Typical Customers, 14 members are in cluster 1 called Occational customer, 13 members are in cluster 2 called Golden Customer, 4 members are in cluster 3 called Superstar, in cluster 4 there are 52 members called Dormant Customers while for cluster 5 there are 28 members called Everydar Shopper.

Keywords: RFM, Customer Segmentation, Clustering, K-Medoids, K-Means

DAFTAR ISI

Halaman

LEMBAR JUDUL TESIS ii

LEMBAR PERNYATAAN ORISINALITAS iii

LEMBAR HALAMAN PENGESAHAN iv

KATA PENGANTAR v

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS vii

ABSTRAK viii

ABSTRACT ix

DAFTAR ISI x

DAFTAR TABEL xii

DAFTAR GAMBAR xiii

DAFTAR LAMPIRAN xiv

DAFTAR RIWAYAT HIDUP xv

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Identifikasi Masalah 3

1.3 Tujuan 3

1.4 Ruang Lingkup Penelitian 4

1.5 Sistematika Penulis 4

BAB II LANDASAN TEORI 6

2.1 Tinjauan Pustaka 6

2.1.1 *Customer Relationship Management* (CRM) 6

2.1.2 Segmentasi Pelanggan 8

2.1.3 RFM (*Recency, Frequency and Monetary*) 11
2.1.4 *Data Mining* 12
2.1.5 *Data Mining* Dalam Kerangka Kerja CRM 16
2.1.6 *Clustering* 17
2.1.7 K-*Means* 18
2.1.8 K-*Medoids* 19
2.2 Tinjauan Studi 20
2.3 Tinjauan Objek Penelitian 23
BAB III METODOLOGI PENELITIAN 25
3.1 Metodologi Penelitian 25
3.1.1 Jenis Penelitian 26
3.1.2 Metode Pengumpulan Data 27
3.1.3 Eksperimen 27
3.1.4 Validasi 33
BAB IV HASIL PENELITIAN 34
4.1 Hasil Penelitian 34
4.1.1 Persiapan *Data Mining* 33
4.2 Pembahasan 40
4.2.1 Pengolahan Data Menggunakan *Rapid Minner* 40
4.2.2 Pengolahan Data Menggunakan Algorithma K-*Means* 43
4.2.3 Pengolahan Data Menggunakan Algorithma K-*Medoids* 43
4.3 Hasil Pengolahan 44
4.3.1 Hasil *Cluster Optimal* 45
4.3.2 Penentuan Karakteristik Pelanngan 48
BAB V KESIMPULAN 55
5.1 Kesimpulan 55
5.2 Saran 56
DAFTAR PUSTAKA
LAMPIRAN

DAFTAR GAMBAR

Gambar Halaman

2.1 Proses *Knowledge Discovery From Data* 13
2.2 Proses CRISP-DM 14
2.3 Kerangka Kerja Teknik *Data Mining* Dalam CRM 17
3.1 Langkah Penelitian 25
3.2 *FlowChart* Cara Kerja *Algorithma* K-*Means* 30
3.3 *FlowChart* Cara Kerja *Algorithma* K-*Medoids* 32
4.1 *Data Base* Transaksi Pulsantara 35
4.2 *Import* Data Set Pada Lembar Kerja *Rapid Minner* 40
4.3 Data Set Berupa Excel Yang Akan Di *Import* 41
4.4 Tampilan *Sheet* Data 41
4.5 Menentukan Anotasi Dari Setiap Tupel Dan Atribut 42
4.6 Data Yang Sudah Dimasukan Ke Read Excel 42
4.7 Pemodelan K-*Means* 43
4.8 Pemodelan K-*Medoids* 44
4.9 Grafik Hasil Pengklasteran Menggunakan K-*Means* 45
4.10 Grafik Hasil Pengklasteran Menggunakan K-*Medoids* 46
4.11 *Centroid Tabel* K-*Medoids* 47

DAFTAR TABEL

Tabel l Halaman

2.1 Karakter Pelanggan Berdasarkan Nilai RFM 12
2.2 Tinjauan Penelitian Sebelumnya Dari *International Journal* 20
2.3 Tinjauan Penelitian Sebelumnya Dari Jurnal Nasional 21
3.1 *Variabel* Proses *Clustering* 29
3.2 Menunjukan *Score Recency-Frequent-Monetary* 29
4.1 Data Set Transaksi Pulsantara 35
4.2 Kategori *Recency* 37
4.3 Kategori *Frequency* 38
4.4 Kategori *Monetary* 38
4.5 Kategori Pembobotan RFM 38
4.6 Nilai Bobot Pada RFM 39
4.7 Perbandingan Performa *Cluster* 46
4.8 Anggota *Cluster* Algorithma K-*Medoids* dengan 6 *Cluster* 48
4.9 Hasil *Cluster* 1 49
4.10 Hasil *Cluster* 2 50
4.11 Hasil *Cluster* 3 51
4.12 Hasil *Cluster* 4 52
4.13 Hasil *Cluster* 5 53

DAFTAR GAMBAR

Gambar Halaman

2.1 Proses *Knowledge Discovery From Data* 13
2.2 Proses CRISP-DM 14
2.3 Kerangka Kerja Teknik *Data Mining* Dalam CRM 17
3.1 Langkah Penelitian 25
3.2 *FlowChart* Cara Kerja *Algorithma* K-*Means* 30
3.3 *FlowChart* Cara Kerja *Algorithma* K-*Medoids* 32
4.1 *Data Base* Transaksi Pulsantara 35
4.2 *Import* Data Set Pada Lembar Kerja *Rapid Minner* 40
4.3 Data Set Berupa Excel Yang Akan Di *Import* 41
4.4 Tampilan *Sheet* Data 41
4.5 Menentukan Anotasi Dari Setiap Tupel Dan Atribut 42
4.6 Data Yang Sudah Dimasukan Ke Read Excel 42
4.7 Pemodelan K-*Means* 43
4.8 Pemodelan K-*Medoids* 44
4.9 Grafik Hasil Pengklasteran Menggunakan K-*Means* 45
4.10 Grafik Hasil Pengklasteran Menggunakan K-*Medoids* 46
4.11 *Centroid Tabel* K-*Medoids* 47

DAFTAR TABEL

Tabel l Halaman

2.1 Karakter Pelanggan Berdasarkan Nilai RFM 12
2.2 Tinjauan Penelitian Sebelumnya Dari *International Journal* 20
2.3 Tinjauan Penelitian Sebelumnya Dari Jurnal Nasional 21
3.1 *Variabel* Proses *Clustering* 29
3.2 Menunjukan *Score Recency-Frequent-Monetary* 29
4.1 Data Set Transaksi Pulsantara 35
4.2 Kategori *Recency* 37
4.3 Kategori *Frequency* 38
4.4 Kategori *Monetary* 38
4.5 Kategori Pembobotan RFM 38
4.6 Nilai Bobot Pada RFM 39
4.7 Perbandingan Performa *Cluster* 46
4.8 Anggota *Cluster* Algorithma K-*Medoids* dengan 6 *Cluster* 48
4.9 Hasil *Cluster* 1 49
4.10 Hasil *Cluster* 2 50
4.11 Hasil *Cluster* 3 51
4.12 Hasil *Cluster* 4 52
4.13 Hasil *Cluster* 5 53

BAB I
PENDAHULUAN

1.1. Latar Belakang

Salah satu bisnis yang memiliki tingkat resiko rendah adalah bisnis pulsa elektronik dikarenakan pulsa seluler atau elektronik merupakan kebutuhan utama yang harus ada, tidak mudah rusak, tidak mempunyai waktu kadaluarsa, tidak membutuhkan biaya pengiriman. Dalam persaingan bisnis di era modern saat ini, pelanggan merupakan hal yang perlu menjadi fokus utama perusahaan untuk mempertahankan eksitensinya. Perusahaan harus merencanakan dan menggunakan strategi yang jelas dalam memperlakukan pelanggan (Maryani & Riana, 2017). Contoh tantangan pada industri telekomunikasi seluler adalah tingginya jumlah pelanggan yang berhenti menggunakan layanan perusahaan dan pindah ke perusahaan kompetitor (I Gede Wiyana Ananta Noor, Ariyanti, & Alamsyah, 2019). Kasus pelanggan yang tidak loyal atau berpindah ke kompetitor lain (churn) merupakan permasalahan utama yang sering dihadapi perusahaan telekomunikasi karena akan berpengaruh terhadap pendapatan yang diterima oleh perusahaan tersebut. Pelanggan yang setia, akan menghasilkan pendapatan tinggi untuk perusahaan, jarang terpengaruh oleh perusahaan pesaing(Ullah et al., 2019) .

Pada bisnis pulsa, data dapat diperoleh berdasarkan data historis, sehingga data akan bertambah secara terus menerus, misalnya data transaksi dari masing-masing agen. Proses transaksi para agen dalam sebuah server pulsa menghasilkan data yang berlimpah berupa profil dari transaksi yang dilakukan agen tersebut. Hal ini akan terjadi secara berulang pada pelaku bisnis pulsa. Penumpukan data transaksi agen secara menerus akan memperlambat pencarian informasi terhadap data tersebut (Hand, 2007).

Data mining merupakan bagian dari *knowledge discovery data* yang merupakan proses ekstraksi informasi yang berguna, tidak diketahui sebelumnya dan tersembunyi dari data (Ramamohan, Vasantharao, Chakravarti, & Ratnam, 2012). Data mining juga banyak digunakan untuk segmentasi pelanggan (Çaliş &

Boyaci, 2015) (Zhang, Zhang, & Qiu, 2018) (Yuliari, Putra, & Rusjayanti, 2015) (Hamdi & Zamiri, 2016) (Wong & Wei, 2018) (Shihab, Afroge, & Mishu, 2019).

Berdasarkan banyaknya data transaksi agen yang tersedia, maka informasi yang belum diketahui atau tersembunyi dapat diketahui dengan cara melakukan pengolahan terhadap data tersebut sehingga berguna bagi pihak pelaku bisnis pulsa (Ramamohan et al., 2012), misalnya informasi mengenai pengelompokan data agen berpotensi memberikan keuntungan yang paling banyak pada perusahaan, yang nantinya dapat membantu perusahan untuk mengambil keputusan dalam pemasaran produk.

Model yang digunakan oleh peneliti adalah RFM (Recency, Frequency, Monetary) yang umum digunakan untuk melakukan pengelompokan waktu kunjungan terakhir, frekuensi kunjungan dan pendapatan yang diperoleh perusahaan (Wongchinsri & Kuratach, 2016). Alasan mengapa menggunakan model RFM adalah mudah digunakan dan cepat di implemntasi pada perusahaan, selain itu RFM mudah dimengerti oleh manajer dan pengambil keputusan dalam pemasaran (Peiman Alipour Sarvari, Ustundag, & Takci, 2014). Pada akhir penelitian diusulkan rekomendasi CRM untuk membuat hubungan bisnis dengan pelanggan menjadi lebih efisien serta dapat memaksimalkan kepuasan pelanggan dan dengan demikian meningkatkan loyalitas pelanggan dan retensi (Roselin, 2014).

Pada penelitan ini dilakukan segmentasi dengan menggunakan metode teknik *clustering* agar proses dan hasil pengelompokannya lebih baik. algoritma yang digunakan dalam pengelompokan pelanggan server Pulsantara menggunakan *K-Means dan K-medoids* yang nantinya akan dibandingkan algoritma yang terbaik antara *K-Means dan K-medoids*. Metode *K-Means* merupakan metode clustering dalam Data mining yang sering digunakan oleh peneliti untuk mengklaster data (Kashwan & Velu, 2013) , (Febrianti, Hafiyusholeh, & Asyhar, 2016) , (Yi, Yang, Qiao, & Xu, 2010) dan yang lainnya. Metode K-Medoids merupakan metode clustering yang cukup efisien untuk data kecil (Zayuka, Nasution, & Purwanto, 2017). Hasil penelitian ini dapat digunakan sebagai sistem pendukung keputusan pada pelaku bisnis pulsa untuk memetakan pelangggan dan mengetahui pelanggan potensial

Berdasarkan latar belakang di atas maka peneliti tertarik untuk melakukan penelitian dengan judul **"ANALISIS PERBANDINGAN ALGORITMA *K-MEANS DAN K-MEDOIDS* UNTUK MENERAPAKAN SEGMENTASI PELANGGAN MENGGUNAKAN METODE RFM".**

1.2. Identifikasi Masalah

Berdasarkan latar belakang dan identifikasi masalah adalah sebagai berikut:

a. Data transaksi yang begitu besar dan belum diolah pada server Pulsantara, sehingga perlu adanya pengolahan data menggunakan datamining.
b. Perlunya penerapan RFM pada server Pulsantara untuk pemetan pelanggan yang dimiliki oleh perusahaan, sehingga dapat diketahui segmen dari para pelanggan tersebut berdasarkan karakteristik yang dimilikinya
c. Perlunya analisis data menggunakan algoritma clustering K-Means dan K-Medoids untuk mengelompokkan data dengan karakteristik yang sama ke suatu klaster bagi perusahaan agar tepat dalam melakukan strategi pemasaran untuk menjaga loyalitas pelanggan ?.

1.3. Tujuan

Tujuan dari penelitian ini adalah :

a. Membuat Customer Profile untuk mengelompokkan/memetakan pelanggan berdasarkan analisis RFM (Recency, Frequency dan Monetary)
b. Mengklasifikasikan dan Mengetahui Karakteristik pelanggan serta melakukan Evaluasi Loyalitas Pelanggan.
c. Mencari tahu agoritma mana yang paling cocok untuk di terapkan pada perusahan Pulsantara jika menggunakan 6 kluter.
d. Mengetahui bagaimana menerapkan Customer Relationship Management dalam melakukan strategi pemasaran yang tepat untuk menjaga loyalitas pelanggan.

1.4. Ruang Lingkup Penelitian

Pada bagian kali ini penulis membatasi ruang lingkup penelitian hanya kepada:

a. Luas lingkup data yang kelola adalah data dari perusahaan Sembilan reload dari tanggal 1 januari 2017 sampai 31 Desember 2018.
b. Hanya mengambil data transaksi yang Sukses.
c. Data atribut yang digunakan adalah Nama, Tanggal, Harga.
d. Membandingkan dua algoritma yaitu *K-Means dan K-medoids* jika menggunakan 4 kluter.

1.5. Sistematika Penulisan

Sistematika penulisan tesis ini terdiri dari 5 (lima) bab, dimana tiap bab terdiri dari beberapa sub bab sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini membahas tentang Latar Belakang Penulisan, Identifikasi Masalah, Batasan Masalah, Rumusan Masalah, Tujuan Manfaat dan Sistematika Penulisan.

BAB 2 LANDASAN/KERANGKA PEMIKIRAN

Pada Bab ini dibahas teori yang melandasi penelitian, dalam bab ini juga diuraikan Tinjauan Pustaka, Tinjauan Studi, Tinjauan Organisasi/Obyek Penelitian, Kerangka Pemikiran.

BAB 3 METODOLOGI PENELITIAN

Membahas mengenai Metodologi/langkah penelitian, pengumpulan data, metode analisis data, kerangka pendekatan berikut jadwal penelitian yang digunakan untuk melakukan segmentasi konsumen

BAB 4 HASIL PENELITIAN DAN PEMBAHASAN

Menampilkan hasil dari pengelompokkan data dan analisa hasil pengolahan data. Pada bab ini juga ditampilkan hasil *K-Means dan K-medoids* berdasarkan analisis *variable Recency, Frequency, Monetar*y dan Pembobotan Kriteria.

BAB 5 PENUTUP

Membahas kesimpulan dari penelitian yang dilakukan dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Pada penelitian ini, tinjauan pustaka yang digunakan merupakan teori – teori yang menjadi landasan dalam penelitian ini. Selain itu kajian pustaka juga didapat melalui referensi dari jurnal – jurnal penelitian baik jurnal nasional maupun internasional.

2.1.1 *Customer Relationship Management* (CRM)

Menurut (Supriatna & Budianto, 2019) CRM adalah strategi inti dalam bisnis yang mengintegrasikan proses-proses dan fungsi-fungsi internal dengan semua jaringan eksternal untuk menciptakan serta mewujudkan nilai bagi para konsumen sasaran secara profitabel. Sedangkan menurut (Yuliari, Putra, & Rusjayanti, 2015) berpendapat bahwa "CRM pada intinya merupakan kolaborasi dengan setiap konsumen yang mampu menciptakan keadaan yang tidak merugikan salah satu pihak (*win-win situation*). anda menambah nilai pada kehidupan sehari-hari setiap konsumen, dan sebagai imbalannya, mereka memberikan kesetiaan kepada anda. Sesungguhnya, proses ini merupakan hal yang berhubungan dengan setiap konsumen secara individual". Menurut (Utami, 2010) pengertian CRM adalah "Suatu proses interaktif yang mengubah data-data pelanggan kedalam kesetiaan pelanggan melalui beberapa kegiatan, yaitu mengumpulkan data pelanggan, menganalisis data pelanggan tersebut dan mengidentifikasi target pelanggan, mengembangkan program CRM, dan menerapkan program CRM". Dari beberapa pendapat di atas maka dapat disimpulkan bahwa *Customer Relationship Management* (CRM) merupakan strategi bisnis yang mengintegrasikan proses dan fungsi internal dengan eksternal untuk menciptakan nilai dan memanjakan pelanggan serta menciptakan keadaan *win-win situation* melalui serangkaian kegiatan mengelola informasi yang rinci tentang masing-masing pelanggan dengan tujuan untuk menciptakan kesetiaan pelanggan agar tidak berpaling kepada pesaing. Pada intinya perusahaan bermaksud membangun ikatan yang lebih kuat dengan para pelanggan yang bertujuan untuk memaksimalkan loyalitas pelanggan.

Konsep *Customer Relationship Management* (CRM) dapat dipahami dalam tiga tataran, yaitu strategis, operasional, dan analitis.

1. CRM strategis terfokus pada upaya untuk mengembangkan kultur usaha yang berorientasi pada pelanggan atau *customer-centric*. Kultur ini ditujukan untuk merebut hati konsumen dan menjaga loyalitas mereka dengan menciptakan serta memberikan nilai bagi pelanggan yang mengungguli para pesaing. Kultur ini tercermin dari perilaku pucuk pimpinan perusahaan, desain sistem formal di dalam perusahaan, dan berbagai mitos dan cerita yang beredar di dalam perusahaan.
2. CRM operasional lebih terfokus pada otomatisasi cara-cara perusahaan dalam berhubungan dengan pelanggan. Berbagai aplikasi perangkat lunak CRM memungkinkan fungsifungsi pemasaran, penjualan, dan penjualan dapat berjalan secara otomatis.
3. CRM analitis digunakan untuk mengekploitasi data konsumen demi meningkatkan nilai mereka (dan nilai perusahaan). sistem ini dikembangkan berdasarkan informasi mengenai konsumen. Data pelanggan dapat diperoleh dari pusat-pusat informasi atau bank data yang dimiliki setiap perusahaan yang relevan, yakni data penjualan (riwayat pembelian barang atau jasa oleh pelanggan), data finansial (riwayat pembayaran atau skor kredit), data pemasaran (respon konsumen terhadap kampanye iklan, data skala loyalitas produk), dan data layanan.

(Utami, 2010) menjelasan proses dari Manajemen Hubungan Pelanggan (*Customer Relationship Management)* adala mengumpulkan data pelanggan, menganalisis data pelanggan dan identitas target pelanggan, mengembangkan Program CRM, dan mengimplementasikan Program CRM. *Customer Relationship Management* (CRM) memungkinkan perusahaan untuk memberikan pelayanan kepada pelanggan secara langsung dan secara maksimal yaitu dengan cara mengembangkan hubungan dengan setiap pelanggan yang berharga melalui penggunaan informasi atau basis data (*database*) pelanggan yang dimiliki perusahaan. Dalam hal ini perusahaan menggunakan strategi *Customer Relationship Management* (CRM) dengan tujuan agar perusahaan tetap fokus terhadap pelanggan berdasarkan informasi atau basis data (*database*) pelanggan yang dimiliki.

2.1.2 Segmentasi Pelanggan

Segmentasi pelanggan adalah proses membagi pelanggan yang berbeda ke kelompok *homogen* untuk membangun strategi pemasaran berdasarkan karakteristik pelanggan. Banyak tipe segmentasi yang berbeda berdasarkan kriteria fisik atau atribut yang digunakan untuk melakukan segmentasi. Sifat dari

segmentasi pelanggan dikelompokan berdasarkan sifat dan karakteristik yang digunakan. Sifat dari segmentasi dapat dibuat dengan aturan bisnis. Pendekatan yang melekat ini tidak menguntungkan akan lebih efisien jika ditangani oleh beberapa segmentasi dasar dan objektifitas berdasarkan persepsi personal dari ahli bisnis. Algoritma *clustering* dapat menganalisa sifat data, mengidentifikasi kelompok pelanggan secara alami dan menawarkan solusi yang diterima dari hasil pengamatan pola data. Pelanyanan Model Data Mining dibangun dengan baik dapat menemukan kelompok dengan profil dan karakteristik yang jelas dan memperkaya skema segmentasi dengan pengertian dan nilai bisnis (Tsiptsis & Chorianopoulos, 2009a)

Untuk dapat menjadi pelanggan yang loyal, seorang pelanggan harus memulai beberapa tahapan. Proses ini berlangsung lama dengan penekanan dan perhatian yang berbeda untuk masing-masing tahap karena setiap tahap mempunyai kebutuhan yang berbeda. Dengan memperhatikan masing-masing tahapan dan memenuhi kebutuhan dalam setiap tahap tersebut, perusahaan memiliki peluang yang lebih besar untuk membentuk calon pembeli menjadi pelanggan loyal dan klien perusahaan. (Hill & Alexander, 2006) menjelaskan bahwa tingkatan loyal terbagi atas 6 tingkat, yaitu :

1. *Suspect*

   Bagian ini termasuk semua pembeli produk atau jasa dalam pemasaran, jadi *suspects* adalah menyadari akan produk atau jasa perusahaan atau tidak mempunyai kecenderungan terhadap pembelian.

2. *Prospects*

   Prospects adalah pelanggan *potensial* yang mempunyai daya tarik terhadap perusahaan tetapi belum mengambil langkah untuk melakukan bisnis dengan perusahaan.

3. *Customers*

   Suatu tipe pembelian produk (walaupun dalam kategori ini termasuk beberapa pembelian ulang) yang tidak memiliki loyalitas pada perusahaan.

4. *Clients*

   Pembelian ulang yang menunjukkan *loyalitas* pada perusahaan tetapi lebih memiliki dorongan pasif daripada aktif terhadap perusahaan.

5. *Advocates*

   Client yang memberikan dorongan yang positif pada perusahaan dengan merekomendasikannya kepada orang lain.

6. *Partners*

   Partners adalah hubungan yang sangat erat antara konsumen dengan supplier yang keduanya saling memperlihatkan keuntungan.

(Hill & Alexander, 2006) membagi tahapan loyalitas pelanggan menjadi 6 tahap dari tingkat suspects hingga tahap partner, untuk lebih jelasnya lagi di bawah ini digambarkan mengenai piramida tentang loyalitas pelanggan.

The Profit Generator System

1. *Suspects*

Adalah semua orang yang mungkin akan membeli produk atau jasa perusahaan. Kita menyebutnya sepagai suspects karena yakin bahwa mereka akan membeli tetapi belum tahu apapun mengenai perusahaan dan barang/ jasa yang ditawarkan.

2. *Prospects*

Adalah orang-orang yang memiliki kebutuhan akan produk atau jasa tertentu dan mempunyai kemampuan untuk membelinya. Para prospects ini meskipun mereka belum melakukan pembelian, mereka telah menyatakan keberadaan perusahaan, barang dan jasa yang ditawarkan karena seseorang telah merekomendasikan barang/ jasa tersebut padanya.

3. Disqualified Prospect

Yaitu prospect yang telah mengetahui keberadaan barang/ jasa terentu, tetapi tidak mempunyai kebutuhan akan barang/ jasa tersebut atau tidak mempunyai kemampuan untuk membeli barang/ jasa tersebut.

4. First Time Customer

Yaitu pelanggan yang membeli untuk yang pertama kalinya, mereka masih menjadi pelanggan yang baru.

5. Repeat Customers

Yaitu pelanggan yang telah melakukan pembelian suatu produk sebanyak 2 kali atau lebih. Mereka adalah yang melakukan pembelian atas produk yang sama sebanyak 2 kali atau membeli dua macam produk yang berbeda dalam 2 kesempatan yang berbeda pula.

6. *Clients*

   Membeli semua barang/jasa yang ditawarkan yang mereka butuhkan, mereka membeli secara teratur, hubungan dengan jenis pelanggan ini sudah kuat dan berlangsung lama yang membuat mereka tidak terpengaruh oleh tarikan pesaing produk lain.

7. *Advocates*

   Seperti layaknya klien, advocates membeli seluruh barang/jasa yang ditawarkan yang ia butuhkan, serta melakukan pembelian secara teratur sebagai tambahan mereka mendorong teman-teman mereka yang lain agar membeli barang/ jasa tersebut. Ia membicarakan tentang barang/jasa tersebut, melakukan pemasaran untuk perusahaan tersebut dan membawa pelanggan untuk perusahaan tersebut.

2.1.3 RFM (*Recency, Frequency and Monetary*)

Model RFM (*Recency, Frequency and Monetary*) telah banyak diterapkan dalam beberapa bidang, terutama dalam dunia pemasaran. Dengan mengadopsi model RFM, seorang pengambil keputusan dapat secara efektif mengidentifikasi pelanggan yang berharga dan akan digunakan sebagai pengembangan strategi pemasaran yang efektif (Shihab, Afroge, & Mishu, 2019). Model RFM sering digunakan untuk segmentasi pasar. RFM mempertahankan informasi tentang waktu pembelian (*recency*) paling akhir, berapa kali pelanggan melakukan pembelian (*frequency*), dan rata-rata uang yang dihabiskan (*monetary*) (Lee, 2014) . Keuntungan model RFM terletak pada relevansinya selama beroperasi pada beberapa variabel yang dapat diamati dan bersifat objektif. Variabel ini digolongkan menurut 3 kriteria, yaitu *recency, frequency*, dan *monetary*.

Berdasarkan nilai RFM yang digunakan dalam analisis clustering pelanggan (Tsiptsis & Chorianopoulos, 2009b), karakteristik pelanggan dibagi menjadi 6 jenis sebagaimana penjabarannya pada Tabel II.1 dibawah ini :

Tabel II.1 Karakter pelanggan berdasarkan Nilai RFM

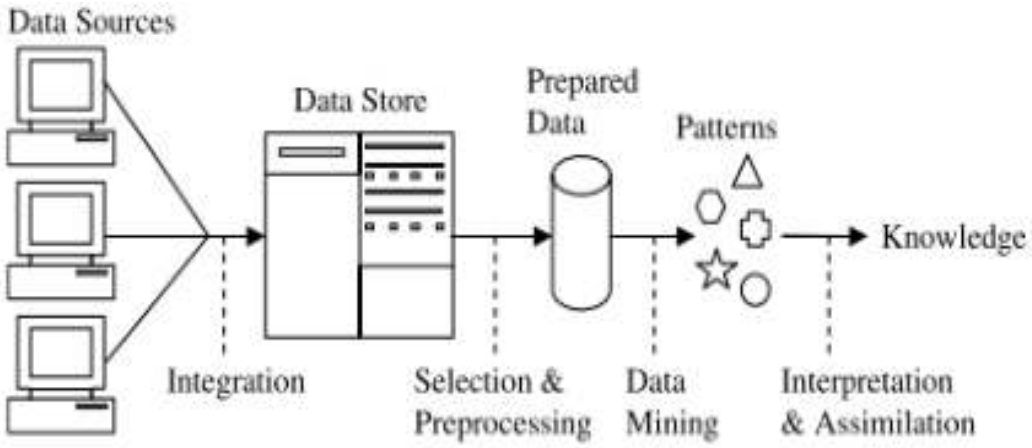
Kelas Pelanggan	Karakteristik
Superstar	Pelanggan dengan tingkat loyalit yang tinggi Memiliki nilai monetary yang tinggi Memiliki frekuensi yang tinggi Memilki nilai transaksi yang tinggi
Golden Customer	Memiliki nilai monetary tertinggi kedua Memilki frekuensi yang tinggi Memiliki nilai transaksi rata - rata yang tinggi
Typical Customer	Memiliki nilai monetary dan frekuemsi rata-rata Memiliki nilai transaksi rata-rata Memiliki nilai frekuensi terendah setelah Dorman *Customer*
Occational customer	Memiliki *recency* rendah (memiliki waktu yang lama dengan rentang waktu trakhir kunjungan) Melakukan pembelanjaan (*monetary*) dalam jumlah besar.
Everydar Shopper	Memiliki peningkatan dalam bertransaksi Melakukan pembelanjaan dalam jumlah kecil Memiliki nilai transaksi yang rendah Memiliki frekuensi dan *monetary* terendah
Dormant Customer	Memiliki waktu yang lama ketika masa terakhir kunjungan (memiliki *recency* terendah)

2.1.4 *Data Mining*

Data mining dapat didefinisikan sebagai proses menemukan dan menggambarkan pola struktural dalam data sebagai alat untuk membantu menjelaskan data dan membuat prediksi dari data tersebut (Witten, Frank, Hall, & Pall, 2017) . *Data mining* merupakan bagian dari *Knowldge Discovery Data* (KDD) yang merupakan proses ekstraksi informasi yang berguna, tidak diketahui sebelumnya, dan tersembunyi dari data (Bramer, 2016) dan juga mengembangkan model yang digunakan untuk memahami fenomena dari analisis data dan prediksi

(Maimon & Rokach, 2010). *Data mining* merujuk pada ekstraksi pengetahuan dari jumlah data yang besar (Han, Kamber, & Pei, 2012) yang tersimpan dalam komputer.

Dibawah ini digambarkan proses ekstraksi data menjadi informasi:



Sumber: (Bramer, 2016)

Gambar II. 1 Proses *Knowledge Discovery From Data*

Dari gambar II.1 menunjukan bahwa data bisa berasal dari mana saja, data-data tersebut kemudian diintegrasikan kedalam sebuah *data store*, dari *data store* data kemudian diseleksi dan diproses sehingga menghasilkan pola-pola dan pengetahuan yang berguna.

Secara garis besar proses KDD dapat jelaskan sebagai berikut (Kusrini & Luthfi, 2009) :

1. *Data Selection*

   Pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/Cleaning*

   Proses *cleaning* antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data. Pada proses ini dilakukan juga proses *enrichment*, yaitu proses memperkaya data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD.

3. *Transformation*

   Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*.

4. *Data Mining*

   Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.

5. *Interpretation/Evaluation*

   Pola informasi yang dihasilkan dari proses *data mining* diterjemahkan menjadi bentuk yang lebih mudah dimengerti oleh pihak yang berkepentingan.

2.1.5 *Data Mining dalam Kerangka Kerja CRM*

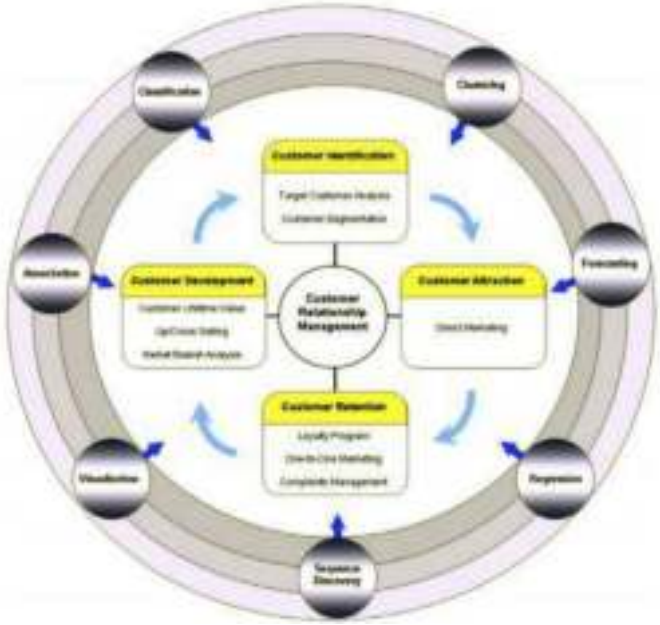
Menurut (Ngai, Xiu, & Chau, 2009) CRM memiliki empat dimensi, yaitu :

1. Identifikasi konsumen (*customer identification*).
2. Membangun daya tarik terhadap konsumen (*customer attraction*).
3. Mempertahankan konsumen yang ada (*customer retention*).
4. Pengembangkan konsumen (*customer development*).

Keempat dimensi dari CRM tersebut dapat dilihat sebagai siklus tertutup dalam *Customer Management System*. Setiap dimensi memiliki tujuan umum yang sama yaitu untuk membangun pemahaman yang lebih mendalam dari konsumen dan meningkatkan nilai konsumen (*customer value*) dari perusahaan. Dalam kaitannya dengan pencapaian tujuan tersebut, teknik *data mining* dapat digunakan untuk menemukan karakteristik dan perilaku konsumen yang tersembunyi dalam *database* yang berukuran besar. Untuk menemukan pola karakteristik dan perilaku konsumen, dapat digunakan metode pemodelan data dalam teknik *data mining*. Beberapa metode pemodelan data yang dapat digunakan, yaitu :

1. *Association*
2. *Classification*
3. *Clustering*
4. *Forecasting*
5. *Regression*
6. *Sequence discovery*
7. *Visualization*

Kerangka kerja teknik *data mining* dalam keempat dimensi CRM dijelaskan melalui Gambar II.3.



Gambar II.3 Kerangka Kerja Teknik *Data Mining* Dalam CRM (Ngai Et Al., 2009)

2.1.6 Clustering

Analisis model *clustering* adalah sebuah teknik dari analisis multivariabel yang digunakan untuk mengelompokkan objek-objek (variabel atau data) sehingga dapat menghasilkan suatu informasi untuk membantu pelaksanaan pengujian terhadap objek dan pada akhirnya dapat menyajikan suatu hipotesis berdasarkan relasi yang terjadi . Prinsip yang digunakan adalah memaksimumkan homogenitas (kesamaan) dalam satukelompok dan juga memaksimumkan heterogenitas (ketidaksamaan) antar kelompok.

Clustering adalah salah satu teknik *unsupervised learning* dimana kita tidak perlu melatih metode tersebut atau dengan kata lain, tidak ada *fase learning*. Tujuan dari metode *clustering* adalah untuk mengelompokkan sejumlah data atau objek kedalam klaster sehingga setiap klaster akan terisi data yang semirip mungkin.

2.1.7 K-*Means*

Algoritma K-*Means* merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K *cluster* yang sudah ditetapkan di awal (Butarbutar, Windarto, Hartama, & Solikhun, 2016) . Algoritam K-*Means* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum Penggunaannya dalam praktek (Parlina, Windarto, Wanto, & M.Ridwan Lubis, 2018). K-*Means* dapat diterapkan pada data yang direpresentasikan dalam r-dimensi ruang tempat. K-*means* mengelompokan set data rdimensi, X= {xi|i=1,...,N} (Sadewo, Windarto, & Hartama, 2017).

Algoritma K-*Means* pengelompokan semua titik data dalam X sehingga setiap titik xi hanya jatuh dalam satu K partisi (Sadewo et al., 2017). Tujuan pengelompokan ini adalah untuk meminimalkan fungsi objek yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Butarbutar et al., 2016).

Parameter yang hrsus dimasukkan ketika menggunakan algoritma K-Means adalah nilai K (Sadewo et al., 2017). Nilai K yang digunkan pada umumnya didasarkan pada informasi yang diketahui sebelumnya mengenai sebenarnya berapa banyak cluster yang muncul dalam X, berapa banyak yang digunakan untuk penerapnnya, atau jenis cluster dicari dengan melakukan percobaan dengan beberapa nilai K. Set representatif cluster dinyatakan C={cj|j=1,…,K. sejumlah K representatif cluster tersebut sebagai *cluster centroid* (titik pusat cluster). Untuk set data dalam X dikelompokan berdasarkan konsep kedekatanatau kemiripan, namun kuantitas yang digunkan untuk mengukurnya adalah ketidak miripan. Metrik yang umum digunakan untuk ketidak miripan tersebut adalah *Euclidean*.

2.1.8 K-Medoids

K-Medoids atau *Partitioning Around Medoids* (PAM) adalah algoritma *clustering* yang mirip dengan K-*Means*. Perbedaan dari kedua algoritma ini yaitu algoritma K-*Medoids* atau PAM menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap cluster, sedangkan K-*Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster* [Kaur, dkk., 2014]. Algoritma K-*Medoids* memiliki kelebihan untuk mengatasi kelemahan pada pada algoritma K-*Means* yang *sensitive* terhadap *noise* dan *outlier*, dimana objek dengan nilai yang besar yang memungkinkan menyimpang pada dari distribusi data. Kelebihan lainnya yaitu hasil proses *clustering* tidak bergantung pada urutan masuk dataset [Furqon, dkk., 2015]. Langkahlangkah algoritma K-*Medoids*:

1. Inisialisasi pusat cluster sebanyak k (jumlah cluster)
2. Alokasikan setiap data (objek) ke clusterterdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan:

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

3. Pilih secara acak objek pada masing-masing cluster sebagai kandidat *medoid* baru.
4. Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat *medoid* baru.
5. Hitung total simpangan (S) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai *medoid*.
6. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoid*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

2.1.9. Validasi

Davies Bouldin Index adalah fungsi rasio dari jumlah antara *cluster scatter* sampai dengan *cluster separation* (Maulik & Bandyopadhyay, 2002). *Davies-Bouldin Index* merupakan metode validasi *cluster* dari hasil *clustering*. Pendekatan pengukuran DBI yaitu memaksimalkan jarak *intercluster* serta meminimalkan jarak *intracluster*. Nilai *purity* adalah kesesuaian antara *cluster* dengan *cluster* ideal, semakin besar nilai *purity* (semakin mendekati 1), semakin baik kualitas *cluster* (Yi, Yang, Qiao, & Xu, 2010). Semakin kecil nilai DB *Index* menunjukan skema *cluster* yang paling optimal. Semakin besar nilai *purity* (mendekati 1) semakin baik kualitas *cluster*.

2.2 Tinjauan Studi

Beberapa penelitian terkait tentang data mining untuk CRM khususnya untuk proses segmentasi pelanggan, serta RFM telah beberapa kali dilakukan Sebelumnya

Tabel II.2 Tinjauan Penelitian Sebelumnya Dari *International Journal*

No	Penelitian	Judul	Metode	Deskripsi
1.	Vahid Golmah (2014) (Golmah, 2014)	*A Case Study of Applying SOM in Market Segmentation of Automobile Insurance Customers*	SOM, CRM	Mencari segmentasi pelanggan asuransi mobil
2.	Prof. Archana Raje, Dr. R.K Srivastava (2014) (Raje & Srivastava,	*The Connection between data mining and segmentation in marketing area*	*Clustering, targetting, positioning*	Diperlukan untuk memahami koneksi antara data mining dan segmentasi dalam pemasaran

	2014)			
3.	Tannane Parsa Kord Aslabi, Rezatoval (2015) (Asiabi & RezaTavoli, 2015)	*A review of different data mining techniques in customer segmentation*	CRM	Menjelaskan CRM dan *data mining* beserta metode-metode yang biasa digunakan untuk segmentasi pelanggan
4.	Ina Maryani, Dwiza Riana, Rachmawati Darma Astuti, Ahmad Ishaq, Sutrisno, Eva Argarini Pratama (Maryani, Riana, Astuti, Ishaq, & Pratama, 2018)	*Customer Segmentation based on RFM model and Clustering Techniques With K-Means Algorithm*	*K-Means clustering*, RFM *Model*	Mencari segmentasi pelanggan yang mendatangkan keuntungan menggunakan K-*Means Clustering*, RFM *Model*
5.	Mediana Aryuni, Evaristus Didik Madyatmadja, Eka Miranda (2018) (Aryuni & Miranda, 2018)	*Customer Segmentation in XYZ Bank using K-Means and K-Medoids Clustering*	RFM, K-*Means*, K-*Medoids*	Membandingkan algoritma K-*Means* dengan K-*Medoids* untuk segmentasi pelanggan BANK XYZ. Hasil penelitian menunjukkan bahwa metode K-*Means*

				mengungguli metode K-*Medoids* berdasarkan jarak *intra cluster* (AWC). Sementara berdasarkan indeks Davies-Bouldin, K-*Means* berkinerja sedikit lebih baik daripada K-*Medoids*

Tabel II.3 Tinjauan Penelitian Sebelumnya dari jurnal nasional

No	Penelitian	Judul	Metode	Deskripsi
1.	Cakra Ramadhan, Yohana Dwi lulu W., Kartina Diah K.W., (2013) (Ramadhana, W, & W, 2013)	*Data mining* dengan algoritma *fuzzy* C-*Means clustering* dalam kasus penjualan di PT Sepatu BATA	*Fuzzy* C-*Means*	Metode perhitungan *fuzzy* C-*Means* cocok digunakan pada aplikasi ini, karena bisa menghasilkan ouput berupa tingkat kelarisan produk pada penjualan sepatu bata
2.	Putu Sukma Kurniawan (2015) (Kurniawan, 2015)	Perancangan data mining untuk analisis keriteria nasabah kredit yang potensial dan	CRM	Manfaat data mining untuk konsep *Customer Relationship Management*

		manfaatnya untuk *customer relationship management* perbankan		perbankan adalah untuk membentuk kelompok nasabah (*customer profilling*) dan untuk bidang marketing dan *customer care*
3.	Beta Estri Adiana, Indah Soesanti, Adhistya Erna Permanasari (2018) (Adiana, Soesanti, & Permanasari, 2018)	Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik *Clustering*	K- *Means* , RFM	Meningkatkan kesetiaan pelanggan dengan mengelompokkan pelanggan menjadi beberapa kelompok dan menentukan strategi pemasaran yang tepat dan efektif untuk setiap kelompok
4	Siti Monalisa (2018) (Monalisa, 2018)	Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-*Means*	RFM, K-*Means*	Mengukur perilaku pembelian pelanggan dengan menggunakan model RFM dengan metode K-*Means*, *silhouette Index*.
5	Arfi Joyendri(2017) (Joyendri,	Strategi *Customer Relationship Management*	RFM, K-*Means*	Menghasilkan pengetahuan yang dihasilkan melalui

	2017)	Untuk Meningkatkan Loyalitas Pelanggan Dan Volume Penjualan Menggunakan Teknik *Clustering* K-*Means*		proses *clustering data mining* yang kemudian dihubungkan dengan strategi pemasaran *customer relationship manajement* yang dapat membantu peningkatan volume penjualan dan loyalitas pelanggan dalam distribusi pemasaran yang selama ini telah dilakukan.
6	Putri Eka Prakasawati, Yulison H. Chrisnanto, Asep Id Hadiana (2019) (Prakasawati, Chrisnanto, & Hadiana, 2019)	Segmentasi Pelanggan Berdasarkan Produk Menggunakan Metode K-*Medoids*	K-*Medoids Clustering*	segmentasi pelanggan berdasarkan produk. Proses pada sistem pengelompokan pelanggan ini menggunakan sebuah algoritma K-*Medoid clustering* untuk mengelompokkan pelanggan berdasarakan

				segmentasi pada produk jumlah pembelian dan area. Dengan data uji sebanyak 6 daerah, 7000 data pelanggan dan 28 produk.

2.3. Tinjauan Obyek Penelitian

Pulsantara merupakan salah satu perusahaan distributor server pulsa elektronik yang berdiri pada tanggal 27 Desember 2013 , didirikan oleh Bayu Setiawan, yang berada pada perumahan Purbalingga. Perusahaan ini memanfaatkan multi chip server guna mendukung proses pengisian pulsa secara otomatis kepada pelanggan. Dengan multi chip server, perusahaan dapat menyediakan semua produk dari pulsa elektronik dengan sistem satu deposit untuk pengisian banyak operator dimana transaksi langsung dilakukan oleh server dan dapat dilakukan 24 jam setiap hari dengan berbasis sms.

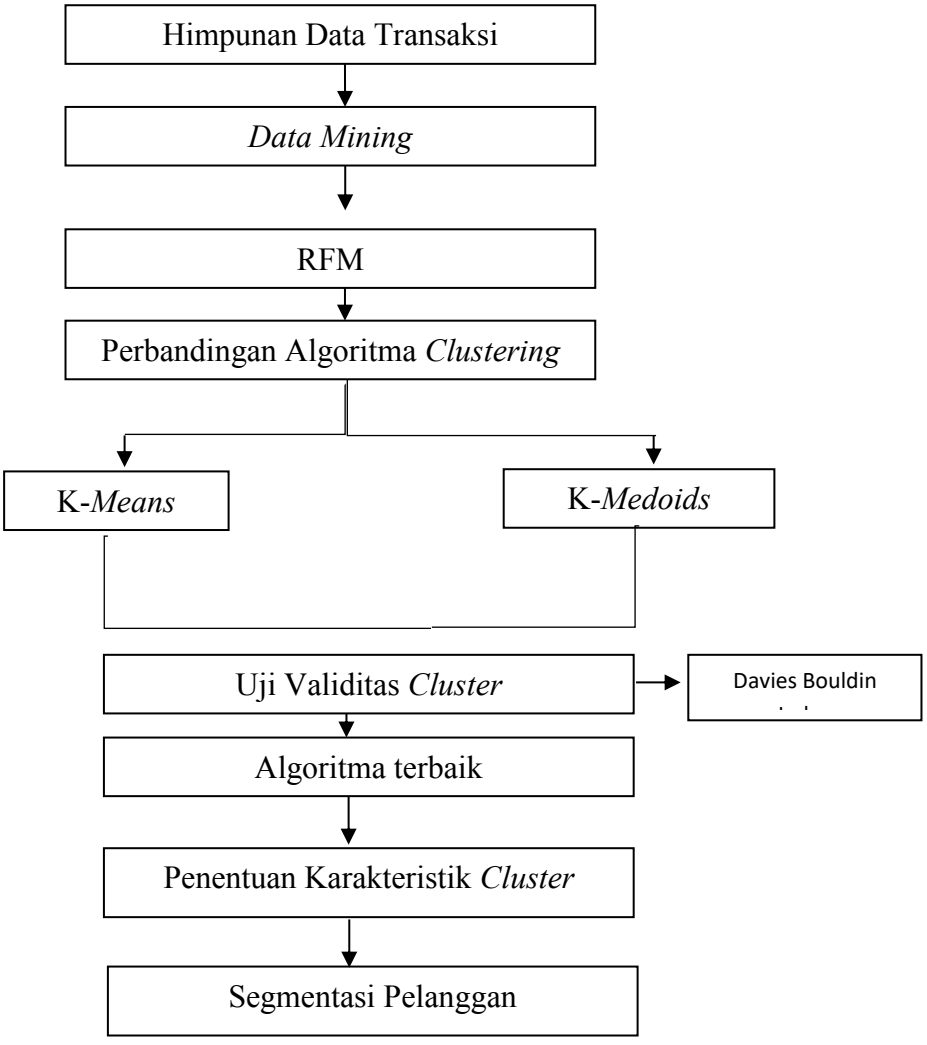
Konsumen yang berminat menjadi pelanggan Pulsantara melakukan pendaftaran dengan mengisi formulir yang berisi data-data calon pelanggan atau agen dan bersedia mengikuti aturan yang diberikan oleh perusahaan. Pelanggan dapat menggunakan pelayanan dari Pulsantara untuk kepentingan usaha atau pribadi. Setiap pelanggan diminta menyimpan sejumlah dana (deposit) di awal sebesar Rp. 100.000. Setiap transaksi pengisian pulsa yang dilakukan oleh pelanggan akan mengurangi saldo deposit yang dimiliki oleh pelanggan. Apabila saldo deposit pelanggan tidak mencukupi maka transaksi pengisian pulsa akan gagal.

BAB III

METODOLOGI PENELITIAN

3.1 Metodelogi Penelitian

Penelitian ini dilakuakan melalui beberapa tahap dari pengumpulan data dari perusahan, peneglompokan data *(clustering)* menggunakan algoritma K-Means dan K-*Medoids*, *validasi cluster* dan melakukan perbandingan dari hasil perhitungan. Berikut langkah penelitian pada Gambar III.1



GambarIII.1 Langkah Penelitian

3.1.1. Jenis Penelitian

Pendekatan penelitian dibagi menjadi dua yaitu pendekatan kualitatif dan pendekatan kuantitatif. Metode penelitian kualitatif berhubungan dengan penilaian subjektif dari sikap, pendapat, dan perilaku. Teknik yang umum digunakan adalah *interview* pada kelompok tertentu dan wawancara yang mendalam (Siyoto & Sodik, 2015).

Metode penelitian kuantitatif digunakan untuk meneliti pada *populasi* atau *sample* tertentu, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif/statistik dengan tujuan untuk menguji hipotesis yang telah ditentukan. Metode penelitian kuantitatif disebut juga dengan metode *discovery* karena dengan metode ini dapat ditemukan dan dikembangkan berbagai iptek baru (Siyoto & Sodik, 2015).

Metode penelitian kuantitatif dapat dibagi menjadi tiga sub kategori yaitu *inferentia*, *experimental*, dan *simulation*. Desain riset yang digunakan dalam penelitian ini adalah eksperimen. Metode penelitian eksperimen dapat diartikan sebagai metode penelitian yang digunakan untuk mencari pengaruh perlakuan tertentu terhadap yang lain dalam kondisi yang terkendalikan (Siyoto & Sodik, 2015). Misalnya pengaruh ruang kerja terhadap produktivitas kerja karyawan.

Metode eksperimen dikarakterisasi dengan lebih banyak mengontrol lingkungan penelitian, dalam hal ini beberapa *variabel* dimanipulasi untuk mengobservasi efek terhadap *variabel* lain. Dalam metode eksperimen peneliti mengukur efek dari eksperimen yang dilakukan dengan sengaja. Tujuan akhir dari eksperimen adalah untuk mengeneralisasi hubungan *variabel* sehingga dapat diterapkan pada populasi yang lebih luas diluar laboratorium (Kothari, 2004).

Metode eksperimen dibagi menjadi dua, yaitu eksperimen absolut dan eksperimen komparatif. Eksperimen absolut mengarah pada dampak yang dihasilkan dari eksperimen, sedangkan eksperimen komparatif membandingkan dua objek yang berbeda (Kothari, 2004). Dalam penelitian ini digunakan metode eksperimen komperatif.

3.1.2. Metode Pengumpulan Data

1. Sumber Data

Metode pengumpulan data untuk mendapatkan sumber data dibagi menjadi dua, yaitu data primer dan data sekunder. Data primer adalah data yang dikumpulkan pertama kali, dan untuk melihat apa yang sesungguhnya terjadi melalui observasi, *interview*, *kuesioner*, dll. Data sekunder adalah data yang telah dikumpulkan dan dianalisis oleh orang lain baik yang telah dipublikasikan maupun yang belum dipublikasikan, misalnya dari dokumentasi, literatur, buku, jurnal, dan informasi lainnya yang ada hubungannya dengan masalah yang diteliti.

Dalam penelitian ini metode pengumpulan data untuk mendapatkan sumber data yang digunakan adalah metode pengumpulan data primer. Data utama diperoleh dari data transaksi pulsa di perusahan Pulsantra, sedangkan data pendukung didapatkan dari buku, jurnal dan publikasi lainnya.

2. Sampel Penelitian

Sampel dari Penelitian ini adalah data transaksi dari "Perusahaan Pulsantara".

3.1.3. Eksperimen

Dalam melakukan eksperimen ini, penulis menggunakan model CRISP-DM (*Cross- Industry Standard Process for Data Mining)* yang terdiri dari enam tahap, yaitu (Maimon & Rokach, 2010):

1. Tahap *Business understanding*

Tahapan ini dilakukan sebelum segmentasi pelanggan. Tahapan ini terdiri dari tiga bagian, yaitu :

a. Pemahaman terhadap tujuan bisnis
b. Penilaian sesuatu
c. Penerjemahan tujuan bisnis dalam data Mining

2. Tahap *data understanding*

Data transaksi diperoleh dari Perushan Pulsantara, yang terdiri dari dari 85010 data data transaksi selama 2 tahun yaitu tahun 2017 dan 2018.

3. Tahap *data preparation*

Persiapan data merupakan tahapan dimana peneliti melakukan *query* data untuk mendapatkan data yang siap untuk di *mining*. Tahapan ini menggunakan *Microsoft Excel*.

4. Tahap *data modelling*

Segmentasi pelanggan menggunakan *proses clustering-data mining* metode *K-Means* dan *K-Medoids*. Proses *clustering* ini dilakukan dengan tiga *variabel*, yaitu *Recency*-Frekuensi-*Monetary*. Dalam konsep RFM ini, setiap dimensi memberikan informasi yang unik mengenai perilaku transaksi masing-masing pelanggan yang dimiliki perusahaan:

a. *Recency* – Berdasarkan penelitian yang telah dilakukan bertahun-tahun, statistik memperlihatkan bahwa pelanggan yang belakangan ini melakukan transaksi pembelian akan cenderung melakukannya lagi dalam waktu dekat.
b. *Frequency* – Sejarah memperlihatkan bahwa pelanggan yang secara berkala melakukan pembelian, akan cenderung melakukannya lagi di kemudian hari.
c. *Monetary Value* – Pelanggan yang telah mengalokasikan cukup banyak dana untuk melakukan pembelian akan cenderung melanjutkan kebiasaan membelinya. Tabel III.1 menunjukkan pengertian ketiga *variabel* tersebut.

Tabel III.1 *Variabel* Proses *Clustering*

Variabel	Keterangan
Recency	Jarak pemakaian (interval) antara pembelian terakhir pada periode tertentu
Frequency	Banyaknya jumlah pembelian dalam periode tertentu
Monetary	Jumlah uang selama periode tertentu

Berikut adalah tahapan penentuan variabel *recency*-frekuensi-*monetary*:

a. Menghitung frekuensi pembelian *customer* setiap bulan dan menjumlahkannya untuk mendapatkan frekuensi pembelian *customer* dalam 2 tahun.
b. Mengitung *recency* setiap *customer* atau *agen*
c. Menghitung total beli selama dua tahun

Setelah proses tersebut berikutnya adalah pembentukan *data processing* pada studi kasus ini bertujuan untuk mencari bentuk aturan data yang sesuai dengan proses data mining. Proses ini dilakukan melalui pengubahan isi data pada

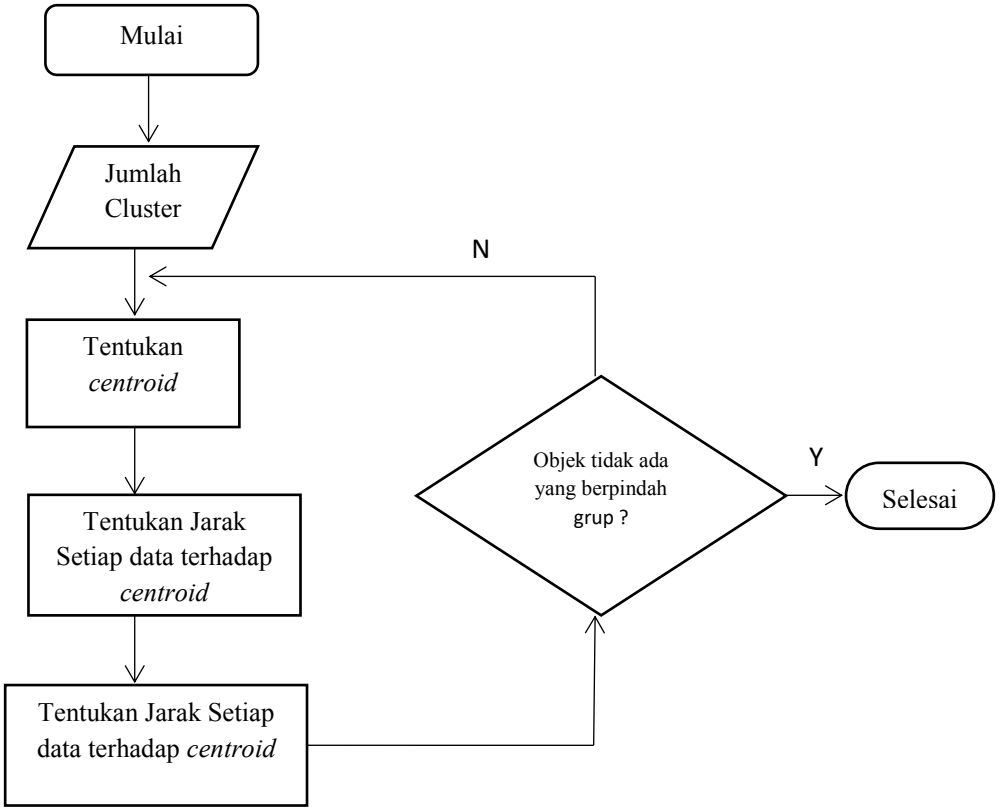
atribute yang ada pada data transaksi penjualan industri knalpot *doctor speed* Purbalingga dengan melakukan proses *konversi* data dari data sumber menjadi data *numerik* sesuai dengan tabel *konversi* sehingga didapatkan angka yang *valid*. Hasil dari proses tersebut kemudian diintegrasikan menjadi sebuah tabel baru sehingga nantinya didapatkan sebuah tabel nilai transaksi yang akan digunakan untuk melakukan proses menjadi beberapa *cluster.* Proses *clustering* ini bertujuan untuk mengidentifikasi dan menentukan jumlah pelanggan pada kelas *profitable* dan *less/unprofitable customer* Setelah mendapatkan tiga variabel tersebut dilakukan proses *scoring*.

Tabel III. 2 Menunjukkan *Score Recency-Frequent-Monetary*

Score	R		F		M	
5	Sangat Pendek	<2bln	Sangat Tinggi	>8000 kali	Sangat Banyak	>30juta
4	Pendek	3-8 bln	Tinggi	8000 - 6000 kali	Banyak	20 - 30 juta
3	Biasa	9-12bln	Biasa	5999 - 3000 kali	Normal	10 - 20 juta
2	Jauh	13 -14bln	Rendah	2999 - 1000 kali	Sedikit	5 - 10 juta
1	Sangat Jauh	>14	Sangat Rendah	<1000 kali	Sangat Sedikit	<5 juta

Score 5 merupakan *score* tertinggi yang artinya terbaik dari data yang ada. Untuk *recency*, *score* 5 menunjukkan *interval* terpendek. Kemudian untuk frekuensi dan *Monetary score* 5 menunjukkan nilai frekuensi dan nilai *monetary* terbesar. Namun, untuk *recency* pada *score* 1 menunjukkan interval terpanjang. Kemudian untuk frekuensi dan *Monetary score* 1 menunjukkan nilai frekuensi dan *Monetary* terkecil. Setelah tahapan *scoring*, dilakukan proses *clustering* dengan *K-Means* dan *K-Medoids* setelah mendapatkan hasil dari kedua algoritma tersebut peneliti akan melakukan perbandingan algorithma mana yang paling baik antara *K-Means* dan *K-Medoids* .

Melakukan Algorithma *K-Means*



Gambar III.2 *Flowchart* Cara Kerja Algorithma K-Means

Dari *flowchart* diatas dapat di jelaskan kerja algoritma K-Means sebagai berikut :

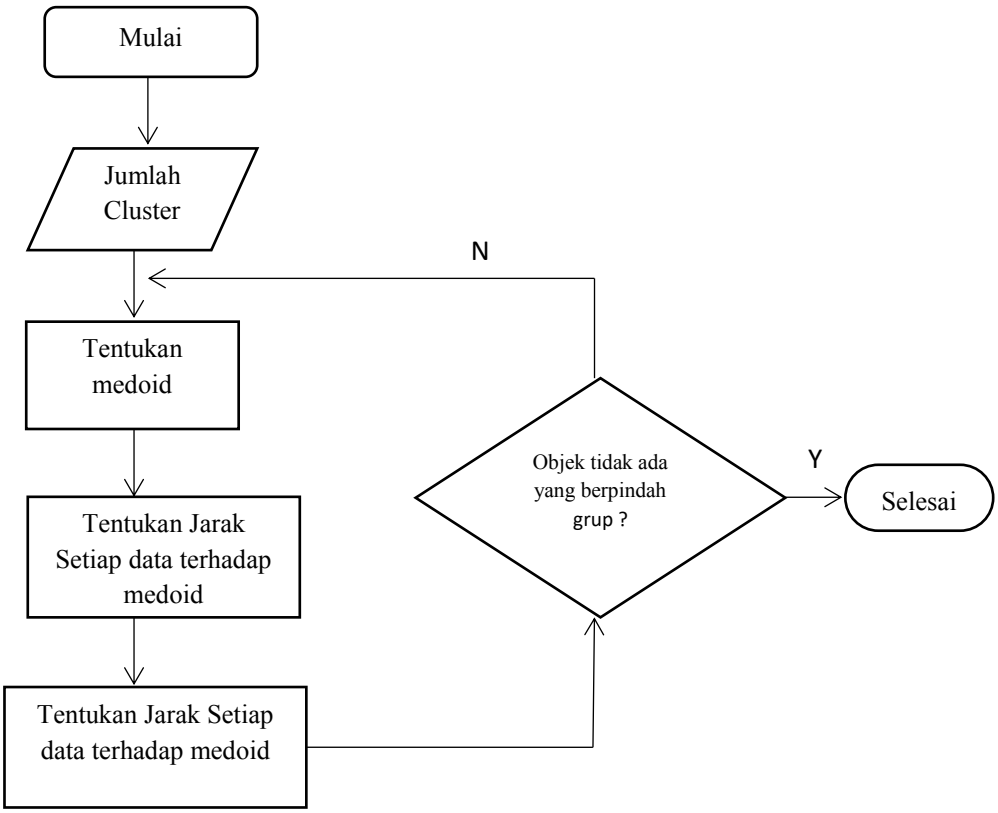
1. Tentukan K sebagai jumlah *cluster* yang ingin dibentuk.
2. Pilih K *centroid* (titik pusat *cluster*) awal secara random. Dalam menentukan n buah pusat *cluster* awal dilakukan pemilihan bilangan random yang merepresentasikan urutan data *input*. Pusat awal *cluster* didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan merandom pusat awal dari data.
3. Hitung jarak setiap data ke masing-masing *centroids*. Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidian distance*. Algoritma perhitungan jarak data dengan pusat *cluster*. Langkah-langkahnya, yaitu:
   a. Ambil nilai data dan nilai pusat *cluster*
   b. Hitung *Euclidian distance* data dengan tiap pusat *cluster*. *Euclidian Distance* merupakan jarak yang didapat dari perhitungan antara semua N

data dengan K *centroid* dimana akan memperoleh tingkat kedekatan dengan kelas yang terdekat dengan populasi data tersebut. Jarak *euclidian ini berguna* untuk menandai adanya persamaan antar tiap *cluster* dengan jarak minimum dan mempunyai persamaan yang lebih tinggi. Persamaan dari *Euclidian matrik* antara titik dan titik adalah:

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

4. Setiap data memilih *centroids* yang terdekat. Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.
5. Tentukan posisi *centroids* yang baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada *centroid* yang sama. Pusat *cluster* yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh *User* atau hasil yang dicapai sudah konvergen (pusat *cluster* baru sama dengan pusat *cluster* lama).
6. Jika posisi *centroids* baru dengan *centroids* yang lama tidak sama, maka ulangi kembali dari langkah 3.

Melakukan Algorithma *K-Medoids*



Gambar III.3 *Flowchart* Cara Kerja *Algorithma* K-Medoids

Dari *flowchart* diatas dapat di jelaskan kerja algoritma K-*Medoids* sebagai berikut :

1. Inisialisasi pusat *cluster* sebanyak k (jumlah *cluster*)
2. Alokasikan setiap data (objek) ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan:

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}; i=1,2,3,\ldots,n$$

3. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat medoid baru.
4. Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat medoid baru.

5. Hitung total simpangan (S) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika S < 0, maka tukar objek dengan data *cluster* untuk membentuk sekumpulan k objek baru sebagai *medoid*.
6. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan medoid, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

3.1.4. Validasi

Davies Bouldin Index adalah fungsi rasio dari jumlah antara *cluster scatter* sampai dengan *cluster separation* (Maulik & Bandyopadhyay, 2002). *Davies-Bouldin Index* merupakan metode validasi *cluster* dari hasil *clustering*. Pendekatan pengukuran DBI yaitu memaksimalkan jarak *intercluster* serta meminimalkan jarak *intracluster*. Nilai *purity* adalah kesesuaian antara *cluster* dengan *cluster* ideal, semakin besar nilai *purity* (semakin mendekati 1), semakin baik kualitas *cluster* (Yi, Yang, Qiao, & Xu, 2010). Semakin kecil nilai DB *Index* menunjukan skema *cluster* yang paling optimal. Semakin besar nilai *purity* (mendekati 1) semakin baik kualitas *cluster*.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

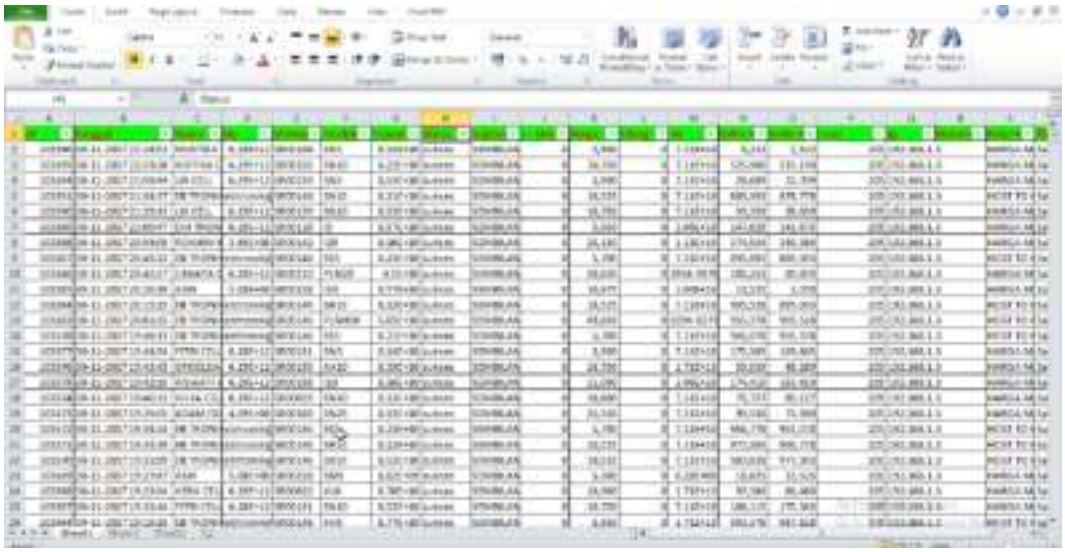
4.1 Hasil Penelitian

Pada bagian hasil dan pembahasan akan diuraikan hasil-hasil dari tahapan penelitian dan dilakukan pula pembahasan dari setiap hasil tersebut. Data yang digunakan dalam penelitian ini adalah data transaksi pelanggan dari tahun 2017 sampai 2018. Dalam penelitian ini dilakukan analisa terhadap hasil *clustering* dan segmentasi data transaksi dalam periode 2 tahun. Hasil *clustering* untuk algoritma *K-Means dan K-Medoids* akan diuji dengan *Davies Bouldin Index,* untuk menguji *validitas cluster.*

4.1.1 Persiapan Data Mining

4.1.1.1.Himpunan Data Transaksi

Pada perusahan Pulsantara terdapat banyak himpunan data tarnsaksi yang belum di manfaatkan sehingga peneliti tertarik untuk memanfaatkan himpunan data tansaksi yang ada untuk menggali informasi tentang loyalitas pelanggan di perusahan Pulsantara



Gambar IV.1 Database transaksi Pulsantara

4.1.1.2. Data Preparation

Pada tahap persiapan data, peneliti memilih atribut yang akan diguankan dan membuang atribut yang tidak dibutuhkan untuk mendapatkan data yang siap diolah seperti pada **Tabel IV.1**.

Tabel IV.1 Data Set Transaksi Pulsantara

Nama Agen	R	F	M
C001	31-Dec-18	4070	Rp45,819,008
C002	18-Jun-18	678	Rp11,756,612
C003	11-Apr-18	704	Rp9,050,500
C004	31-Dec-18	72	Rp1,052,500
C005	31-Dec-18	154	Rp1,405,100
C006	31-Dec-18	60	Rp627,100
C007	31-Dec-18	858	Rp10,802,164
C008	31-Dec-18	216	Rp1,909,500
.....	.....	.....	.....
.....	.....	.....	.....
.....	.....	.....	.....
C132	16-Dec-18	1522	Rp46,019,740
C133	30-May-17	680	Rp8,175,400

4.1.1.3.Data Modeling

Pada tahapan ini peneliti akna melakukan proses clustering dengan menggunakan metode 2 metode clustering yaitu *K-means* dan *K-Medoids*. Namun sebelum melakuakn tahap clustering perlu dilakukan pembobotan terlebih dahulu. Untuk pembobotan diambil dari data *recency*, *frequensi*, *Monetary* harus dibagi ke dalam 5 skala/*score*. Tabel IV.2 menunjukkan skala *score recency*, *frekuensi*, dan *Monetary* untuk data penjualan tahun 2017-2018.

Berikut proses menentukan skala/bobot yang digunakan

a. Hitunglah *mean* (nilai rata-rata) dari seluruh data dengan rumus :

$$\bar{x} = \frac{\sum x}{n} \text{ …….. (1)}$$

Keterangan :

$\overline{x}$: Rata-rata

$\sum x$: Hasil penjumlahan data

n : Banyak data

b. Hitunglah standar deviasi dari seluruh data dengan rumus :

$$S = \sqrt{\sum \frac{(x-\overline{x})^2}{n-1}} \text{ …….. (2)}$$

Keterangan :

S : Standar Deviasi

X : Nilai data

$\overline{x}$: Nilai rata-rata

n : Banyak data

berdasarkan rumus standar nilai mean dan standar deviasi yang diperoleh makan dihasilkan tabel pembobotan *Recency, Frequency dan Monetary* secara berturut-turut pada **Tabel IV.2.** sampai **Tabel IV.5**:

Tabel IV.2. Kategori *Recency*

Score	R	
5	Sangat Pendek	R < $\overline{x} - S$
4	Pendek	$\overline{x} - S$ < R < $\overline{x}$
3	Biasa	$\overline{x}$ < R < S
2	Jauh	S < R < $\overline{x} + S$
1	Sangat Jauh	R> $\overline{x} + S$

Tabel IV.3 Kategori *Frequensi*

Score	F	
5	Sangat Tinggi	F> $\overline{x} + S$
4	Tinggi	S < F < $\overline{x} + S$
3	Biasa	$\overline{x}$ < F < S
2	Rendah	$\overline{x} - S$ < F < $\overline{x}$
1	Sangat Rendah	F < $\overline{x} - S$

Tabel IV.4. Kategori *Monitary*

Score	M	
5	Sangat Banyak	M > $\overline{x} + S$
4	Banyak	S < M < $\overline{x} + S$
3	Normal	$\overline{x}$ < M < S
2	Sedikit	$\overline{x} - S$ < M < $\overline{x}$
1	Sangat Sedikit	M < $\overline{x} - S$

Tabel IV.5. Kreteria Pembobtan RFM

Score	R		F		M	
5	Sangat Pendek	<2bln	Sangat Tinggi	>8000 kali	Sangat Banyak	>30juta
4	Pendek	3-8 bln	Tinggi	8000 - 6000 kali	Banyak	20 - 30 juta
3	Biasa	9-12bln	Biasa	5999 - 3000 kali	Normal	10 - 20 juta
2	Jauh	13 -14bln	Rendah	2999 - 1000 kali	Sedikit	5 - 10 juta
1	Sangat Jauh	>14	Sangat Rendah	<1000 kali	Sangat Sedikit	<5 juta

Setelah menentukan kreteria pembobotan yang tertera pada **Tabel IV.5** maka tahap selanjutnya adalh memberikan nilai bobot pada data yang ditunujkan pada **Tabel IV.1**

Tabel IV.6. Nilai Bobot pada RFM

KODE CUSTOMER	R	F	M
C001	5	3	5
C002	4	1	3
C003	3	1	2
C004	5	1	1
C005	5	1	1
C006	5	1	1
.....	.....	.....	.....
.....	.....	.....	.....
.....	.....	.....	.....
C132	5	2	5
C133	1	1	2

Tabel IV.6 menunjukkan transformasi data setelah dilakukan *score recency*, *frekuensi*, dan *Monetary* data transaksi Pulsantara dan sudah siap untuk dilakukan proses *data mining*.

Setelah semua data transaksi ditransformasi ke dalam bentuk angka, maka data-data tersebut telah dapat dikelompokan dengan menggunakan algoritma *K-Means* dan *K-Medoids.*

4.2. Pembahasan

Pada tahap ini peneliti menggunakan *software Rapid miner* untuk mengolah data yang sudah diberikan pembobotan. Langkah-langkah sebagi berikut :

a. Buka *software rapid miner*, masukan data *training* yang telah disiapkan pada tahap data *preparation*, pada tab operators pilih Import – Data – Read Excel tarik ke area *main process*. Pada area parameters klik Import Configuration Wizard.



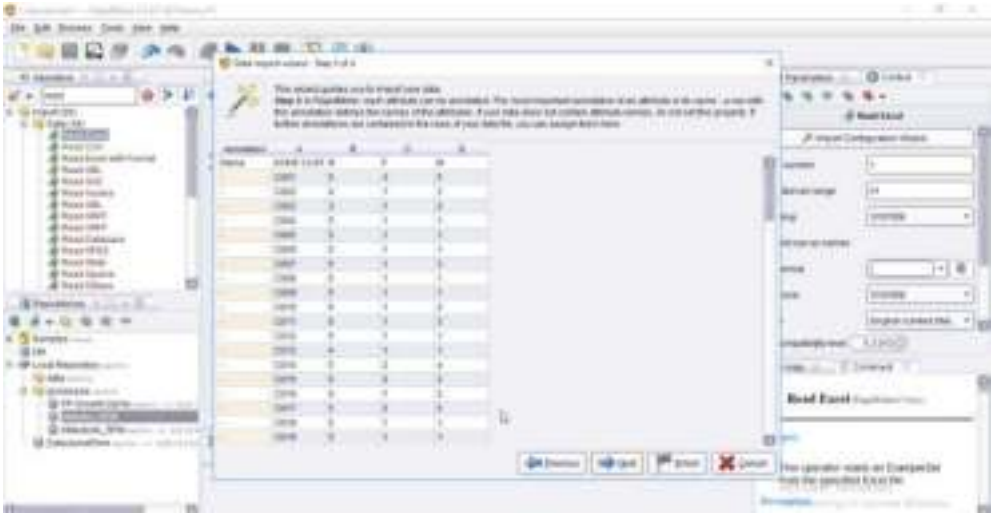
Gambar IV.2 Import data set pada lebar kerja Rapid Miner

b. Tampil jendela *Import Configuration Wizard*, pada langkah ini tentukan data berformat .xls yang akan digunakan sebagai *data training* kemudian klik tombol *next.*



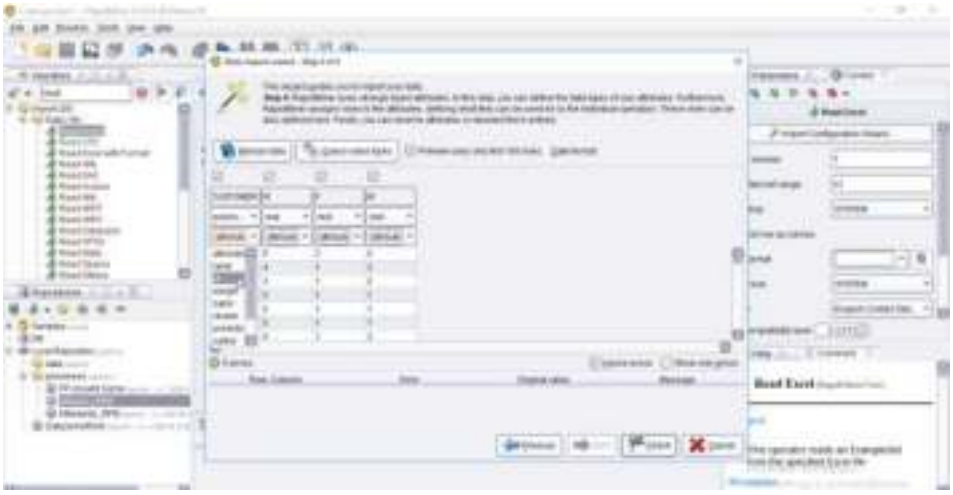
Gambar IV.3 data set berupa excel yang akan *di-import*

c. Langkah selanjutnya yaitu memilih *sheet* yang akan digunakan sebagai data *training*



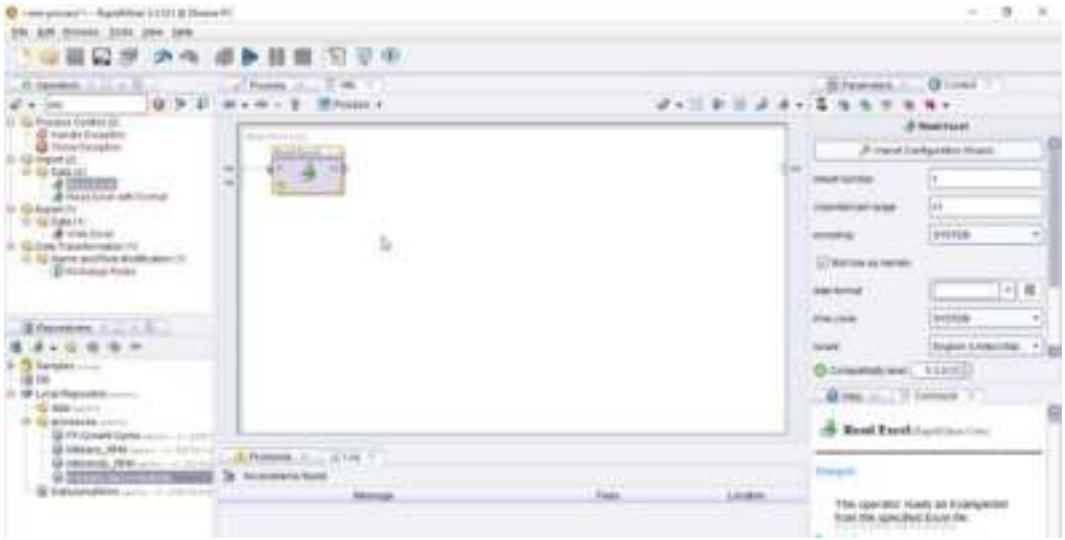
Gambar IV.4 Tampilan sheet data

d. langkah ke selanjutnya menentukan anotasi dari setiap *tupel dan atribut*, langkah ini dapat diabaikan. Kemudian klik *Next*.



Gambar IV.5 Menentukan Anotasi Dari Setiap *Tupel* dan Atribut

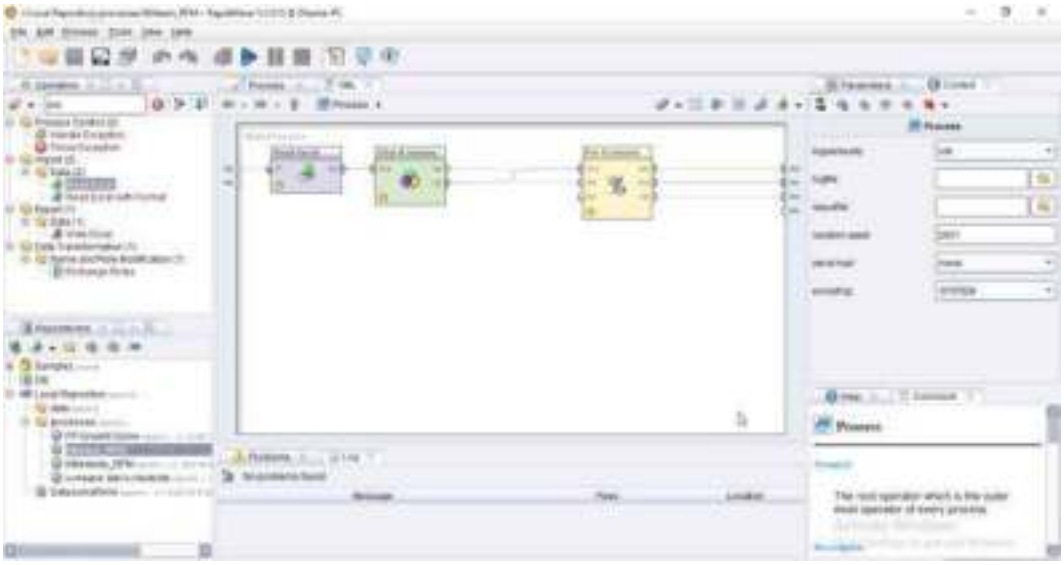
e. Langkah ke empat yaitu menentukan nama atribut dan tipe data dari masing-masing atribut. Pada atribut *Customer* ubah tipe data menjadi label. Klik *finish*



Gambar IV.6 Data Yang sudah di masukan ke read excel

4.2.1. Pengolahan data menggunakan Algorithma *K-Means*

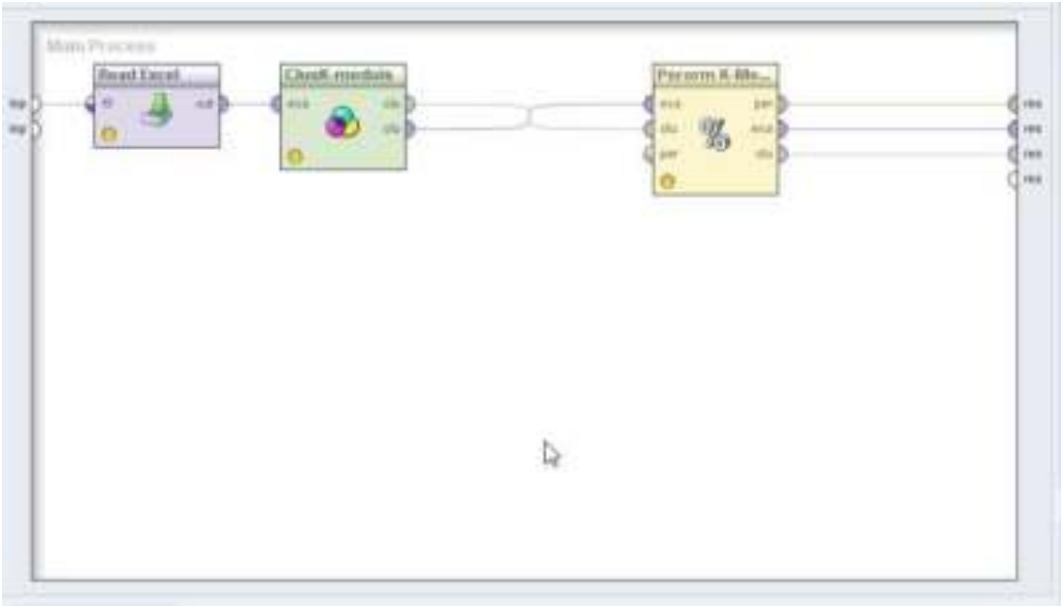
Setelah data training disiapkan, proses pembuatan model dapat dilakukan dengan cara pilih tab *Operators* - *Modelling – Clustering* and *Segmentation – K-means*. Geser *K-Means* ke area *main process* dan hubungkan seperti terlihat pada gambar berikut:



Gambar IV.7 Pemodelan K-Means

4.2.2. Pengolahan data menggunakan Algorithma *K-Means*

Setelah data training disiapkan, proses pembuatan model dapat dilakukan dengan cara pilih tab *Operators* - *Modelling – Clustering* and *Segmentation – K-Medoids*. Geser *K-Medoids* ke area *main process* dan hubungkan seperti terlihat pada gambar berikut:



Gambar IV.8 Pemodelan *K-Medoids*

4.2 Pembahasan

Sebelumnya telah dilakukan proses *modeling clustering* terhadap customer Pulsantara menggunakan Algoritma K-Means dan K-Medoids berdasarkan variabel RFM yang telah diperoleh dari proses identifikasi KM pada perusahaan. Sebelum melakukan proses clustering ditentukan terlebih dahulu jumlah cluster yang ingin dibentuk, yaitu sebanyak 4 cluster sampai dengan 6 cluster. Dalam menentukan jumlah cluster ini telah dilakukan diskusi terlebih dahulu dengan pihak perusahaan. Penentuan jumlah cluster tersebut didasarkan pada interpertasi subjektif. Interpretasi ini didasarkan bahwa jika semakin banyak cluster yang terbentuk maka *knowledge discovery in data base* yang terbentuk akan semakin general dan tidak spesifik. Sedangkan jika *cluster* yang dibentuk terlalu sedikit memungkinkan terdapat anggota yang tidak mewakili kelompok. Dipilih jumlah 4 *cluster* ini diharapkan dapat memberikan hasil yang optimal pada proses *clustering customer*. Setelah ditentukan jumlah cluster yang akan dicoba atau di test untuk mengetahui algoritma terbaik melalui nilai DBI.

4.2.1. Hasil *Cluster Optimal*

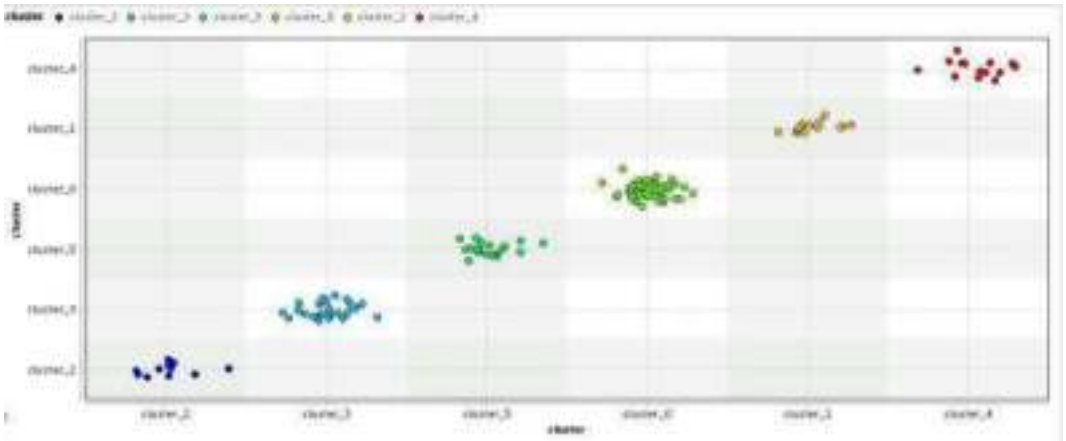
Dari hasil pengolahan data yang dilakukan berdasarkan *dataset* transaksi pelanggan menggunakan *K-Means* dan *K-Medoids* melalui 4 *iterasi* sampai 6 *iterasi* menghasilkan nilai DBI setiap *cluster* dari masing- masing algoritma seperti pada tabel IV.7 .

Tabel IV.7 Perbandingan *Performa Cluster*

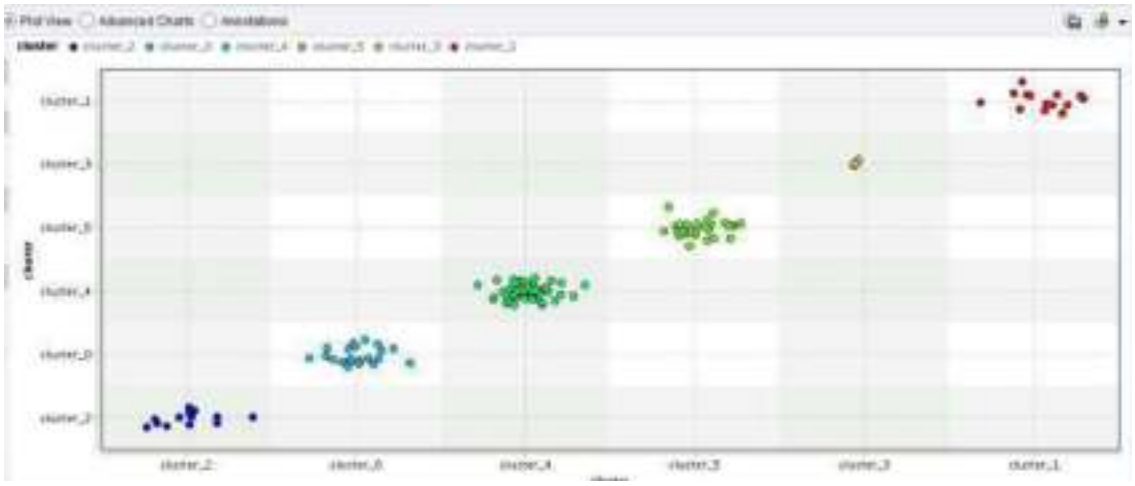
Jumlah Cluster	DB-*index*	
	K-Means	K-Medoids
4	0.816	0.820
5	0.680	0.797
6	0.843	0.660

Hasilnya, untuk performa cluster terbaik ada pada algoritma *K-Medoids* dengan jumlah 6 *cluster* karena memiliki nilai *Davies Bouldien Index* (DI) yang paling kecil diantara jumlah cluster yang lain yaitu *0,660* tercantum dalam Tabel IV.7

Pada Gambar IV.9 dan Gambar IV.10 merupakan penyebaran data hasil clustering dari ke-4 cluster yang terbetuk menggunakan algorithma *K-Means* dan *K-Medoids*, dari grafik tersebut dapat dilihat bahwa anggota dari setiap cluster terpisah kedalam setiap masing-masing cluster, tidak ada anggota cluster yang termasuk kedalam 2 cluster. Hal ini menunjukan bahwa anggota dari hasil clustering yang terbetuk sudah dapat mewakili setiap kelompok.



Gambar IV.9 Grafik Hasil Pengklasteran menggunakan K-Means



Gambar IV.10 Grafik Hasil Pengklasteran menggunakan K-Medoids

Tabel IV. 1 Anggota Cluster dan Validasi

Clustering	K-Means		K-Medoids	
	Jumlah anggota Cluster	Hasil Validasi DB	Jumlah anggota Cluster	Hasil Validasi DB
Cluster 1	2	0. 450%	12	0.632%
Cluster 2	107		2	
Cluster 3	2		74	
Cluster 4	22		45	

Suatu *cluster* dinyatakan *konvergen* apabila sudah tidak terjadi adanya perubahan atau perpindahan anggota dari satu *cluste*r ke *cluster* lain. Selain itu, cluster yang telah *konvergen* juga ditandai oleh tidak adanya perubahan nilai *DB-index*. Dari kedua skenario ujicoba menggunakan 4 sampai 6 jumlah cluster dari algorithma *K-Means* dan *K-Medoids* yang telah dilakukan, akan dibuat suatu peringkat berdasarkan nilai DB-*index* terbaik. Tingkat optimal suatu *cluster* dapat diukur Nilai *DB Index*. Nilai *DB Index* (Yi, Yang, Qiao, & Xu, 2010) untuk jumlah *cluster* yang terbaik adalah yang memiliki index terkecil diantara *cluster* lainnya. Maka dapat ditarik kesimpulan anatar Algorithma K-Means dan Algorithma K-Medoids untuk segmentasi pelanggan dengan jumlah 6 cluster adalah algorithma Algorithma K-Medoids yaitu 0.660%.

4.2.2 Penentuan Karakteristik Pelanggan

Setelah dilakukan pemilihan algorithma terbaik yaitu K-Medoids dan proses clustering yang diperoleh hasil dari Ke-6 cluster yang terbentuk memiliki karakteristik berbeda-beda, sehingga dapat diketahui cluster mana yang merupakan kelompok customer yang potensial dan yang tidak. Tahap selanjutnya adalah mentukan profil customer yang telah terbentuk, untuk menentukan profil pelanggan dari ke-6 cluster yang terbentuk, dibawah ini akan dijelaskan hasil clustering dengan menggunakan interpretasi subjektif dalam mengidentifikasi pola data yang terbentuk atau yang disebut sebagai knowledge discovery in data mining berdasarkan nilai variabel RFM pada setiap cluster.

Cluster 1 yang memiliki 27 anggota, . Nilai rata-rata recency pada cluster 1 adalah 59 hari atau 1 bulan lebih 19 hari. Rata-rata frekuensi pembelian yang dilakukan customer cluster 1 dapat dikatakan berada pada nilai frekuensi yang rendah, yaitu hanya 1532 kali transaksi. Sedangkan untuk rata-rata nilai monetary yang dikeluarkan pelanggan sebesar Rp17,614,914, nilai tersebut menjadi nilai monetary kedua terendah dari keempat cluster. Yang mempengaruhi cluster 1 adalah frekuensi dan monetary. Berdasarkan identifikasi knowledge yang diperoleh, cluster 1 termasuk kedalam jenis Typical Customer, jenis customer ini memiliki nilai recency , frekuensi dan monetary rata-rata, sehingga jenis customer seperti ini akan memberikan sedikit keuntungan bagi perusahaan.

Tabel IV.2 Hasil Cluster 1

No	Nama Agen	Tanggal terahir Pembelian	R	F	M
1	C002	18-Jun-18	206	678	Rp11,756,612
2	C007	31-Dec-18	10	858	Rp10,802,164
3	C011	30-Dec-18	11	634	Rp10,568,674
4	C014	31-Dec-18	10	1984	Rp20,234,624
5	C015	31-Dec-18	10	1092	Rp13,753,840
6	C020	31-Dec-18	10	864	Rp13,966,782
7	C021	21-Dec-18	20	1384	Rp19,043,260
8	C022	31-Dec-18	10	1196	Rp17,561,530
9	C026	31-Dec-18	10	2770	Rp28,604,980
10	C028	31-Dec-18	10	2356	Rp26,896,754
11	C032	31-Dec-18	10	1064	Rp8,522,220

12	C037	30-Dec-18	11	1086	Rp11,672,432
13	C038	31-Dec-18	10	2024	Rp22,007,962
14	C042	31-Dec-18	10	1332	Rp18,981,256
15	C045	31-Dec-18	10	1492	Rp29,999,754
16	C048	12-Apr-18	273	1330	Rp15,337,972
17	C050	31-Dec-18	10	1880	Rp19,399,862
18	C054	11-Nov-18	60	2436	Rp25,860,596
19	C060	29-Oct-18	73	1050	Rp12,880,018
20	C064	9-Dec-18	32	1030	Rp11,459,380
21	C067	15-Jul-18	179	1480	Rp13,407,910
22	C070	30-Aug-18	133	1352	Rp11,772,114
23	C075	15-Jul-18	179	2482	Rp24,691,970
24	C082	19-Jun-18	205	2936	Rp23,598,090
25	C104	11-Dec-18	30	1508	Rp15,270,842
26	C107	31-Dec-18	10	1752	Rp27,453,092
27	C116	10-Nov-18	61	1330	Rp10,098,000
Rata-rata			59	1532	Rp17,614,914

Cluster 2 memiliki 33 anggota, jika dilihat pada tabel VI.3 cluster 2 memiliki nilai rata-rata recency jauh yaitu 394 hari atau 1 tahun 1 bulan 4 hari , frekuensi yang rendah yaitu 373 tarnsaksi dan monetary Rp3,515,116. Berdasarkan knowledge yang diperoleh dari data transaski Pulsantara, cluster 2 ini merupakan customer yang kurang berpotensil memberikan keuntungan bagi Pulsantra, hal ini dikarenakan customer pada cluster 2 memiliki nilai rata-rata monetary yang sedikit sehingga kurang memberikan keuntungan bagi Pulsantara , selain itu juga dikarenakan customer pada cluster ini melakukan transaksi terakhir pada waktu yang cukup lama. Cluster ini dapat dikategorikan kedalam jenis Occational customer.

Tabel IV.3 Hasil Cluster 2

No	Nama Agen	Tanggal terahir Pembelian	R	F	M
1	C003	11-Apr-18	274	704	Rp9,050,500
2	C052	11-Jan-18	364	1384	Rp17,263,320
3	C057	10-Jan-18	365	776	Rp7,997,548
4	C058	10-Apr-18	275	836	Rp8,456,038
5	C061	10-May-18	245	82	Rp701,594
6	C086	13-May-18	242	56	Rp508,150
7	C087	5-May-18	250	214	Rp1,851,150
8	C091	14-May-18	241	324	Rp2,593,716
9	C092	26-Apr-18	259	210	Rp4,052,960
10	C093	30-Apr-18	255	176	Rp4,146,300
11	C094	22-Apr-18	263	244	Rp2,597,300
12	C095	3-Feb-18	341	58	Rp601,950
13	C096	15-Mar-18	301	26	Rp196,500
14	C099	13-Jan-18	362	66	Rp689,100
15	C100	26-Jan-18	349	12	Rp395,200
16	C101	19-Jan-18	356	60	Rp744,500
17	C102	25-Jan-18	350	56	Rp462,750
18	C103	3-Mar-18	313	68	Rp1,907,260
19	C105	12-Mar-17	669	708	Rp7,085,960
20	C106	12-May-17	608	486	Rp4,437,538
21	C109	24-Jul-17	535	24	Rp1,306,800
22	C110	12-Sep-17	485	106	Rp864,340
23	C111	8-Sep-17	489	690	Rp3,575,240
24	C112	4-Apr-18	281	246	Rp864,000
25	C114	30-Dec-17	376	1742	Rp18,651,200
26	C118	27-May-17	593	514	Rp1,530,980
27	C121	18-Feb-18	326	552	Rp1,759,240
28	C122	28-Apr-17	622	46	Rp174,860
29	C123	2-Mar-18	314	552	Rp1,968,500
30	C125	2-Mar-18	314	108	Rp397,480
31	C129	12-Feb-17	697	38	Rp195,380
32	C131	18-Jan-17	722	468	Rp796,060
33	C133	30-May-17	590	680	Rp8,175,400
Rata - Rata			394	373	Rp3,515,116

Cluster 3 memiliki 62 anggota, jika dilihat pada tabel VI.4 cluster 3 memiliki nilai rata-rata recency pendek yaitu 93 hari atau 3 bulan 1 hari hal ini memnadakn bahwa pelanggan pada cluster 3 melakuakn pembelian belum lama, namu memiliki frekuensi yang rendah yaitu 276 tarnsaksi dan monetary Rp2,736,874 . Berdasarkan knowledge yang diperoleh dari data transaski Pulsantara, cluster 3 ini merupakan customer yang baru yang kemungkinan akan berpotensil memberikan keuntungan bagi Pulsantara, hal ini dikarenakan customer pada cluster 3 memiliki nilai recency yang pendek, rata-rata monetary yang sangat sedikit, selain itu memiliki frekuensi pembelian yang sangat rendah. Cluster ini dapat dikategorikan kedalam jenis Dormant Customer

Tabel IV.4 Hasil Cluster 3

No	Nama Agen	Tanggal terahir Pembelian	R	F	M
1	C004	31-Dec-18	10	72	Rp1,052,500
2	C005	31-Dec-18	10	154	Rp1,405,100
3	C006	31-Dec-18	10	60	Rp627,100
4	C008	31-Dec-18	10	216	Rp1,909,500
5	C009	27-Dec-18	14	30	Rp251,800
6	C010	31-Dec-18	10	468	Rp6,218,446
7	C012	29-Dec-18	12	180	Rp2,266,850
8	C013	12-Aug-18	151	302	Rp3,905,680
9	C016	31-Dec-18	10	274	Rp5,333,312
10	C018	31-Dec-18	10	100	Rp2,825,542
11	C019	29-Dec-18	12	24	Rp175,600
12	C024	16-Dec-18	25	98	Rp1,598,112
13	C027	30-Dec-18	11	460	Rp6,742,282
14	C029	20-Dec-18	21	336	Rp6,784,610
15	C030	29-Dec-18	12	404	Rp3,630,106
16	C031	31-Dec-18	10	204	Rp3,104,520
17	C033	30-Dec-18	11	732	Rp7,845,624
18	C034	31-Dec-18	10	148	Rp1,097,870
19	C035	28-Dec-18	13	620	Rp7,880,618
20	C039	26-Dec-18	15	612	Rp4,992,860
21	C040	12-Jul-18	182	230	Rp2,976,424
22	C044	30-Dec-18	11	104	Rp1,287,040
23	C046	31-Dec-18	10	54	Rp593,990
24	C047	19-Dec-18	22	260	Rp3,374,680
25	C049	16-Dec-18	25	152	Rp2,699,200
26	C051	30-Sep-18	102	756	Rp8,898,800

27	C055	11-Jun-18	213	880	Rp8,839,204
28	C059	24-Oct-18	78	454	Rp5,515,130
29	C062	10-Oct-18	92	38	Rp417,720
30	C063	22-Oct-18	80	838	Rp9,695,030
31	C065	31-Jul-18	163	314	Rp3,476,588
32	C066	31-Jul-18	163	12	Rp97,980
33	C068	14-Sep-18	118	80	Rp1,160,860
34	C069	9-Sep-18	123	16	Rp122,600
35	C071	7-Sep-18	125	472	Rp2,989,200
36	C072	22-Jul-18	172	30	Rp295,480
37	C074	13-Jul-18	181	10	Rp106,590
38	C076	21-Aug-18	142	358	Rp5,297,740
39	C077	27-Sep-18	105	140	Rp2,896,250
40	C078	7-Aug-18	156	1324	Rp8,773,840
41	C079	19-Aug-18	144	364	Rp5,914,450
42	C080	16-Jun-18	208	24	Rp399,450
43	C081	23-Jun-18	201	54	Rp925,650
44	C083	29-Jun-18	195	24	Rp599,858
45	C084	28-Jun-18	196	56	Rp786,840
46	C085	23-May-18	232	308	Rp3,904,800
47	C088	31-Dec-18	10	140	Rp1,301,100
48	C089	31-May-18	224	16	Rp168,950
49	C090	5-Sep-18	127	392	Rp4,042,300
50	C097	2-Aug-18	161	38	Rp389,500
51	C098	2-Jul-18	192	12	Rp90,200
52	C108	6-Jun-18	218	690	Rp2,547,200
53	C113	4-Jul-18	190	72	Rp130,860
54	C115	27-Aug-18	136	468	Rp1,619,300
55	C117	24-Oct-18	78	464	Rp1,973,300
56	C119	18-Nov-18	53	108	Rp654,920
57	C120	12-Dec-18	29	24	Rp174,820
58	C124	31-Dec-18	10	546	Rp1,750,400
59	C126	8-Jul-18	186	486	Rp1,175,860
60	C127	5-Dec-18	36	196	Rp596,700
61	C128	9-Jun-18	215	582	Rp1,181,300
62	C130	3-Sep-18	129	50	Rp196,060
Rata - Rata			93.871	276.29	Rp2,736,874

Cluster 4 memiliki 11 anggota. Cluster 4 ini dapat dikatakan sebagai cluster outlier, hal ini dapat dilihat dari data yang dimiliki customer SG, SG memiliki nilai recency , frekuensi dan monetary paling tinggi dibandingkan cluster yang lainnya. Nilai recency yang dimiliki pelanggan pada cluster 4 sebesar 46 , karena semakin kecil nilai recency nya maka semakin tinggi nilai transaksinya. Nilai 46 ini menandakan bahwa transaksi terakhir pelanngan cluster 4 adalah 1 bulan terakhir. Frekuensi pada cluster 4 melakukan transaksi pembelian produk ke Pulsantara sebanyak 9700 kali Sehingga cluster 4 dapat dikatakan memiliki loyalitas yang tinggi terhadap perusahan Pulsantara. SG memiliki rata-rata nilai monitry yang sangat tinggi sebesar Rp140,270,302. Nilai monetary tersebut sangat besar jika di bandingkan dengan nilai monetary cluster lain. Berdasarkan identifikasi knowledge pada Pulsantara, Cluster 4 merupakan customer yang sangat berpotensial dengan frekuensi pembelian yang tinggi dan besarnya nilai monetary yang dikeluarkan SG dalam transaksi akan memberikan keuntungan yang besar bagi perusahan pulsantara. Oleh karena itu cluster 4 dapat dikategorikan kedalam jenis customer Superstar, yang memiliki nilai recency , frekuensi dan monetary yang tinggi. Customer jenis ini juga merupakan customer yang memiliki nilai loyalti yang besar terhadap perusahaan. Oleh karena itu perusahan pulsantara harus mempertahankan customer Cluster 4. Tabel IV.5 dibawah ini menunjukan hasil output clustering cluster 4.

Tabel IV.5 Hasil Cluster 4

No	Nama Agen	Tanggal terahir Pembelian	R	F	M
1	C001	31-Dec-18	10	4070	Rp45,819,008
2	C017	31-Dec-18	10	21786	Rp37,633,420
3	C023	31-Dec-18	10	4904	Rp85,650,444
4	C025	31-Dec-18	10	2848	Rp39,105,146
5	C036	31-Dec-18	10	3982	Rp43,387,626
6	C041	31-Dec-18	10	9982	Rp184,124,460
7	C043	31-Dec-18	10	2702	Rp43,806,630
8	C053	11-Apr-18	274	26688	Rp514,077,400
9	C056	30-Nov-18	41	3418	Rp43,361,958
10	C073	30-Sep-18	102	24808	Rp459,987,490
11	C132	16-Dec-18	25	1522	Rp46,019,740
Rata - Rata			46.5455	9700.91	Rp140,270,302

BAB V

KESIMPULAN

5.1 Kesimpulan

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut :

1. Berdasarkan hasil *clustering* dengan metode *K-Means* dan *K-Medoids* jumlah segmen pelanggan yang terbentuk adalah 4 *cluster*/segmen pelanggan dari keseluruhan jumlah transaksi 85010 transaksi dengan jumlah agen sebanyak 133.
2. Algorithma terbaik antara K-Means dan Algorithma K-Medoids untuk segmentasi pelanggan dengan jumlah cluster 4 adalah algorithma K-Means dikarenakan nilai DB Index K-means yaitu 0.816% lebih kecil dibandingkan menggunakan Algorithma K-Medoids yaitu 0.820%.
3. *Cluster* yang terbentuk menggunakan algoritma *K-Means* antara lain: 27 anggota berada pada *cluster 1*, 33 anggota berada pada *cluster 2*, 62 anggota berada pada *cluster 3* dan 11 anggota berada pada *cluster 4*.
4. *Cluster* yang terbentuk menggunakan algoritma *K-Medoids* antara lain: 17 anggota berada pada *cluster 1*, 80 anggota berada pada *cluster 2*, 14 anggota berada pada *cluster 3* dan 22 anggota berada pada *cluster 4*.
5. Cluster 1 yang memiliki 27 anggota . Nilai rata-rata recency pada cluster 1 adalah 59 hari atau 1 bulan lebih 19 hari. Rata-rata frekuensi pembelian yang dilakukan customer cluster 1 dapat dikatakan berada pada nilai frekuensi yang rendah, yaitu hanya 1532 kali transaksi. Sedangkan untuk rata-rata nilai monetary yang dikeluarkan pelanggan sebesar Rp17,614,914, nilai tersebut menjadi nilai monetary kedua terendah dari keempat cluster. Yang mempengaruhi cluster 1 adalah frekuensi dan monetary. Berdasarkan identifikasi knowledge yang diperoleh, cluster 1 termasuk kedalam jenis Typical Customer, jenis customer ini memiliki nilai recency , frekuensi dan monetary rata-rata, sehingga jenis customer seperti ini akan memberikan sedikit keuntungan bagi perusahaan.
6. Cluster 2 memiliki 33 anggota, cluster 2 memiliki nilai rata-rata recency jauh yaitu 394 hari atau 1 tahun 1 bulan 4 hari , frekuensi yang rendah yaitu 373 tarnsaksi dan monetary Rp3,515,116. Berdasarkan knowledge yang diperoleh dari data transaski Pulsantara, cluster 2 ini merupakan

customer yang kurang berpotensil memberikan keuntungan bagi Pulsantra, hal ini dikarenakan customer pada cluster 2 memiliki nilai rata-rata monetary yang sedikit sehingga kurang memberikan keuntungan bagi Pulsantara , selain itu juga dikarenakan customer pada cluster ini melakukan transaksi terakhir pada waktu yang cukup lama. Cluster ini dapat dikategorikan kedalam jenis Occational customer.

7. Cluster 3 memiliki 62 anggota, cluster 3 memiliki nilai rata-rata recency pendek yaitu 93 hari atau 3 bulan 1 hari hal ini memnadakn bahwa pelanggan pada cluster 3 melakuakn pembelian belum lama, namu memiliki frekuensi yang rendah yaitu 276 tarnsaksi dan monetary Rp2,736,874 . Berdasarkan knowledge yang diperoleh dari data transaski Pulsantara, cluster 3 ini merupakan customer yang baru yang kemungkinan akan berpotensil memberikan keuntungan bagi Pulsantara, hal ini dikarenakan customer pada cluster 3 memiliki nilai recency yang pendek, rata-rata monetary yang sangat sedikit, selain itu memiliki frekuensi pembelian yang sangat rendah. Cluster ini dapat dikategorikan kedalam jenis Dormant Customer
8. Cluster 4 memiliki 11 anggota. Cluster 4 ini dapat dikatakan sebagai cluster outlier, hal ini dapat dilihat dari data yang dimiliki customer SG, SG memiliki nilai recency , frekuensi dan monetary paling tinggi dibandingkan cluster yang lainnya. Nilai recency yang dimiliki pelanggan pada cluster 4 sebesar 46 , karena semakin kecil nilai recency nya maka semakin tinggi nilai transaksinya

5.2 Saran

Saran diberikan agar penelitian ini dapat berkembang untuk penelitian-penelitian selanjutnya, berikut saran-saran yang diusulkan:

1. Penelitian ini dapat dikembangkan dengan menggunakan metode-metode *clustering* lainnya untuk melakukan perbandingan.
2. Variasi tentang rentang waktu transaksi untuk data yang akan di mining
3. Untuk penentuan cluster optimal dapat digunakan pendekatan lainnya tidak hanya *Davies Bouldien Index*

DAFTAR PUSTAKA

Adiana, B. E., Soesanti, I., & Permanasari, A. E. (2018). Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering. *JUTEI*, *1*(2), 23–32. https://doi.org/10.21460/jutei.2017.21.76

Aryuni, M., & Miranda, E. (2018). Customer Segmentation in XYZ Bank using K-Means and K-Medoids Clustering. In *2018 International Conference on Information Management and Technology (ICIMTech)* (pp. 412–416). IEEE.

Asiabi, T. T. P. K., & RezaTavoli. (2015). A Review of Different Data Mining Techniques in Customer Segmentation. *Journal of Advances in Computer Research Quarterly*, *6*(3), 51–64.

Butarbutar, N., Windarto, A. P., Hartama, D., & Solikhun. (2016). Komparasi Kinerja Algoritma Fuzzy C-Means Dan K-Means Dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa. *JURASIK (Jurnal Riset Sistem Informasi & Teknik Informatika)*, *1*(1), 46–55.

Çaliş, A., & Boyaci, A. (2015). Data mining application in banking sector with clustering and classification methods. In *Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management Dubai, United Arab Emirates (UAE)* (pp. 978–985).

Febrianti, F., Hafiyusholeh, M., & Asyhar, A. H. (2016). PERBANDINGAN PENGKLUSTERAN DATA IRIS MENGGUNAKAN METODE K-MEANS DAN FUZZY C-. *JURNAL MATEMATIKA "MANTIK," 02*(01), 7–13.

Golmah, V. (2014). A Case Study of Applying SOM in Market Segmentation of Automobile Insurance Customers. *International Journal of Database Theory and Application*, *7*(1), 25–36.

Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques*. Berlin: Springer.

Hamdi, K., & Zamiri, A. (2016). Identifying and Segmenting Customers of Pasargad Insurance Company Through RFM Model (RFM). *Internasional Bussness Managemen*, *10*(18), 4209–4214.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques Third Edition*. USA: Elsevier Inc.

Hand, D. J. (2007). Principles of Data Mining. In *Drug Safety* (pp. 621–622).

Hill, N., & Alexander, J. (2006). *The Handbook of Customer Satisfaction and Loyalty Measurement* (3nd Revise). New York: Gower Publishing Ltd;

I Gede Wiyana Ananta Noor, Ariyanti, M., & Alamsyah, A. (2019). Telecom Customer ' s Segmentation Using Decision Tree to Increase Active Electronic Money Subscribers. In *1st International Conference on*

Economics, Business, Entrepreneurship, and Finance (ICEBEF 2018) Telecom (Vol. 65, pp. 628–632).

Joyendri, A. (2017). STRATEGI CUSTOMER RELATIONSHIP MANAGEMENT UNTUK MENINGKATKAN LOYALITAS PELANGGAN DAN VOLUME PENJUALAN MENGGUNAKAN TEKNIK CLUSTERING K-MEANS. *TELEMATIKA*, *14*(02), 75–82.

Kashwan, K. R., & Velu, C. M. (2013). Customer Segmentation Using Clustering and Data Mining Techniques. *International Journal of Computer Theory and Engineering*, *5*(6), 1–6. https://doi.org/10.7763/IJCTE.2013.V5.811

Kothari, C. R. (2004). *Research Methodology Methods and Techniques*. India: New Age International Limited.

Kurniawan, P. S. (2015). Perancangan Data Mining untuk Analisis Kriteria Nasabah Kredit yang Potensial dan Manfaatnya untuk Customer Relationship Management Perbankan. *Journal of Accounting and Investment*, *1*(11), 155–174. https://doi.org/10.18196/JAI-2015.0040

Kusrini Luthfi, E. T. (2009). *Algoritma Data mining*. (T. A. Prabawati, Ed.). Yogyakarta: C.V Andi OFFSET. Retrieved from https://books.google.co.id/books?id=-Ojclag73O8C&pg=PA3&dq=data+mining+adalah&hl=id&sa=X&ved=0ahUKEwijrefgpYnZAhXBPY8KHWeJCQ4Q6AEIKzAA#v=onepage&q=data mining adalah&f=false

Lee, Y. J. (2014). Privacy-preserving Data Mining for Personalized Marketing. *International Journal of Computer Communications and Networks (IJCCN)*, *4*(1), 1–9.

Maimon, O., & Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook Second Edition*. London: Springer.

Maryani, I., & Riana, D. (2017). Clustering and profiling of customers using RFM for customer relationship management recommendations. In *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017* (pp. 2–7). https://doi.org/10.1109/CITSM.2017.8089258

Maryani, I., Riana, D., Astuti, R. D., Ishaq, A., & Pratama, E. A. (2018). Customer Segmentation based on RFM model and Clustering Techniques With K-Means Algorithm. In *2018 Third International Conference on Informatics and Computing (ICIC)* (pp. 1–6). IEEE.

Maulik, U., & Bandyopadhyay, S. (2002). Performance Evaluation of Some Clustering Algorithms and Validity Indices. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, *24*(12), 1650–1654.

Monalisa, S. (2018). Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-Means. *QUERY: Jurnal Sistem Informasi*,

2(1), 9–15.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Expert Systems with Applications Application of data mining techniques in customer relationship management : A literature review and classification. *Expert Systems With Applications*, *36*(2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Parlina, I., Windarto, A. P., Wanto, A., & M.Ridwan Lubis. (2018). MEMANFAATKAN ALGORITMA K-MEANS DALAM MENENTUKAN PEGAWAI YANG LAYAK MENGIKUTI ASESSMENT CENTER UNTUK CLUSTERING PROGRAM SDP. *CESS (Journal of Computer Engineering System and Science)*, *3*(1), 87–93.

Peiman Alipour Sarvari, Ustundag, A., & Takci, H. (2014). Performance Evaluation of Different Customer Segmentation Approaches Based on RFM and Demographics Analysis. *Kybernetes*, *43*(8), 1209–1223. https://doi.org/10.1108/K-01-2015-0009

Prakasawati, P. E., Chrisnanto, Y. H., & Hadiana, A. I. (2019). SEGMENTASI PELANGGAN BERDASARKAN PRODUK MENGGUNAKAN METODE K- MEDOIDS. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, *3*(1), 335–339. https://doi.org/10.30865/komik.v3i1.1610

Raje, P. A., & Srivastava, R. K. (2014). The Connection between Data Mining and Segmentation in Marketing Area. *International Journal of Innovative Research in Computer and and Communication Engineering*, *2*(12), 7415–7418.

Ramadhana, C., W, Y. D. L., & W, K. D. K. (2013). Data Mining dengan Algoritma Fuzzy C-Means Clustering Dalam Kasus Penjualan di PT Sepatu Bata. *Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2013 (Semantik 2013)*, *2013*(November), 54–60.

Ramamohan, Y., Vasantharao, K., Chakravarti, C. K., & Ratnam, a S. K. (2012). A Study of Data Mining Tools in Knowledge Discovery Process. *International Journal of Soft Computing and Engineering*, *2*(3), 191–194.

Roselin, R. (2014). Customer Behaviour Analysis for Credit Card Proposers Based on Data Mining Techniques. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, *1*(11), 62–66.

Sadewo, M. G., Windarto, A. P., & Hartama, D. (2017). PENERAPAN DATAMINING PADA POPULASI DAGING AYAM RAS PEDAGING DI INDONESIA BERDASARKAN PROVINSI MENGGUNAKAN K-MEANS CLUSTERING. *Jurnal Nasional Informatika Dan Teknologi Jaringan*, *2*(1), 60–67.

Shihab, S. H., Afroge, S., & Mishu, S. Z. (2019). RFM Based Market Segmentation Approach Using Advanced K-means and Agglomerative Clustering: A Comparative Study. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 7–9).

IEEE.

Siyoto, S., & Sodik, M. A. (2015). *Dasar Metodologi Penelitia*. Yogyakarta: Literasi Media Publishing.

Supriatna, A., & Budianto, H. (2019). Penerapan Customer Relationship Management (CRM) Sebagai Upaya Meningkatkan Kepuasan Dan Loyalitas Pelanggan Pada Gelora Mukti Sport Berbasis Web. *Jurnal Nuansa Informatika*, *13*(2), 11–18.

Tsiptsis, K., & Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation*. (L. John Wiley & Sons, Ed.) (firs). United Kingdom: John Wiley & Sons, Ltd.

Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., & Kim, S. W. (2019). A Churn Prediction Model Using Random Forest : Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector. *IEEE Access*, *7*(60), 60134–60149. https://doi.org/10.1109/ACCESS.2019.2914999

Utami, C. W. (2010). *Manajemen Ritel : Strategi dan Implementasi OperasionalBisnis Ritel Moderen DI Indonesia* (second). Jakarta: Salemba Empat.

Witten, I. H., Frank, E., Hall, M. A., & Pall, C. J. (2017). *Data Mining Partical Machine Learning Tools and Tecniques* (Fourth). United States: Elsevier Inc.

Wong, E., & Wei, Y. (2018). Customer online shopping experience data analytics Integrated customer segmentation and customised services prediction model. *International Journal of Retail & Distribution Management*, *46*(4), 406–420. https://doi.org/10.1108/IJRDM-06-2017-0130

Wongchinsri, P., & Kuratach, W. (2016). A survey - Data mining frameworks in credit card processing. In *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2016* (pp. 1–6). https://doi.org/10.1109/ECTICon.2016.7561287

Yi, B., Yang, F., Qiao, H., & Xu, C. (2010). An Improved Initialization Center Algorithm for K-means Clustering. In *International Conference on Computational Intelligence and Software Engineering* (pp. 1–4).

Yuliari, N. P. P., Putra, I. K. G. D., & Rusjayanti, N. K. D. (2015). CUSTOMER SEGMENTATION THROUGH FUZZY C-MEANS AND FUZZY RFM METHOD. *Journal of Theoretical and Applied Information Technology*, *78*(3), 380–385.

Zayuka, H., Nasution, S. M., & Purwanto, Y. (2017). PERANCANGAN DAN ANALISIS CLUSTERING DATA MENGGUNAKAN METODE K-MEDOIDS UNTUK BERITA BERBAHASA INGGRIS DESIGN AND ANALYSIS OF DATA CLUSTERING USING K-MEDOIDS METHOD. In

e-Proceeding of Engineering : (Vol. 4, pp. 2182–2190).

LAMPIRAN

1. **Surat Riset**



DISTRIBUTOR SERVER PULSA ELEKTRONIK
PULSANTARA PURBALINGGA

Jl. Selakambang No. 32 Rt 02 Rw 04 Penaruban, Kel. Kaligondang Kab. Purbalingga, 53391

SURAT KETERANGAN
NO. 027/S.Ket/IV/2019

Yang bertanda tangan dibawah ini :

Nama : Bayu Setiawan

Jabatan : Pemilik Server Pulsantara

Dengan ini menerangkan bahwa :

Nama : Rachmawati Darma Astuti

NIM : 14002140

Kampus : STMIK Nusa Mandiri Jakarta

Adalah benar telah melaksanakan riset atau penelitian pada Pulsantara Purwokerto dari tanggal 05 Juni 2019 s/d 05 Juli 2019.

Demikian Surat keterangan ini dibuat untuk dapat dipergunakan sebagimana mestinya.

Purbalingga 06 Juni 2019

Pemilik

Pulsantra

Bayu Setiawan

Lampiran – Data yang sudah dipilih atributnyamenggunakan RFM

Nama Agen	R		F	M
C001	31-Dec-18	10	4070	Rp45,819,008
C002	18-Jun-18	206	678	Rp11,756,612
C003	11-Apr-18	274	704	Rp9,050,500
C004	31-Dec-18	10	72	Rp1,052,500
C005	31-Dec-18	10	154	Rp1,405,100
C006	31-Dec-18	10	60	Rp627,100
C007	31-Dec-18	10	858	Rp10,802,164
C008	31-Dec-18	10	216	Rp1,909,500
C009	27-Dec-18	14	30	Rp251,800
C010	31-Dec-18	10	468	Rp6,218,446
C011	30-Dec-18	11	634	Rp10,568,674
C012	29-Dec-18	12	180	Rp2,266,850
C013	12-Aug-18	151	302	Rp3,905,680
C014	31-Dec-18	10	1984	Rp20,234,624
C015	31-Dec-18	10	1092	Rp13,753,840
C016	31-Dec-18	10	274	Rp5,333,312
C017	31-Dec-18	10	21786	Rp37,633,420
C018	31-Dec-18	10	100	Rp2,825,542
C019	29-Dec-18	12	24	Rp175,600
C020	31-Dec-18	10	864	Rp13,966,782
C021	21-Dec-18	20	1384	Rp19,043,260
C022	31-Dec-18	10	1196	Rp17,561,530
C023	31-Dec-18	10	4904	Rp85,650,444
C024	16-Dec-18	25	98	Rp1,598,112
C025	31-Dec-18	10	2848	Rp39,105,146
C026	31-Dec-18	10	2770	Rp28,604,980
C027	30-Dec-18	11	460	Rp6,742,282
C028	31-Dec-18	10	2356	Rp26,896,754
C029	20-Dec-18	21	336	Rp6,784,610
C030	29-Dec-18	12	404	Rp3,630,106
C031	31-Dec-18	10	204	Rp3,104,520
C032	31-Dec-18	10	1064	Rp8,522,220
C033	30-Dec-18	11	732	Rp7,845,624
C034	31-Dec-18	10	148	Rp1,097,870
C035	28-Dec-18	13	620	Rp7,880,618
C036	31-Dec-18	10	3982	Rp43,387,626
C037	30-Dec-18	11	1086	Rp11,672,432
C038	31-Dec-18	10	2024	Rp22,007,962
C039	26-Dec-18	15	612	Rp4,992,860
C040	12-Jul-18	182	230	Rp2,976,424

C041	31-Dec-18	10	9982	Rp184,124,460
C042	31-Dec-18	10	1332	Rp18,981,256
C043	31-Dec-18	10	2702	Rp43,806,630
C044	30-Dec-18	11	104	Rp1,287,040
C045	31-Dec-18	10	1492	Rp29,999,754
C046	31-Dec-18	10	54	Rp593,990
C047	19-Dec-18	22	260	Rp3,374,680
C048	12-Apr-18	273	1330	Rp15,337,972
C049	16-Dec-18	25	152	Rp2,699,200
C050	31-Dec-18	10	1880	Rp19,399,862
C051	30-Sep-18	102	756	Rp8,898,800
C052	11-Jan-18	364	1384	Rp17,263,320
C053	11-Apr-18	274	26688	Rp514,077,400
C054	11-Nov-18	60	2436	Rp25,860,596
C055	11-Jun-18	213	880	Rp8,839,204
C056	30-Nov-18	41	3418	Rp43,361,958
C057	10-Jan-18	365	776	Rp7,997,548
C058	10-Apr-18	275	836	Rp8,456,038
C059	24-Oct-18	78	454	Rp5,515,130
C060	29-Oct-18	73	1050	Rp12,880,018
C061	10-May-18	245	82	Rp701,594
C062	10-Oct-18	92	38	Rp417,720
C063	22-Oct-18	80	838	Rp9,695,030
C064	9-Dec-18	32	1030	Rp11,459,380
C065	31-Jul-18	163	314	Rp3,476,588
C066	31-Jul-18	163	12	Rp97,980
C067	15-Jul-18	179	1480	Rp13,407,910
C068	14-Sep-18	118	80	Rp1,160,860
C069	9-Sep-18	123	16	Rp122,600
C070	30-Aug-18	133	1352	Rp11,772,114
C071	7-Sep-18	125	472	Rp2,989,200
C072	22-Jul-18	172	30	Rp295,480
C073	30-Sep-18	102	24808	Rp459,987,490
C074	13-Jul-18	181	10	Rp106,590
C075	15-Jul-18	179	2482	Rp24,691,970
C076	21-Aug-18	142	358	Rp5,297,740
C077	27-Sep-18	105	140	Rp2,896,250
C078	7-Aug-18	156	1324	Rp8,773,840
C079	19-Aug-18	144	364	Rp5,914,450
C080	16-Jun-18	208	24	Rp399,450
C081	23-Jun-18	201	54	Rp925,650
C082	19-Jun-18	205	2936	Rp23,598,090
C083	29-Jun-18	195	24	Rp599,858
C084	28-Jun-18	196	56	Rp786,840
C085	23-May-18	232	308	Rp3,904,800

C086	13-May-18	242	56	Rp508,150
C087	5-May-18	250	214	Rp1,851,150
C088	31-Dec-18	10	140	Rp1,301,100
C089	31-May-18	224	16	Rp168,950
C090	5-Sep-18	127	392	Rp4,042,300
C091	14-May-18	241	324	Rp2,593,716
C092	26-Apr-18	259	210	Rp4,052,960
C093	30-Apr-18	255	176	Rp4,146,300
C094	22-Apr-18	263	244	Rp2,597,300
C095	3-Feb-18	341	58	Rp601,950
C096	15-Mar-18	301	26	Rp196,500
C097	2-Aug-18	161	38	Rp389,500
C098	2-Jul-18	192	12	Rp90,200
C099	13-Jan-18	362	66	Rp689,100
C100	26-Jan-18	349	12	Rp395,200
C101	19-Jan-18	356	60	Rp744,500
C102	25-Jan-18	350	56	Rp462,750
C103	3-Mar-18	313	68	Rp1,907,260
C104	11-Dec-18	30	1508	Rp15,270,842
C105	12-Mar-17	669	708	Rp7,085,960
C106	12-May-17	608	486	Rp4,437,538
C107	31-Dec-18	10	1752	Rp27,453,092
C108	6-Jun-18	218	690	Rp2,547,200
C109	24-Jul-17	535	24	Rp1,306,800
C110	12-Sep-17	485	106	Rp864,340
C111	8-Sep-17	489	690	Rp3,575,240
C112	4-Apr-18	281	246	Rp864,000
C113	4-Jul-18	190	72	Rp130,860
C114	30-Dec-17	376	1742	Rp18,651,200
C115	27-Aug-18	136	468	Rp1,619,300
C116	10-Nov-18	61	1330	Rp10,098,000
C117	24-Oct-18	78	464	Rp1,973,300
C118	27-May-17	593	514	Rp1,530,980
C119	18-Nov-18	53	108	Rp654,920
C120	12-Dec-18	29	24	Rp174,820
C121	18-Feb-18	326	552	Rp1,759,240
C122	28-Apr-17	622	46	Rp174,860
C123	2-Mar-18	314	552	Rp1,968,500
C124	31-Dec-18	10	546	Rp1,750,400
C125	2-Mar-18	314	108	Rp397,480
C126	8-Jul-18	186	486	Rp1,175,860
C127	5-Dec-18	36	196	Rp596,700
C128	9-Jun-18	215	582	Rp1,181,300
C129	12-Feb-17	697	38	Rp195,380
C130	3-Sep-18	129	50	Rp196,060

C131	18-Jan-17	722	468	Rp796,060
C132	16-Dec-18	25	1522	Rp46,019,740
C133	30-May-17	590	680	Rp8,175,400

Lampiran – Data yang sudah diolah menggunakan metode RFM

KODE CUSTOMER	R	F	M
C001	5	3	5
C002	4	1	3
C003	3	1	2
C004	5	1	1
C005	5	1	1
C006	5	1	1
C007	5	1	3
C008	5	1	1
C009	5	1	1
C010	5	1	2
C011	5	1	3
C012	5	1	1
C013	4	1	1
C014	5	2	4
C015	5	2	3
C016	5	1	2
C017	5	5	5
C018	5	1	1
C019	5	1	1
C020	5	1	3
C021	5	2	3
C022	5	2	3
C023	5	3	5
C024	5	1	1
C025	5	2	5
C026	5	2	4
C027	5	1	2
C028	5	2	4
C029	5	1	2
C030	5	1	1
C031	5	1	1
C032	5	2	2
C033	5	1	2
C034	5	1	1
C035	5	1	2
C036	5	3	5
C037	5	2	3
C038	5	2	4
C039	5	1	1
C040	4	1	1
C041	5	5	5

C042	5	2	3
C043	5	2	5
C044	5	1	1
C045	5	2	4
C046	5	1	1
C047	5	1	1
C048	3	2	3
C049	5	1	1
C050	5	2	3
C051	4	1	2
C052	2	2	3
C053	3	5	5
C054	4	2	4
C055	4	1	2
C056	5	3	5
C057	2	1	2
C058	3	1	2
C059	4	1	2
C060	4	2	3
C061	3	1	1
C062	4	1	1
C063	4	1	2
C064	5	2	3
C065	4	1	1
C066	4	1	1
C067	4	2	3
C068	4	1	1
C069	4	1	1
C070	4	2	3
C071	4	1	1
C072	4	1	1
C073	4	5	5
C074	4	1	1
C075	4	2	4
C076	4	1	2
C077	4	1	1
C078	4	2	2
C079	4	1	2
C080	4	1	1
C081	4	1	1
C082	4	2	4
C083	4	1	1
C084	4	1	1
C085	4	1	1
C086	3	1	1

C087	3	1	1
C088	5	1	1
C089	4	1	1
C090	4	1	1
C091	3	1	1
C092	3	1	1
C093	3	1	1
C094	3	1	1
C095	3	1	1
C096	3	1	1
C097	4	1	1
C098	4	1	1
C099	2	1	1
C100	3	1	1
C101	3	1	1
C102	3	1	1
C103	3	1	1
C104	5	2	3
C105	1	1	2
C106	1	1	1
C107	5	2	4
C108	4	1	1
C109	1	1	1
C110	1	1	1
C111	1	1	1
C112	3	1	1
C113	4	1	1
C114	2	2	3
C115	4	1	1
C116	4	2	3
C117	4	1	1
C118	1	1	1
C119	5	1	1
C120	5	1	1
C121	3	1	1
C122	1	1	1
C123	3	1	1
C124	5	1	1
C125	3	1	1
C126	4	1	1
C127	5	1	1
C128	4	1	1
C129	1	1	1
C130	4	1	1
C131	1	1	1

C132	5	2	5
C133	1	1	2

STMIK NUSA MANDIRI	**Lembar Konsultasi Bimbingan Tesis** **Program Studi Ilmu Komputer (S2)** **STMIK Nusa Mandiri Jakarta**

NIM : 14002140
Nama Lengkap : Rachmawati Darma Astuti
Dosen Pembimbing : Dr. Dwiza Riana S.Si, MM, M.Kom
Judul Tesis : Analisis Perbandingan Algoritma K-Means Dan K-Medoids Untuk Menerapkan Segmentasi Pelanggan

No	Tanggal Bimbingan	Materi Bimbingan	Paraf Dosen Pembimbing
1.	30 September 2019	Pengajuan Proposal Tesis	
2.	08 Oktober 2019	Pengajuan BAB I	
3.	20 Oktober 2019	ACC Pengajuan BAB I dan Pengajuan BAB II	
4.	30 Oktober 2019	ACC Pengajuan BAB II dan Pengajuan BAB III	
5.	12 November 2019	ACC Pengajuan BAB III dan Pengajuan BAB IV	
6.	04 Desember 2019	Pengujian dan Pembahasan BAB V	
7.	17 Desember 2019	ACC Pengajuan BAB IV dan Pengajuan BAB V	
8.	08 Januari 2020	ACC Keseluruhan	

Catatan untuk Dosen Pembimbing
Bimbingan Tesis

- Dimulai pada tanggal : 30 September 2019
- Diakhiri pada tanggal : 08 Januari 2020
- Jumlah pertemuan bimbingan : 8 Kali

Disetujui Oleh,
Dosen Pembimbing

Dr. Dwiza Riana S.Si, MM, M.Kom

Dipindai dengan CamScanner

DAFTAR RIWAYAT HIDUP

A. Biodata mahasiswa

NIM	: 14002145
Nama Lengkap	: Rachmawati Darma Astuti
Tempat & Tanggal Lahir	: Jakarta, 27 Desember 1995
Alamat Lengkap	: Jl.Cerme RT 02/RW 05 Maoslor, Cilacap

B. Pendidikan

1. SD N 02 Maos lor, lulus tahun ajaran 2006 - 2007
2. SMP N 03 Maos, lulus tahun ajaran 2009 - 2010
3. SMK YPE Sampang, lulus tahun ajaran 2012 – 2013
4. Diploma III AMIK BSI Purwokerto, lulus tahun ajaran 2016
5. S1 STMIK Nusa Mandiri Jakarta, lulus tahun ajaran 2017
6. S2 SMIK Nusa Mandiri Jakarta, Sampai dengan sekarang

65

C. Riwayat Pengalaman Berorganisasi / Pekerjaan

1. Sebagai Seksi Bidang Majalah Dinding OSIS SMK YPE Sampang 2010/2011
2. Sebagai Anggota Bhyangkara Polsek Sampang 2010/2011
3. Sebagai Wakil Ketua OSIS SMK YPE Sampang 2011/2012
4. Sebagai Sekretaris I Bhyangkara Polsek Sampang 2011/2012
5. Sebagai Mentri Dalam Negri di SEMA AMIK BSI Purwokerto 2014-2015
6. Magang bagian Marketing Comunication AMIK BSI Purwokerto 2014 s/d 2016
7. Asisten Instruktur AMIK BSI Purwokerto 2017
8. Instruktur Program Studi Teknologi Komputer Universitas Bina Sarana Informtaika 2018 s/d serkarang



Jakarta, 08 Januari 2020

Rachmawati Darma Astuti