



**APLIKASI DATA MINING DENGAN K-MEANS CLUSTER UNTUK
MEMPREDIKSI PRODUK POTENSIAL DAN PENENTUAN
PERSEDIAAN PRODUK**

Disusun dan Diajukan untuk Memenuhi Persyaratan Ujian Akhir Memperoleh
Gelar Sarjana Komputer pada Fakultas Sains dan Teknologi
Universitas Pembangunan Panca Budi
Medan

SKRIPSI

OLEH

NAMA : DOLY ZAINUDDIN HASIBUAN
NPM : 1524370854
PROGRAM STUDI : SISTEM KOMPUTER

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN PANCA BUDI
MEDAN
2019**

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Doly Zainuddin Hasibuan
NPM : 1524370854
Prodi : Sistem Komputer
Konsentrasi : Kjk (Keamanan Jaringan Komputer)
Judul Skripsi : Aplikasi Data Mining Untuk Memprediksi Produk
Potensial dan Penentuan Persediaan Produk

Dengan ini menyatakan bahwa :

1. Tugas Akhir/Skripsi saya bukan hasil Plagiat
2. Saya tidak akan menuntut perbaikan nilai indeks Prestasi Kumulatif (IPK) setelah ujian Sidang Meja Hijau
3. Skripsi saya dapat dipublikasikan oleh pihak lembaga, dan saya tidak akan menuntut akibat publikasi tersebut

Demikian pernyataan ini saya perbuat dengan sebenar-benarnya, terimakasih

Medan, 30 Agustus 2019

METERAI TEMPEL nyataan
4D1C8ADF548990979
6000 ENAM RIBU RUPIAH

DOLY ZAINUDDIN HASIBUAN

1524370854



UNIVERSITAS PEMBANGUNAN PANCA BUDI
FAKULTAS ILMU KOMPUTER
 Jl. Jend. Gatot Subroto Km. 4,5 Telp (061) 8455571
 website : www.pancabudi.ac.id email: unpab@pancabudi.ac.id
 Medan - Indonesia

Universitas : Universitas Pembangunan Panca Budi
 Fakultas : ILMU KOMPUTER
 Dosen Pembimbing I : Kana Saputra S, S.Pd., M.kom
 Dosen Pembimbing II : Zulham Sitorus, S.Kom., M. Kom
 Nama Mahasiswa : DOLY ZAINUDDIN HASIBUAN
 Jurusan/Program Studi : Sistem Komputer
 Nomor Pokok Mahasiswa : 1524370854
 Jenjang Pendidikan : Strata Satu (S-1)
 Judul Tugas Akhir/Skripsi : Aplikasi Data Mining Dengan K-Means Cluster
 Untuk Memprediksi Produk Potensial Dan
 Penentuan Persediaan Produk.

TANGGAL	PEMBAHASAN MATERI	PARAF	KETERANGAN
18/9 - 18	Revisi jurnal ; jenis jenis produk.		
16/10 - 18	Aplikasi pada perencanaan di jenis produk hpl maka		
18/10 - 18	Rancangan di awal pada mulai ts kmi		
26/10 - 18	Literatur lebih jelas di awal.		
23/1 - 19	Acc. Jurnal hasil		

25/2 - 19. Acc. hasil uji coba
 19. Acc. hasil

Medan, 29 Januari 2018
 Diketahui/Disetujui oleh :
 Dekan,



Sri Shindi Indira, S.T.,M.Sc.



UNIVERSITAS PEMBANGUNAN PANCA BUDI
FAKULTAS ILMU KOMPUTER

Jl. Jend. Gatot Subroto Km. 4,5 Telp (061) 8455571
 website : www.pancabudi.ac.id email: unpab@pancabudi.ac.id
 Medan - Indonesia

Universitas : Universitas Pembangunan Panca Budi
 Fakultas : ILMU KOMPUTER
 Dosen Pembimbing I : Kana Saputra S, S.Pd., M.Kom / Dr. Muhammed Iqbal S.Kom., M.Kom
 Dosen Pembimbing II : Zulhan Situmorang, S. Kom., M.Kom
 Nama Mahasiswa : DOLY ZAINUDDIN HASIBUAN
 Jurusan/Program Studi : Sistem Komputer
 Nomor Pokok Mahasiswa : 1524370854
 Jenjang Pendidikan : Strata Satu (S-1)
 Judul Tugas Akhir/Skripsi : Aplikasi Data Mining Dengan K-Means Cluster Untuk Memprediksi Produk Potensial Dan Penentuan Perediaan Produk.

TANGGAL	PEMBAHASAN MATERI	PARAF	KETERANGAN
9/6 -17	~ perbaikan latar belakang dan tambah referensi	[Signature]	
21/8 -18	~ Acc BAB I	[Signature]	
5/9 -18	~ Acc BAB II	[Signature]	
18/9 -18	~ pelajaran babal variabel yg digunakan dan hapus apa!	[Signature]	
16/10 -18	~ Acc BAB III	[Signature]	
18/10 -18	~ Acc BAB IV	[Signature]	
26/10 -18	~ Acc BAB V dan sensor	[Signature]	
3/11 2018	- Acc Sudz	[Signature]	
17/11 2018	Acc Sudz	[Signature]	

Medan, 29 Januari 2018
 Diketahui/Disetujui oleh :
 Dekan,



Sri Shindi Indira, S.T., M.Sc.

Dinyatakan tidak ada sangkut paut dengan UPT: Perpustakaan
02 JUL 2019
UPT. Perpustakaan

Permohonan Meja Hijau



Medan, 25 Juni 2019
Kepada Yth : Bapak/Ibu Dekan
Fakultas SAINS & TEKNOLOGI
UNPAB Medan
Di -



Yang terhormat, saya yang bertanda tangan di bawah ini :

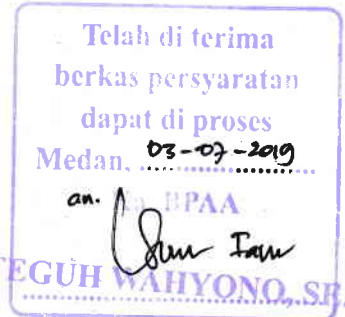
Nama : Doly Zainuddin Hasibuan
Tempat/Tgl. Lahir : MEDAN / 28 Maret 1993
Nama Orang Tua : Marwan Hasibuan
P. M : 1524370854
Kultur : SAINS & TEKNOLOGI
Program Studi : Sistem Komputer
No. HP : 081264551083
Alamat : Jl. Medan Batang Kuis Dusun II Gg. Ismail Sei Rotan

Sehubungan dengan ini saya bermohon kepada Bapak/Ibu untuk dapat diterima mengikuti Ujian Meja Hijau dengan judul Aplikasi data minning dengan k-means cluster untuk memprediksi produk potensial dan penentuan persediaan produk, Selanjutnya saya menyatakan :

- Melampirkan KKM yang telah disahkan oleh Ka. Prodi dan Dekan
- Tidak akan menuntut ujian perbaikan nilai mata kuliah untuk perbaikan indeks prestasi (IP), dan mohon diterbitkan ijazahnya setelah lulus ujian meja hijau.
- Telah tercap keterangan bebas pustaka
- Terlampir surat keterangan bebas laboratorium
- Terlampir pas photo untuk ijazah ukuran 4x6 = 5 lembar dan 3x4 = 5 lembar Hitam Putih
- Terlampir foto copy STTB SLTA dilegalisir 1 (satu) lembar dan bagi mahasiswa yang lanjutan D3 ke S1 lampirkan ijazah dan transkripnya sebanyak 1 lembar.
- Terlampir pelunasan kwintasi pembayaran uang kuliah berjalan dan wisuda sebanyak 1 lembar
- Skripsi sudah dijilid lux 2 exemplar (1 untuk perpustakaan, 1 untuk mahasiswa) dan jilid kertas jeruk 5 exemplar untuk penguji (bentuk dan warna penjilidan diserahkan berdasarkan ketentuan fakultas yang berlaku) dan lembar persetujuan sudah di tandatangani dosen pembimbing, prodi dan dekan
- Soft Copy Skripsi disimpan di CD sebanyak 2 disc (Sesuai dengan Judul Skripsinya)
- Terlampir surat keterangan BKKOL (pada saat pengambilan ijazah)
- Setelah menyelesaikan persyaratan point-point diatas berkas di masukan kedalam MAP
- Bersedia melunaskan biaya-biaya yang dibebankan untuk memproses pelaksanaan ujian dimaksud, dengan perincian sbb :

1. [102] Ujian Meja Hijau	: Rp.	750.000
2. [170] Administrasi Wisuda	: Rp.	1,500,000
3. [202] Bebas Pustaka	: Rp.	100,000
4. [221] Bebas LAB	: Rp.	5,000
Total Biaya	: Rp.	2.355.000

2.355.000 dlm 03/07/19



Ukuran Toga : M

Hormat saya
Doly Zainuddin Hasibuan
1524370854

Catatan :

- 1. Surat permohonan ini sah dan berlaku bila ;
 - a. Telah dicap Bukti Pelunasan dari UPT Perpustakaan UNPAB Medan.
 - b. Melampirkan Bukti Pembayaran Uang Kuliah aktif semester berjalan
- 2. Dibuat Rangkap 3 (tiga), untuk - Fakultas - untuk BPAA (asli) - Mhs.ybs.



YAYASAN PROF. DR. H. KADIRUN YAHYA
UNIVERSITAS PEMBANGUNAN PANCA BUDI
LABORATORIUM KOMPUTER
Jl. Jend. Gatot Subroto Km 4,5 Sei Sikambing Telp. 061-8455571
Medan - 20122

KARTU BEBAS PRAKTIKUM

Yang bertanda tangan dibawah ini Ka. Laboratorium Komputer dengan ini menerangkan bahwa :

Nama : Doly Zainuddin Hasibuan
N.P.M. : 1524370854
Tingkat/Semester : Akhir
Fakultas : SAINS & TEKNOLOGI
Jurusan/Prodi : Sistem Komputer

Benar dan telah menyelesaikan urusan administrasi di Laboratorium Komputer Universitas Pembangunan Panca Budi Medan.



Plagiarism Detector v. 1092 - Originality Report:

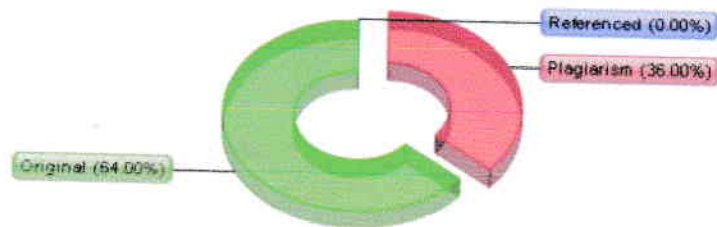
Analyzed document: 28-01-19 3:17:44 PM

"DOLY ZAINUDDIN
HASIBUAN_1524370854_SISTEM
KOMPUTER.docx"

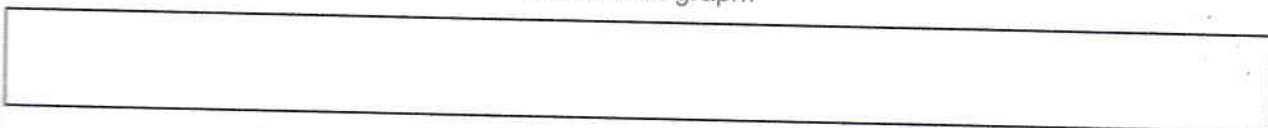
Licensed to: Universitas Pembangunan Panca Budi_License2



Relation chart:



Distribution graph:



Comparison Preset: Rewrite. Detected language: Indonesian

Top sources of plagiarism:

% 9	wrds: 826	https://rezqiwati.wordpress.com/2009/04/24/data-mining-%E2%80%93-proses-tahapan-dan-penera...
% 9	wrds: 826	https://rezqiwati.wordpress.com/2009/04/24/data-mining-
% 8	wrds: 735	http://xerma.blogspot.com/2015/01/pengertian-data-mining-apa-itu-data.html

[Show other Sources:]

Processed resources details:

176 - Ok / 53 - Failed

[Show other Sources:]

Important notes:

Wikipedia: Wiki Detected!	Google Books: [not detected]	Ghostwriting services: [not detected]	Anti-cheating: [not detected]
---	-------------------------------------	--	--------------------------------------



UNIVERSITAS PEMBANGUNAN PANCA BUDI
FAKULTAS ILMU KOMPUTER

Jl. Jend. Gatot Subroto Km 4,5 ☎ 061 50 200 502 PO BOX 1099 Medan

or : 2436 / 17 / F.Ilkom / 2017

p : 1 (satu) eks.

l : **Tugas Bimbingan Skripsi/Tugas Akhir**

ada : Yth. Bapak/Ibu

1. **Kana Saputra S, S.Pd.,M.Kom**

(Pembimbing I)

2. **Zulham sitorus, S.kom.,M.kom**

(Pembimbing II)

Di -

Tempat

Dengan hormat, sehubungan permohonan mahasiswa untuk melaksanakan pembuatan Skripsi/ Tugas Akhir, yang diajukan oleh :

N a m a : **Doly Zainuddin Hasibuan**

N.P.M : 1524370854

Prog. Studi : Sistem Komputer

Judul : Aplikasi Data Mining Dengan K-Means Cluster Untuk Memprediksi Produk Potensial Dan Penentuan Persediaan Produk

Jadwal Seminar :

Jadwal Sidang :

Sehubungan dengan hal tersebut, maka kami menugaskan Bapak/Ibu sebagai dosen pembimbing guna penyelesaian Skripsi/Tugas Akhir mahasiswa tersebut. Dalam proses bimbingan tidak dibenarkan menawarkan bantuan untuk pembuatan skripsi, tata cara penulisan Skripsi/Tugas Akhir sesuai dengan ketentuan yang ditetapkan oleh Fakultas.

Demikian disampaikan, atas perhatian dan kerja sama Bapak/Ibu diucapkan terima kasih.

Medan, 07 Juni 2017

An. Dekan,





UNIVERSITAS PEMBANGUNAN PANCA BUDI FAKULTAS ILMU KOMPUTER

Jl. Jend. Gatot Subroto Km 4,5 Medan Fax. 061-8458077 PO.BOX : 1099 MEDAN

PROGRAM STUDI SISTEM KOMPUTER
PROGRAM STUDI TEKNIK KOMPUTER

(TERAKREDITASI)
(TERAKREDITASI)

PERMOHONAN MENGAJUKAN JUDUL SKRIPSI


Yang bertanda tangan di bawah ini :

Nama Lengkap : Doly Zainuddin Hasibuan
 Tempat/Tgl. Lahir : / 28 Maret 1993
 Nomor Pokok Mahasiswa : 1524370854
 Program Studi : Sistem Komputer
 Konsentrasi : Keamanan Jaringan Komputer
 Jumlah Kredit yang telah dicapai : 141 SKS, IPK 2.94

Dengan ini mengajukan judul skripsi sesuai dengan bidang ilmu, dengan judul:


No.	Judul Skripsi	Persetujuan
1.	Sistem pendukung keputusan karyawan berprestasi berdasarkan kinerja menggunakan metode simple additive weighting	<input type="checkbox"/>
2.	Aplikasi sistem pendukung keputusan seleksi beasiswa dengan menggunakan metode profile matching	<input type="checkbox"/>
3.	Aplikasi data mining dengan k-means cluster untuk memprediksi produk potensial dan penentuan persediaan produk	<input checked="" type="checkbox"/> <i>Y</i>


Judul yang disetujui oleh Kepala Program Studi diberikan tanda


 (Ir. Bhakti Alamsyah, M.T., Ph.D.)

Medan, 19 Mei 2017
 Pemohon,

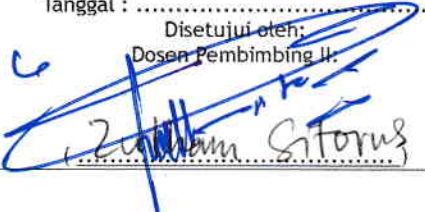
 (Doly Zainuddin Hasibuan)

Nomor :
 Tanggal : 20/05/2017
 Disetujui oleh:
 Ka. Dekan

 (Ir. Bhakti Alamsyah, M.T., Ph.D.)

Tanggal :
 Disetujui oleh:
 Dosen Pembimbing I:

 (Karna Saputra)

Tanggal :
 Disetujui oleh:
 Ka. Prodi Sistem Komputer

 (Eko Hariyanto, S.Kom., M. Kom)

Tanggal :
 Disetujui oleh:
 Dosen Pembimbing II:

 (Zulfaham Sitornus)

No. Dokumen: FM-LPPM-08-01 Revisi: 02 Tgl. Eff: 20 Des 2015

Handwritten marks

ABSTRAK

DOLY ZAINUDDIN HASIBUAN

**Aplikasi Data Mining Dengan K-Means Cluster Untuk Memprediksi Produk
Potensial Dan Penentuan Persediaan Produk
2019**

PT. Cipta Mebelindo Lestari adalah perusahaan yang bergerak pada jasa furniture lebih dari 30 tahun, kegiatan produksi pada produk furniture tentunya menjadikan stok barang bertambah, apabila kegiatan penjualan menurun maka tentunya akan terjadi penimbunan stok yang berlebihan. Untuk mengetahui produk yang paling diminati tentu memerlukan waktu dalam menganalisa data penjualan barang pada PT. MCL. Pembuatan aplikasi data mining dengan K-Means Cluster untuk memprediksi produk potensial dan penentuan persediaan produk memberikan hasil berupa cluster atau pengelompokan data penjualan produk yang paling diminati konsumen, dengan iterasi yang dilakukan berulang-ulang pada data penjualan yang telah di input ke dalam database, sehingga aplikasi dapat mengolah data dan memberikan hasil berupa laporan yang akan dijadikan bahan pertimbangan PT.CML dalam melakukan kegiatan produksi. Dengan mengetahui barang furniture yang paling diminati konsumen, maka produksi barang tidak akan terjadi penimbunan stok sehingga memberikan siklus produksi yang lebih baik bagi PT. CML.

Kata kunci : *Furniture*, data mining, k-means cluster

DAFTAR ISI

HALAMAN JUDUL

LEMBAR PENGESAHAN

SURAT PERNYATAAN

LEMBAR PENGESAHAN

Halaman

ABSTRAK i

KATA PENGANTAR ii

DAFTAR ISI iv

DAFTAR GAMBAR vi

DAFTAR TABEL..... vii

BAB I. PENDAHULUAN 1

1.1 Latar Belakang Masalah. 1

1.2 Rumusan Masalah..... 3

1.3 Batasan Masalah 3

1.4 Tujuan Penelitian 3

1.5 Manfaat Penelitian. 4

1.6 Metodologi Penelitian 4

1.7 Sistematika Penulisan..... 5

BAB II. LANDASAN TEORI 8

2.1 Pt. Cipta Mebelindo Lestari.. 8

2.1.1 Sejarah Perusahaan..... 8

2.1.2 Visi dan Misi Perusahaan 8

2.2 Bidang Produksi 9

2.3 Knowledge Discovery In Database 9

2.3.1 Database. 11

2.4 Tahap – Tahap Knowledge DiscoveryIn Database..... 11

2.5 Data Mining 14

2.5.1 Fungsi Data Mining..... 15

2.5.2 Tujuan Data Mining. 19

2.6 Algoritma K-Mean Clustering 21

2.7 UML (Unified Modelling Language)..... 23

2.7.1 Use Case Diagram..... 24

2.7.2	<i>Activity Diagram</i>	25
2.7.3	Class Diagram.....	26
2.8	Flowchart	29
2.9	Software Pendukung	31
BAB III. ANALISIS DAN PERANCANGAN SISTEM.....		34
3.1	Analisis.....	34
3.1.1	Analisis Sistem	34
3.1.2	Analisis Masalah	34
3.2	Algoritma K-Means Clustering	35
3.2.1	Contoh Kasus	36
3.3	Analisis Kebutuhan Perangkat	50
3.3.1	Analisis Kebutuhan Perangkat Keras.....	51
3.3.2	Analisis Kebutuhan Perangkat Lunak.....	51
BAB IV. IMPLEMENTASI SISTEM.		52
4.1	Implementasi Aplikasi	52
4.2	Pengujian Aplikasi	52
4.2.1	Perangkat Keras (Hardware).....	52
4.2.2	Perangkat Lunak (Software)	53
4.3	Instalasi Aplikasi.....	53
4.4	Pengujian Data Dengan Aplikasi RapidMiner.....	54
4.4.1	Tampilan Utama.....	54
4.4.2	Upload Data Penjualan.....	54
4.4.3	<i>Import Configuration Wizard</i>	55
4.4.4	Proses Metode K-Means	56
4.4.5	Hasil Perhitungan	57
BAB V. PENUTUP.....		60
5.1	Kesimpulan	60
5.2	Saran	61
DAFTAR PUSTAKA		
BIOGRAFI PENULIS		
LAMPIRAN - LAMPIRAN		

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam *database*, data *warehouse*, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti database sistem, data *warehousing*, *statistik*, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi. Selain itu, data mining didukung oleh ilmu lain seperti *neural network*, pengenalan pola, *spatial data analysis*, *image database*, *signal processing* (Gunawan Abdillah, et.al, 2016).

Algoritma K-means pernah digunakan untuk menentukan klasifikasi potensi pemakaian air pelanggan baru di pdam tirta raharja (Gunawan Abdillah, et.al,2016). Selain itu, algoritma K-Means pernah digunakan untuk menentukan calon penerima beasiswa bidik misi di Polbeng, (Jaroji, et.al, 2016). Berdasarkan penelitian terdahulu, penulis akan menerapkan metode K-Means untuk memprediksi produk potensial.

Clustering merupakan pengelompokan data-data yang di mana akan di kelompokkan dan juga dihitung secara matematika maupun dengan teknologi komputer. Dalam clustering, terdapat banyak tehnik-tehnik yang digunakan. Salah satunya adalah dengan menggunakan Algoritma K-Means Cluster. Algoritma K-Means merupakan salah satu metode data clustering non-hierarki yang mengelompokkan data dalam bentuk satu atau lebih cluster atau kelompok.

PT. Cipta Mebelindo Lestari adalah perusahaan yang bergerak dibidang mebel dan furniture di Indonesia. Sebagai perusahaan nasional selama ini pemesanan dilakukan oleh konsumen dengan cara mengunjungi *outlite* atau melalui website. Masalah yang terjadi pada perusahaan adalah tidak adanya aplikasi yang dapat memberikan tolak ukur penjualan barang produksi kepada perusahaan. Dengan tidak adanya sistem yang terkomputerisasi selama ini, perusahaan tidak dapat melakukan perkiraan produksi barang, sehingga barang yang telah menjadi stok menjadi berlebih dan proses produksi kurang efisien.

Algoritma K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data kedalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum penggunaanya dalam praktik. Secara historis, K-Means menjadi salah satu algoritma yang paling penting dalam bidang data mining (Wu dan Kumar, dalam Eko Prasetyo 2014:189).

Diharapkan dengan adanya sebuah aplikasi terkomputerisasi yang dapat melakukan prediksi produk potensial untuk menentukan persediaan produk, maka produksi lebih efisien dan jumlah stok pada perusahaan tidak mengalami penumpukan stok barang. Berdasarkan uraian latar belakang dan uraian permasalahan diatas maka diambil judul skripsi **“Aplikasi Data Mining Dengan K-Means Cluster Untuk Memprediksi Produk Potensial Dan Penentuan Persediaan Produk”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka perlu dirumuskan suatu masalah yang akan dipecahkan/diselesaikan pada penelitian /perancangan ini antarat lain adalah sebagai berikut :

- a. Bagaimana menganalisa data untuk memprediksi produk potensial dan penentuan persediaan produk?
- b. Bagaimana menerapkan metode K-Means Cluster untuk memprediksi produk potensial dan penentuan persediaan produk?
- c. Bagaimana merancang aplikasi data minining menggunakan metode K-Means Cluster?

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut :

- a. Pembahasan pada penelitian ini membahas Perancangan aplikasi data mining dengan menggunakan metode K-Means Cluster.
- b. Aplikasi hanya mendata stok barang dan memprediksi stok barang dan jenis barang produksi.
- c. Aplikasi yang digunakan dalam perancangan ini adalah *RapidMiner Studio*.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut :

- a. Untuk menganalisa data untuk memprediksi produk potensial dan penentuan persediaan produk.

- b. Menerapkan metode K-Means Cluster untuk memprediksi produk potensial dan penentuan persediaan produk.
- c. Untuk merancang aplikasi data minining menggunakan metode K-Means Cluster.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian yang dilakukan bagi skripsi ini antara lain adalah sebagai berikut :

- a. Bagi Perusahaan.
Membantu pihak perusahaan untuk memprediksi produk paling potensial dan stok barang produksi
- b. Bagi mahasiswa.
Sebagai suatu hasil karya ilmu yang didapta selama masa perkuliahan dimana karya tersebut dapat bermanfaat untuk orang lain.
- c. Bagi Universitas Panca Budi Terutama Fakultas Ilmu Komputer.
Sebagai bahan referensi untuk dapat dikembangkan dalam perancangan tampilan dan aplikasi kedepannya.

1.6 Metode Penelitian

Metodologi penelitian merupakan sekumpulan peraturan, kegiatan, dan prosedur yang digunakan oleh pelaku suatu disiplin ilmu.

- a. Studi *Literatur*
Metode pengumpulan data yang dilakukan melalui membaca dan mempelajari referensi berupa makalah,jurnal ilmiah dan buku yang

berkaitan dengan penelitian. Fasilitas *internet* juga digunakan media untuk mencari data atau informasi yang di publikasikan di dunia maya yang berkaitan dengan objek penelitian.

b. Wawancara (*Interview*)

Metode pengumpulan data yang dilakukan melalui tanya jawab secara langsung kepada narasumber yaitu bagian administrasi untuk mendapatkan informasi yang berkaitan dengan objek penelitian.

c. *Observasi*

Metode pengumpulan data dengan mengadakan penelitian dan peninjauan langsung pada PT. Cipta Mebelindo Lestari.

d. Perancangan Aplikasi

Melakukan perancangan aplikasi dengan menggunakan UML sebagai desain aplikasi.

e. Implementasi dan pengujian

Melakukan pengujian Perangkat Lunak berdasarkan implementasi dan analisis yang telah dilakukan

1.7 Sistematika Penulisan

Untuk mempermudah penulisan skripsi ini, sistematika penulisan yang diharapkan dalam penelitian ini adalah sebagai berikut :

BAB I : PENDAHULUAN

Dalam bab ini dibahas mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat dari penelitian, metode penelitian dan sistematika penulisan.

BAB II : LANDASAN TEORI

Dalam bab ini dijelaskan mengenai landasan teori tentang data mining serta membahas pengertian dari metode yang diangkat dan aplikasi yang akan digunakan dalam membangun aplikasi data mining dengan menggunakan smart sistem.

BAB III : ANALISIS DAN PERANCANGAN SISTEM

Dalam bab ini menguraikan analisa dan perancangan, algoritma sistem dalam memprediksi hasil perhitungan dari masalah yang ada dan pemodelan/perancangan sistem.

BAB IV : IMPLEMENTASI SISTEM

Dalam bab ini berisi tentang hasil dari sistem yang telah diselesaikan dan pengujian terhadap sistem.

BAB V : PENUTUP

Dalam bab ini berisi kesimpulan dari masalah yang dihadapkan memberikan saran-saran untuk kemajuan ataupun pengembangan dimasa yang akan datang.

BAB II

LANDASAN TEORI

2.1 PT. Cipta Mebelindo Lestari

2.1.1 Sejarah Perusahaan

Cipta Mebel memulai bisnis rumahnya dengan hanya lima karyawan pada tahun 1979. Pada tahun 1981, karena permintaan pasar yang melonjak, kami pindah dari bisnis rumahan ke bisnis manufaktur furnitur. Karyawan meningkat dari lima hingga tiga ratus lima puluh. Pada tahun 1987, Cipta Mebel mendaftarkan perusahaannya sebagai organisasi dan mengubah namanya menjadi PT. Cipta Mebelindo Lestari (PT CML).

2.1.2 Visi Dan Misi Perusahaan

Visi Dari PT. Cipta Mebelindo Lestari adalah

“We strive to be a reliable and trustworthy company in Southeast Asia that delivers furniture products beyond customers’ expectations.”

Kami berusaha menjadi perusahaan yang dapat diandalkan dan dapat dipercaya di Asia Tenggara yang memberikan produk furnitur di luar harapan pelanggan.

Sedangkan Misi dari PT. Cipta Mebelindo Lestari adalah :

- 1) Kami terus mengevaluasi proses manufaktur kami untuk menyediakan pelanggan dengan daya tahan produk, desain dan nilai yang lebih baik.
- 2) Kami percaya dalam membangun dan mempertahankan hubungan yang dapat dipercaya / dapat diandalkan yang kuat dimana kedua pelanggan

dan perusahaan kami dapat bekerja dan tumbuh bersama untuk mencapai tujuan individu.

- 3) Kami berdedikasi untuk menciptakan perusahaan yang memungkinkan karyawan tumbuh dalam perusahaan melalui kesempatan yang sama.

2.2 Bidang Produksi

Cipta Mebelindo Lestari adalah produsen furnitur kayu yang berbasis di Indonesia. Reputasi kami sebagai produsen furnitur kamar tidur, produsen bunk bed dan produsen furnitur bergensi di Indonesia terutama berasal dari produk unggulan kami seperti Richmond, Claude, Ezra, Linear, dan furnitur Seattle Wood mengambil bagian besar dalam gaya hidup manusia. Sejak dulu, kami selalu menempatkan furnitur kayu di rumah kami untuk membuat rumah kami terlihat indah, tetapi yang penting adalah memiliki banyak manfaat untuk membuat hidup kita lebih mudah. Kayu telah digunakan di seluruh dunia menyediakan banyak jenis furnitur, dari kursi ke meja, dari meja ke lemari. Kayu dipilih karena keserbagunaan dan keindahan yang dibawa ke rumah serta segala yang bisa dibuat untuk kebutuhan perlengkapan interior rumah.

2.3 Knowledge Discovery In Database

Secara sederhana *Data Mining* adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan (Eko Prasetyo, 2012:2). Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*)

untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Fajar Astuti Hermawati, 2009:3). Defenisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan defenisi-defenisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *KnowledgeDiscovery In Database*(KDD) adalah penerapan metode saintifik pada *Data Mining*.

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. *Data mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. *Data mining* bidang dari beberapa keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar. (Kusrini & Emha Taufiq Luthfi, 2010:3-4)

Data mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola, dan trend baru yang bermakna dengan menyaring data yang sangat besar, yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik statistik dan matematik (Kusrini, 2007:82).

Data Mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari

suatu kumpulan data (*Data Mining*, sering juga disebut sebagai *Knowledge Discovery In Database* (KDD)). KDD bertujuan untuk memanfaatkan data dalam basis data dengan dengan mengolahnya sehingga menghasilkan informasi baru yang sempurna (Prasetyo Eko, 2014; 1-2).

2.3.1 Database

Database (Connoly dan Begg, 2010 : 54-66) adalah suatu pembagian kumpulan data yang berisi secara logika, dan keterangan dari masing-masing data yang didesain untuk mendapatkan informasi yang dibutuhkan sebuah organisasi. *Database system* adalah kumpulan program aplikasi yang berinteraksi dengan basis data bersama dengan *Database Management System* (DBMS) dan basis data itu sendiri, sedangkan *Database Management System* (DBMS) adalah merupakan sistem perangkat lunak yang memungkinkan pengguna untuk mendefinisikan, membuat, memelihara dan kontrol ke akses *database*. *Database* (Mcleod, 2007 : 124), adalah kumpulan dari semua data berbasis komputer pada suatu perusahaan.

Dari teori-teori tersebut dapat disimpulkan bahwa *Database* adalah sejumlah data yang terorganisasi dengan *record* dan *field*-nya yang terstruktur dan saling terhubung untuk menyediakan informasi yang dibutuhkan oleh perusahaan.

2.4 Tahap-Tahap Knowledge Discovery In Database

Sebagai suatu rangkaian proses, *Data Mining* dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 2.1. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge base*. Tahap-tahap *Data Mining* ada 6 yaitu :

a. Pembersihan data (*Data Cleaning*)

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevandengan hipotesa *Data Mining* yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performasi dari teknik *Data Mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

b. Integrasi data (*Data Integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk *data mining* tidak hanya berasal dari satu *database* tetapi juga berasal dari beberapa *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

c. Seleksi Data (*Data Selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

d. Transformasi data (*Data Transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *Data Mining*. Beberapa metode *Data Mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan *Clustering* hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

e. Proses *Mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

f. Evaluasi pola (*Pattern Evaluation*),

Untuk mengidentifikasi pola-pola menarik kedalam *Knowledge Based* yang ditemukan. Dalam tahap ini hasil dari teknik *Data Mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses *Data Mining*, mencoba *Data Mining* lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

g. Presentasi pengetahuan (*Knowledge Presentation*),

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses *Data Mining* adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Adakalanya hal ini harus melibatkan orang-orang yang tidak memahami *Data Mining*. Karenanya presentasi hasil *Data Mining* dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses *Data Mining*. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *Data Mining* (Han,2006).

2.5 Data Mining

Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata *mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu *Data Mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, *statistic* dan *database*.

Data Mining adalah proses menerapkan metode ini untuk data dengan maksud untuk mengungkap pola-pola tersembunyi. Dengan arti lain *Data mining* adalah proses untuk penggalian pola-pola dari data. *Data mining* menjadi alat yang semakin penting untuk mengubah data tersebut menjadi informasi. Hal ini sering digunakan dalam berbagai praktek profil, seperti pemasaran, pengawasan, penipuan deteksi dan penemuan ilmiah. Telah digunakan selama bertahun-tahun

oleh bisnis, ilmuwan dan pemerintah untuk menyaring volume data seperti catatan perjalanan penumpang penerbangan, data sensus dan supermarket scanner data untuk menghasilkan laporan riset pasar. Alasan utama untuk menggunakan *Data Mining* adalah untuk membantu dalam analisis koleksi pengamatan perilaku.

Data tersebut rentan terhadap *Collinearity* karena diketahui keterkaitan.

Fakta yang tak terelakkan *Data Mining* adalah bahwa subset/set data yang dianalisis mungkin tidak mewakili seluruh domain, dan karenanya tidak boleh berisi contoh-contoh hubungan kritis tertentu dan perilaku yang ada di bagian lain dari domain. Untuk mengatasi masalah semacam ini, analisis dapat ditambah menggunakan berbasis percobaan dan pendekatan lain, seperti *Choice Modelling* untuk data yang dihasilkan manusia. Dalam situasi ini, yang melekat dapat berupa korelasi dikontrol untuk, atau dihapus sama sekali, selama konstruksi desain eksperimental.

Beberapa teknik yang sering disebut-sebut dalam literatur *Data Mining* dalam penerapannya antara lain: *clustering*, *classification*, *association rule mining*, *neural network*, *genetic algorithm* dan lain-lain. Yang membedakan persepsi terhadap *Data Mining* adalah perkembangan teknik-teknik *Data Mining* untuk aplikasi pada database skala besar. Sebelum populernya *Data Mining*, teknik-teknik tersebut hanya dapat dipakai untuk data skala kecil saja.

2.5.1 Fungsi Data Mining

Teknik – teknik data mining telah digunakan untuk menemukan pola yang tersembunyi dan memprediksi tren masa depan. Dan keuntungan kompetitif dari

data mining termasuk dengan meningkatnya pendapatan, berkurangnya pengeluaran, dan kemampuan pemasaran yang meningkat. (Pujari et. All, 2012)

Data mining dibagi menjadi dua kategori utama (Han dan Kamber, 2006 : 21- 29) yaitu:

A. Prediktif

Tujuan dari tugas prediktif adalah untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai atribut-atribut lain. Atribut yang diprediksi umumnya dikenal sebagai target atau variable tak bebas, sedangkan atribut-atribut yang digunakan untuk membuat prediksi dikenal sebagai *explanatory* atau *variable* bebas.

B. Deskriptif

Tujuan dari tugas deskriptif adalah untuk menurunkan pola-pola (korelasi, *trend*, *cluster*, teritori, dan anomali) yang meringkas hubungan yang pokok dalam data. Tugas *data mining* deskriptif sering merupakan penyelidikan dan seringkali memerlukan teknik *post-processing* untuk validasi dan penjelasan hasil.

Fungsi dari *data mining* juga ada dalam dunia kesehatan, dimana data mining telah digunakan untuk untuk meningkatkan diagnosis dan pengobatan atau lebih mengerti perilaku dari pasien. (Sandra et all, 2009)

Data mining juga memiliki beberapa fungsionalitas yaitu *Concept/Class Description: Characterization and Discrimination, Mining Frequent Patterns, Associations, and Correlations, Classification and Prediction, Cluster Analysis, Outlier analysis*, dan *Evolution analysis*. (Han dan Kamber, 2006 : 21 – 27)

Berikut adalah penjelasan dari masing-masing fungsi diatas:

1. *Concept/Class Description: Characterization and Discrimination*

Data characterization adalah ringkasan dari semua karakteristik atau fitur dari data yang telah diperoleh dari target kelas. Data yang sesuai dengan kelas yang telah ditentukan oleh pengguna biasanya dikumpulkan di dalam *database*. Misalnya, untuk mempelajari karakteristik produk perangkat lunak dimana pada tahun lalu seluruh penjualan telah meningkat sebesar 10%, data yang terkait dengan produk-produk tersebut dapat dikumpulkan dengan menjalankan sebuah *query SQL*. Sedangkan, *data discrimination* adalah perbandingan antara fitur umum objek data target kelas dengan fitur umum objek dari satu atau satu set kelas lainnya. target diambil melalui *query database*. Misalnya, pengguna mungkin ingin membandingkan fitur umum dari produk perangkat lunak yang pada tahun lalu penjualannya meningkat sebesar 10% tetapi selama periode yang sama seluruh penjualan juga menurun setidaknya 30%.

2. *Mining Frequent Patterns, Associations, and Correlations*

Frequent Patterns adalah pola yang sering terjadi di dalam data. Ada banyak jenis dari *frequent patterns*, termasuk di dalamnya pola, sekelompok *item set*, *sub-sequence*, dan sub-struktur. Sebuah *frequent patterns* biasanya mengacu pada satu set item yang sering muncul bersama-sama dalam suatu kumpulan data transaksional, misalnya seperti susu dan roti.

Associations Analysis adalah pencarian aturan-aturan asosiasi yang menunjukkan kondisi-kondisi nilai atribut yang sering terjadi bersama-sama dalam sekumpulan data. Analisis asosiasi sering digunakan untuk menganalisa *Market Basket Analysis* dan data transaksi.

3. *Classification and Prediction*

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan memprediksikan kelas untuk data yang tidak diketahui kelasnya. Model yang diturunkan didasarkan pada analisis dari training data (yaitu objek data yang memiliki label kelas yang diketahui). Model yang diturunkan dapat direpresentasikan dalam berbagai bentuk seperti *If-then* klasifikasi, *decision tree*, dan sebagainya.

Teknik *classification* bekerja dengan mengelompokkan data berdasarkan *data training* dan nilai atribut klasifikasi. Aturan pengelompokan tersebut akan digunakan untuk klasifikasi data baru ke dalam kelompok yang ada. *Classification* dapat direpresentasikan dalam bentuk pohon keputusan (*decision tree*). Setiap *node* dalam pohon keputusan menyatakan suatu tes terhadap atribut *dataset*, sedangkan setiap cabang menyatakan hasil dari tes tersebut. Pohon keputusan yang terbentuk dapat diterjemahkan menjadi sekumpulan aturan dalam bentuk *IF condition THEN outcome*. (Mewati Ayub, 2007 : 7).

Dalam banyak kasus, pengguna ingin memprediksikan nilai-nilai data yang tidak tersedia atau hilang (bukan label dari kelas). Dalam kasus ini nilai data yang akan diprediksi merupakan data *numeric*. Disamping itu, prediksi lebih menekankan pada identifikasi *trend* dari distribusi berdasarkan data yang tersedia.

4. *Cluster Analysis*

Cluster adalah kumpulan objek data yang mirip satu sama lain dalam kelompok yang sama dan berbeda dengan objek data di kelompok lain.

Sedangkan, *Clustering* atau Analisis *Custer* adalah proses pengelompokan satu set benda-benda fisik atau abstrak kedalam kelas objek yang sama. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu *cluster* dan semakin besar perbedaan tiap *cluster* maka kualitas analisis *cluster* semakin baik.

2.5.2 Tujuan Data Mining

Tujuan dari *data mining* (Hoffer, Prescott, dan McFadden, 2007) adalah:

1. *Explanatory*

Untuk menjelaskan beberapa kondisi penelitian, seperti mengapa penjualan truk *pick-up* meningkat di Colorado.

2. *Confirmatory*

Untuk mempertegas hipotesis, seperti halnya dua kali pendapatan keluarga lebih suka dipakai untuk membeli peralatan keluarga dibandingkan dengan satu kali pendapatan keluarga.

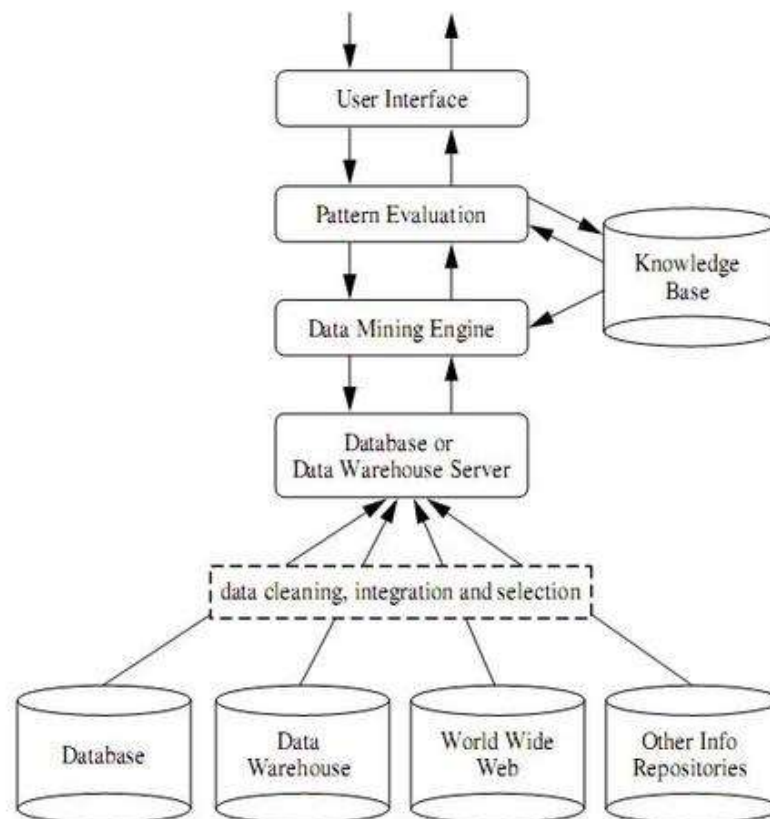
3. *Exploratory*

Untuk menganalisa data yang memiliki hubungan yang baru. Misalnya, pola apa yang cocok untuk kasus penggelapan kartu kredit.

2.5.3 Arsitektur Data Mining

Data mining merupakan proses pencarian pengetahuan yang menarik dari data berukuran besar yang disimpan dalam basis data, *data warehouse* atau tempat penyimpanan informasi lainnya. Dengan demikian arsitektur sistem *data mining* memiliki komponen-komponen utama (Han dan Kamber, 2006) yaitu:

- a. *Database, data warehouse, World Wide Web*, atau tempat penyimpanan informasi lainnya: bisa berbentuk satu atau banyak *database, data warehouse, spreadsheet*, ataupun tempat penyimpanan informasi lainnya. *Data Cleaning, Data Integration* dan *Data Selection* dapat dijalankan pada data tersebut.
- b. *Database* dan *data warehouse server*. Komponen ini bertanggung jawab dalam pengambilan data yang relevan, berdasarkan permintaan pengguna.
- c. *Knowledge Based*. Komponen ini merupakan domain *knowledge* yang digunakan untuk memandu pencarian atau mengevaluasi pola-pola yang dihasilkan. Pengetahuan tersebut meliputi hirarki konsep yang digunakan untuk mengorganisasikan atribut atau nilai atribut kedalam level abstraksi yang berbeda. Pengetahuan tersebut juga dapat berupa kepercayaan pengguna (*user belief*), yang dapat digunakan untuk menentukan kemenarikan pola yang diperoleh.
- d. *Data mining engine*. Bagian ini merupakan komponen penting dalam arsitektur sistem *data mining*. Komponen ini terdiri dari modul-modul fungsional seperti karakterisasi, asosiasi, klasifikasi, dan analisis *cluster*.
- e. *Graphical user interface (GUI)*. Modul ini berkomunikasi dengan pengguna dan *data mining*. Melalui komponen ini, pengguna berinteraksi dengan sistem menggunakan *query*.



Gambar 2.1 Arsitektur sistem *data mining*

2.6 Algoritma K-Means Clustering

Algoritma K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data kedalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum penggunaannya dalam praktik. Secara historis, K-Means menjadi salah satu algoritma yang paling penting dalam bidang data mining (Wu dan Kumar, dalam Eko Prasetyo 2014:189).

Algoritma K-Means merupakan salah satu metode data clustering non-hierarki yang mengelompokkan data dalam bentuk satu atau lebih cluster atau

kelompok. Langkah-langkah melakukan clustering dengan metode K-Means adalah :

- a. Pilih jumlah cluster K.
- b. Inisialisasi K pusat cluster ini bisa dilakukan dengan berbagai cara.
- c. Alokasi semua data/objek ke cluster terdekat.
- d. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang.

$$D(X_1, X_2) = \|X_1 - X_2\|_2 = \sqrt{\sum_{i=1}^P |X_{2j} - X_{1j}|^2}$$

- e. Tugaskan lagi setiap objek memakai pusat cluster yang baru.

Langkah-langkah *K-Means* adalah sebagai berikut :

- a. Langkah pertama: Tanyakan kepada pemakai Algoritma K-Means, sebutlah sebanyak K kelompok.
- b. Langkah kedua: Secara sembarang, pilihlah K buah catatan (dari sekian catatan yang ada) sebagai pusat-pusat kelompok awal.
- c. Langkah ketiga: Untuk setiap catatan, tentukan pusat kelompok terdekatnya dan tetapkan catatan tersebut sebagai anggota dari kelompok yang terdekat pusat kelompoknya. Hitung rasio antara besaran Between Cluster Variation (BCV) dengan Within Cluster Variation, lalu bandingkan rasio tersebut dengan rasio sebelumnya (bila sudah ada). Jika rasio tersebut membesar, lanjutkan ke langkah keempat. Jika tidak, hentikan prosesnya

$$BCV = \sqrt{d(m1, m2) + d(m1, m3) + d(m1, m3)}$$

- d. Langkah keempat: Perbaharui pusat-pusat kelompok (berdasarkan kelompok yang didapat dari langkah ketiga) dan kembalilah ke langkah ketiga.

2.7 UML (*Unified Modelling Language*)

Unified Modelling Language (UML) adalah bahasa standar untuk melakukan spesifikasi, visualisasi, konstruksi, dan dokumentasi dari komponen-komponen perangkat lunak, dan digunakan untuk pemodelan bisnis. UML menggunakan notasi grafis untuk menyatakan suatu desain. Pemodelan dengan UML berarti menggambarkan yang ada dalam dunia nyata ke dalam bentuk yang dapat dipahami dengan menggunakan notasi standar UML.

Pembangunan utama UML adalah diagram. Beberapa diagram ada yang rinci (jenis *timing diagram*) dan bersifat umum (*classdiagram*). Para pengembang sistem berorientasi objek menggunakan bahasa model untuk menggambarkan, membangun, dan mendokumentasikan sistem yang mereka rancang. UML memungkinkan para anggotanya untuk bekerja sama dengan bahasa model yang sama dalam mengaplikasikan beragam sistem. Intinya UML merupakan alat komunikasi yang konsisten dalam men-*support* para pengembang sistem saat ini.

Saat ini piranti perangkat lunak semakin luas dan besar lingkungannya, piranti saat ini seharusnya dirancang dengan memperhatikan hal-hal seperti *scability*, *security*, dan eksekusi dalam keadaan apapun. Selain itu arsitekturnya harus didefinisikan dengan jelas, agar *bug* mudah ditemukan dan diperbaiki, bahkan oleh orang lain selain *programmer* aslinya. Keuntungan lain dari perancangan


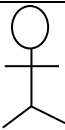

arsitektur yang matang adalah dimungkinkannya pengguna kembali modul atau komponen untuk aplikasi piranti lunak lain yang membutuhkan fungsionalitas yang sama.

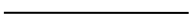
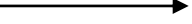
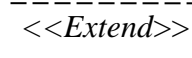
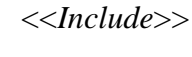
2.7.1 Use Case Diagram

Use Case Diagram adalah suatu representasi/model yang digunakan pada rekayasa perangkat lunak yang menunjukkan sekumpulan *use case* dan *actor* serta hubungan diantara keduanya. Aktor adalah orang atau sistem lain yang berhubungan dengan sistem.

Use Case Diagram menekankan pada “siapa” melakukan “apa” dalam lingkungan sistem perangkat lunak yang akan dibangun. *Use Case Diagram* menjelaskan manfaat suatu sistem jika dilihat menurut pandangan orang yang berada di luar sistem. Diagram ini menunjukkan fungsionalitas suatu sistem atau kelas dan bagaimana sistem tersebut berinteraksi dengan dunia luar. Berikut tabel yang menyajikan notasi *Use Case Diagram*.

Tabel 1. Simbol *Use Case Diagram*

No.	Gambar	Nama	Keterangan
1.		<i>Use Case</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan hasil bagi suatu <i>actor</i> .
2.		<i>Actor</i>	Menggambarkan <i>user</i> atau eksternal entitas atau menggambarkan menerima informasi dari sistem.
3.		<i>System Boundary</i>	Menggambarkan jangkauan sistem.

4.		<i>Association</i>	Apa yang menghubungkan antara objek satu dengan yang lainnya.
5.		<i>Generalization</i>	Dibuat ketika ada sebuah keadaan perlakuan khusus.
6.		<i>Extend</i>	Perluasan dari <i>use case</i> lain jika kondisi atau syarat terpenuhi.
7.		<i>Include</i>	Menjelaskan bahwa <i>use case</i> termaksud dalam <i>use case</i> lain.

Sumber : Rekayasa Perangkat Lunak, Rosa A.s 2014:146




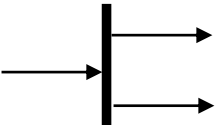
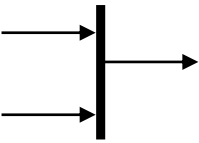
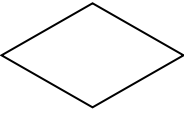
2.7.2 Activity Diagram

Activity Diagram merupakan *state* diagram khusus dimana sebagian besar *state* adalah *action* dan sebagian besar transisi di-*trigger* oleh selesainya *state* sebelumnya (*internal processing*). Digram aktivitas lebih memfokuskan diri pada eksekusi dan alur sistem dari pada bagaimana sistem itu dirakit.

Ketika digunakan dalam pemodelan *software*, diagram aktivitas merepresentasikan pemanggilan suatu fungsi tertentu misalnya *call*. Sedangkan bila digunakan dalam pemodelan bisnis, diagram ini menggambarkan aktivitas yang dipicu oleh kejadian-kejadian diluar, seperti pemasangan atau kejadian-kejadian *internet*. Sama seperti *state*, standar UML menggunakan segi empat dengan sudut membulat untuk menggambarkan aktivitas. *Decision* digunakan untuk menggambarkan *behavior* dalam kondisi tertentu.

Diagram ini juga dapat digunakan untuk memodelkan aksi dan hasil ketika operasi berlangsung. Berikut tabel yang menyajikan notasi *Activity diagram* dibawah ini.

Tabel 2. Simbol *Activity diagram*

No.	Gambar	Nama	Keterangan
1.		<i>Start Point</i>	Titik awal yang menunjukkan dimulainya suatu aliran aktivitas dalam suatu sistem pada sebuah <i>activity diagram</i> .
2.		<i>End point</i>	Titik akhir yang menunjukkan akhir suatu aliran aktivitas dalam suatu sistem pada sebuah <i>activity diagram</i> .
3.		<i>Activity</i>	Simbol yang menggambarkan sebuah pekerjaan atau tugas dalam aliran aktivitas dalam suatu sistem.
4.		<i>Fork</i>	Simbol yang digunakan untuk menggambarkan proses percabangan dalam aliran aktivitas sistem.
5.		<i>Joint</i>	Simbol yang digunakan untuk menggambarkan proses penggabungan
6.		<i>Decision</i>	Suatu titik atau <i>point</i> pada <i>activity diagram</i> yang mengindikasikan suatu kondisi

Sumber :Rekayasa Perangkat Lunak, Rosa A.s (2014:186)

2.7.3 Class Diagram

Class adalah sebuah spesifikasi yang jika diinstansi akan menghasilkan sebuah objek dan merupakan inti dari pengembangan dan desain berorientasi

objek. *Class* menggambarkan keadaan (atribut/properti) suatu sistem, sekaligus menawarkan layanan untuk memanipulasi keadaan tersebut (metoda/fungsi). *Class diagram* menggambarkan struktur dan deskripsi *class*, *package* dan objek beserta hubungan satu sama lain seperti *containment*, pewarisan, asosiasi, dan lain-lain. *Class* memiliki tiga area pokok antara lain yaitu adalah :

1. Nama (*stereotype*)
2. Atribut
3. Metoda

Atribut dan metoda dapat *memiliki* salah satu sifat berikut :

- a. *Private*, tidak dapat dipanggil dari luar *class* yang bersangkutan.
- b. *Protected*, hanya dapat dipanggil oleh *class* yang bersangkutan dan anak-anak yang mewarisinya.
- c. *Public*, dapat dipanggil oleh siapa saja.

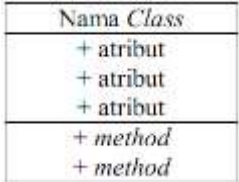

Class dapat merupakan implementasi dari sebuah *interface*, yaitu *class* abstrak yang hanya memiliki metoda. *Interface* tidak dapat langsung diinstansikan, tetapi harus diimplementasikan dahulu menjadi sebuah *class*. Dengan demikian *interface* mendukung resolusi metoda pada saat *run-time*.

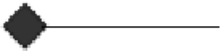
Ada beberapa hubungan antar *class* dalam *class*, hubungan itu antara lain sebagai berikut :

1. Asosiasi, yaitu hubungan statis antar *class*. Umumnya menggambarkan *class* yang memiliki atribut berupa *class* lain, atau *class* yang harus mengetahui eksistensi *class* lain. Panah *navigability* menunjukkan arah *query* antar *class*.
2. Agregasi, yaitu hubungan yang menyatakan bagian (“terdiri atas..”).



3. Pewarisan, yaitu hubungan hirarkis antar *class*. *Class* dapat diturunkan dari *class* lain dan mewarisi semua atribut dan metoda *class* asalnya dan menambahkan fungsionalitas baru, sehingga ia disebut anak dari *class* diwarisinya. Kebalikan dari pewarisan adalah generalisasi.
4. Hubungan dinamis, yaitu rangkaian pesan (*message*) yang di-*passing* dari satu *class* kepada *class* lain. Hubungan dinamis dapat digambarkan dengan menggunakan *sequence diagram* yang akan dijelaskan kemudian.

Tabel 3. Simbol *Class Diagram*

No.	Simbol	Nama	Keterangan
1.		Class	Class Class adalah blok - blok pembangun pada pemrograman berorientasi obyek. Sebuah class digambarkan sebagai sebuah kotak yang terbagi atas 3 bagian. Bagian atas adalah bagian nama dari class. Bagian tengah mendefinisikan property/atribut class. Bagian akhir mendefinisikan methodmethod dari sebuah class.
2.		Association	Association Sebuah asosiasi merupakan sebuah relationship paling umum antara 2 class dan dilambangkan oleh sebuah garis yang menghubungkan antara 2 class. Garis ini bisa melambangkan tipe-tipe relationship dan juga dapat menampilkan hukum-hukum multiplisitas pada sebuah relationship.

			(Contoh: One-to-one, one-to-many, many-to-many).
3.		composition	Sebuah relationship composition digambarkan sebagai garis dengan ujung berbentuk jajaran genjang berisi/solid.

Tabel 3. Simbol Class *Diagram* (lanjutan)



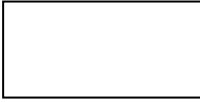
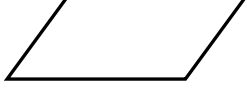
No.	Simbol	Nama	Keterangan
4.		Dependency	Umumnya penggunaan dependency digunakan untuk menunjukkan operasi pada suatu class yang menggunakan class yang lain. Sebuah dependency dilambangkan sebagai panah bertitik-titik.
5.		Aggregation	Aggregation mengindikasikan keseluruhan bagian relationship dan biasa disebut relasi

Sumber : Grady Booch, 2005

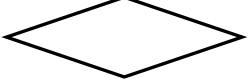
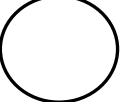
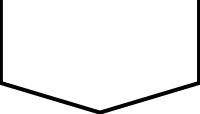

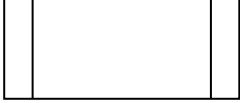
2.8 Flowchart

Sistem *flowchat* yaitu bagan yang (*flow*) program atau prosedur sistem secara logika. Simbol-simbol *flowchart* pada tabel sebagai berikut :

Tabel 4. Simbol-Simbol Flowchart

No	Simbol	Nama	Fungsi
1		Terminator	Permulaan/ akhir program
3		Preparation	Proses input/output data parameter, dan informasi.
4		Proses	Proses penghitungan/proses pengolahan data.
5		<i>Input/Output Data</i>	Proses <i>input/output</i> data, parameter. Informasi.

Tabel 4. Simbol-Simbol Flowchart (lanjutan)

No	Simbol	Nama	Fungsi
6		<i>Decision</i>	Perbandingan pernyataan, penyeleksian data yang memberikan pilihan untuk langkah selanjutnya.
7		<i>One Page Connector</i>	Penghubung bagian-bagian <i>flowchart</i> yang berada pada satu halaman.
8		<i>Off-line connector</i>	Simbol masuk atau keluarnya suatu prosedur pada lembar kertas lainnya
9		<i>Document</i>	Simbol untuk data yang terbentuk kertas maupun untuk informasi
10		<i>Predefined</i>	Proses untuk menyatukan sekumpulan langkah proses yang ditulis sebagai prosedur

Sumber : Rekayasa Perangkat Lunak, Rosa A.s (2014:176)

2.9 Software Pendukung

Adapun software pendukung dalam pembuatan aplikasi sistem pendukung keputusan yang akan dirancang antara lain adalah :

a. RapidMiner Studio

RapidMiner dilisensikan di bawah GNU Affero General Public License versi 3 dan saat ini tersedia dalam versi 5.2. Ini pada awalnya dikembangkan mulai tahun 2001 di kursi untuk arti Pejabat intelijen dari Universitas Dortmund dengan nama “Yale”. Sejak tahun 2007, program ini telah terus oleh Rapid-I GmbH, yang didirikan oleh mantan anggota kursi, dan itu telah meningkat dengan pesat sejak saat itu. Pengenalan RapidAnalytics pada tahun 2010 berarti sekarang ada versi server yang sesuai tersedia, yang membuat kolaborasi dan efisiensi penggunaan sumber daya komputer mungkin.

RapidMiner sebelumnya dikenal sebagai YALE (*Yet Another Learning Environment*), mulai dikembangkan pada tahun 2001 oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer dari Unit Kecerdasan Buatan Universitas Teknik Dortmund. Mulai tahun 2006, perkembangannya didorong oleh Rapid-I, sebuah perusahaan yang didirikan oleh Ingo Mierswa dan Ralf Klinkenberg pada tahun yang sama. Pada tahun 2007, nama perangkat lunak itu berubah dari YALE menjadi RapidMiner. Pada tahun 2013, perusahaan melakukan rebranding dari Rapid-I menjadi RapidMiner.

Rapid Miner adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan

lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambahan teks, dan analisis prediktif. Hal ini digunakan untuk bisnis dan komersial, juga untuk penelitian, pendidikan, pelatihan, rapid prototyping, dan pengembangan aplikasi serta mendukung semua langkah dalam proses pembelajaran mesin termasuk persiapan data, hasil visualisasi, validasi model, dan optimasi. RapidMiner dikembangkan pada model inti terbuka. Dengan RapidMiner Studio Free Edition, yang terbatas untuk 1 prosesor logika dan 10.000 baris data, tersedia di bawah lisensi AGPL. Harga komersial dimulai dari \$2.500 dan tersedia dari pengembang.

Salah satu keuntungan terbesar adalah tanpa diragukan lagi fakta bahwa RapidMiner tersedia untuk download gratis di Community Edition. Siswa karena itu dapat menginstalnya pada komputer pribadi mereka hanya dalam cara yang sama seperti universitas dapat membuat instalasi RapidMiner tersedia di komputer lembaga. Jadi memulai cepat dan gratis. Berkat distribusi yang luas RapidMiner ini, siswa juga memiliki kesempatan selama masa studi mereka bekerja dengan alat yang mereka dapat benar-benar menggunakan di tempat kerja nanti.

RapidMiner adalah salah satu software untuk pengolahan data mining. Pekerjaan yang dilakukan oleh RapidMiner text mining adalah berkisar dengan analisis teks, mengekstrak pola-pola dari data set yang besar dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan database. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi bermutu tertinggi dari teks yang diolah.

RapidMiner menyediakan prosedur data mining dan machine learning, di dalamnya termasuk: ETL (extraction, transformation, loading), data preprocessing, visualisasi, modelling dan evaluasi. Proses data mining tersusun atas operator-operator yang nestable, dideskripsikan dengan XML, dan dibuat dengan GUI. Penyajiannya dituliskan dalam bahasa pemrograman Java.

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1. Analisis

3.1.1 Analisis Sistem

Analisis yang dilakukan pada PT. Cipta Mebelindo terhadap produk mebel yang diminati konsumen dan produk yang diproduksi oleh perusahaan. Pengumpulan data penjualan pada PT. Cipta Mebelindo dilakukan untuk mengetahui jenis produk mebel yang diminati oleh konsumen, setelah data didapat kemudian dikelompokkan menjadi beberapa item yang paling banyak terjadi transaksi pada perusahaan.

Proses analisis pada sistem yang akan dirancang menggunakan metode K-Means Cluster untuk melakukan pengelompokan yang nantinya akan memberikan hasil berupa pengelompokan data produk yang paling diminati oleh konsumen, sehingga pada bagian produksi di perusahaan tidak akan terjadi penumpukan stok produk pada item tertentu.

3.1.2 Analisis Masalah

Analisis permasalahan adalah suatu proses mengumpulkan dan menginterpretasikan kenyataan-kenyataan yang ada, mendiagnosa persoalan dan menggunakan keduanya untuk memperbaiki sistem. Tahapan analisis merupakan sistem informasi yang dibutuhkan kedalam tiap bagian sama dan komponen

dengan maksud mengidentifikasi dan mengevaluasi permasalahan-permasalahan yang dihadapi dalam perusahaan.

Berdasarkan analisa yang dilakukan pada penelitian ini, proses penjualan barang dilakukan dengan cara manual atau tidak terkomputerisasi dengan baik. Sehingga data yang ada dalam bentuk catatan yang dibuat dalam beberapa *letter file*.

3.2 Algoritma K-Means Clustering

Algoritma sistem adalah tahapan penyelesaian yang dibuat guna menyelesaikan permasalahan dalam bentuk sistematis pada sebuah sistem dengan cara menganalisa permasalahan dan merubahnya dalam bentuk penyelesaian yang terstruktur. Algoritma K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data kedalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum penggunaannya dalam praktik. Algoritma K-Means merupakan salah satu metode data clustering non-hierarki yang mengelompokkan data dalam bentuk satu atau lebih cluster atau kelompok.

Langkah-langkah melakukan clustering dengan metode K-Means adalah:

1. Pilih jumlah cluster K.
2. Inisialisasi K pusat cluster ini bisa dilakukan dengan berbagai cara.
3. Alokasi semua data/objek ke cluster terdekat.
4. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang.
5. Tugaskan lagi setiap objek memakai pusat cluster yang baru

Sedangkan untuk langkah-langkah pada metode K-Means adalah sebagai berikut:

1. Tentukan jumlah cluster
2. Alokasikan data ke dalam kelompok secara acak
3. Hitung pusat cluster (centroid) menggunakan mean utk masing-masing kelompok
4. Alokasikan masing-masing data ke centroid terdekat
5. Kembali ke langkah 3, jika masih ada data yang berpindah cluster atau jika nilai centroid diatas nilai ambang, atau jika nilai pada fungsi obyektif yang digunakan masih diatas ambang

3.2.1 Contoh Kasus

Berikut ini merupakan tabel penjualan barang mebel dari tahun 2013 sampai dengan tahun 2017 untuk 20 item yang diambil dari perusahaan untuk data perhitungan pada metode.

Tabel 3.1. Penjualan Barang

NAMA BARANG	THN 2013	THN 2014	THN 2015	THN 2016	THN 2017
Lemari Pakaian Mewah Racoco Flower	114	114	139	129	117
AIRLAND 505 ESSENTIAL 27 cm (Medium firm)	92	114	99	117	139
INOAC YUKATA (BEST SELLER)	93	128	84	122	117
LEMARI MIRACLE 2 pintu XL	96	123	88	105	94
Meja Rias Volga - Kaca Sliding	135	118	86	85	66
Ranjang Besi Levander + Kasur	122	69	139	139	123
VIP FILING CABINET V-302	77	125	121	89	76
SOFA BED TURTLE	125	60	110	80	75
DIRBT-MDT	119	119	134	83	85
Rak Buku MPR 199	87	105	93	76	105
Meja UNO Modern Series	135	115	126	106	96

Tabel 3.1. Penjualan Barang (lanjutan)

NAMA BARANG	THN 2013	THN 2014	THN 2015	THN 2016	THN 2017
Meja Rapat Euro (kaki besi)	75	139	104	126	135
PARTISI 4 STAFF VARIASI G	63	66	99	130	128
SOFA Kennedy 321 tanpa meja	107	124	84	138	133
Meja Tamu Milan	130	88	126	99	87
Kitchen Set Mutiara S Series	85	124	133	134	71
MEJA MAKAN GRANIT 4 KURSI	78	85	111	64	93
Kursi Taman Celebrity 08	77	61	135	75	87
SD 1512 TECIDO	115	113	118	133	124
Lunar SR 9390	73	124	93	125	107

Sumber : PT. Cipta Mebelindo Lestari

Berikut ini merupakan pengkodean barang yang terjual berdasarkan nama data barang diatas.

Tabel 3.2. Daftar pengkodean Dan Jumlah Terjual

NAMA BARANG	KODE	THN 2013	THN 2014	THN 2015	THN 2016	THN 2017
Lemari Pakaian Mewah Racoco Flower	KD_1	114	114	139	129	117
AIRLAND 505 ESSENTIAL 27 cm (Medium firm)	KD_2	92	114	99	117	139
INOAC YUKATA (BEST SELLER)	KD_3	93	128	84	122	117
LEMARI MIRACLE 2 pintu XL	KD_4	96	123	88	105	94
Meja Rias Volga - Kaca Sliding	KD_5	135	118	86	85	66
Ranjang Besi Levander + Kasur	KD_6	122	69	139	139	123
VIP FILING CABINET V-302	KD_7	77	125	121	89	76
SOFA BED TURTLE	KD_8	125	60	110	80	75
DIRBT-MDT	KD_9	119	119	134	83	85
Rak Buku MPR 199	KD_10	87	105	93	76	105
Meja UNO Modern Series	KD_11	135	115	126	106	96
Meja Rapat Euro (kaki besi)	KD_12	75	139	104	126	135
PARTISI 4 STAFF VARIASI G	KD_13	63	66	99	130	128
SOFA Kennedy 321 tanpa meja	KD_14	107	124	84	138	133
Meja Tamu Milan	KD_15	130	88	126	99	87
Kitchen Set Mutiara S Series	KD_16	85	124	133	134	71
MEJA MAKAN GRANIT 4 KURSI	KD_17	78	85	111	64	93

Tabel 3.2. Daftar pengkodean Dan Jumlah Terjual (lanjutan)

NAMA BARANG	KODE	THN 2013	THN 2014	THN 2015	THN 2016	THN 2017
Kursi Taman Celebrity 08	KD_18	77	61	135	75	87
SD 1512 TECIDO	KD_19	115	113	118	133	124
Lunar SR 9390	KD_20	73	124	93	125	107

Sumber : PT. Cipta Mebelindo Lestari

1. Tentukan jumlah cluster

Berdasarkan data penjualan setiap tahunnya ada lima cluster yang akan dijadikan centroid awal dengan jumlah 5 cluster.

2. Alokasikan data ke dalam kelompok secara acak

Data penjualan dialokasikan berdasarkan urutan kode yang dapat dilihat pada tabel 2.

3. Hitung pusat cluster (centroid) menggunakan K-means untuk masing-masing kelompok

Pusat cluster yang diambil secara acak adalah:

Pusat cluster 0 = KD_1 C_0 114 114 139 129 117

Pusat cluster 1 = KD_2 C_1 92 114 99 117 139

Pusat cluster 2 = KD_6 C_2 122 69 139 139 123

Pusat cluster 3 = KD_4 C_3 96 123 88 105 94

Pusat cluster 4 = KD_8 C_4 125 60 110 80 75

Rumus menghitung jarak pusat cluster *euclidian* adalah sebagai berikut.

$$D(X_1, X_2) = \|X_1 - X_2\|_2 = \sqrt{\sum_{j=1}^p |X_{2j} - X_{1j}|^2}$$

1. Jarak objek A

a. Pusat cluster 0

$$d = \sqrt{(114 - 114)^2 - (114 - 114)^2 - (139 - 139)^2 - (129 - 129)^2 - (117 - 117)^2}$$

$$= 0$$

b. Pusat cluster 1

$$d = \sqrt{(114 - 92)^2 - (114 - 114)^2 - (139 - 99)^2 - (129 - 117)^2 - (117 - 139)^2}$$

$$= 2250$$

c. Pusat cluster 2

$$d = \sqrt{(114 - 122)^2 - (114 - 69)^2 - (139 - 139)^2 - (129 - 139)^2 - (117 - 123)^2}$$

$$= 2169$$

d. Pusat cluster 3

$$d = \sqrt{(114 - 96)^2 - (114 - 123)^2 - (139 - 88)^2 - (129 - 105)^2 - (117 - 94)^2}$$

$$= 3805$$

e. Pusat cluster 4

$$d = \sqrt{(114 - 125)^2 - (114 - 60)^2 - (139 - 110)^2 - (129 - 80)^2 - (117 - 75)^2}$$

$$= 7933$$

Proses perhitungan diatas dilakukan untuk masing-masing item terjual sehingga didapat hasil perhitungan pada tabel 3.

Tabel 3.3. Hasil Perhitungan Pusat Cluster Baru Iterasi 0

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	0	2250	2169	3805	7933	0
KD_2	2250	0	4395	2375	8535	0
KD_3	3291	931	6860	862	8860	862
KD_4	3805	2375	7540	0	5468	0
KD_5	7383	6581	11388	1252	4056	1252
KD_6	2169	4395	0	7540	6710	0
KD_7	3763	5373	8214	1692	4476	1692
KD_8	7933	8535	6710	5468	0	0
KD_9	3195	5349	7108	2720	4172	2720
KD_10	5177	2959	7740	1320	3268	1320
KD_11	1161	2743	4116	1552	4408	1161
KD_12	2222	764	6485	2655	12043	764
KD_13	4077	2623	1774	5184	5528	1774
KD_14	3469	817	6166	2638	11518	817
KD_15	2661	4471	3434	2788	1550	1550
KD_16	2306	6176	5827	3407	7597	2306
KD_17	6462	5924	7609	3673	1253	1253
KD_18	6678	8588	5517	7021	843	843
KD_19	508	866	2421	2703	8093	508
KD_20	2373	1243	5642	618	7486	618

4. Alokasikan masing-masing data ke centroid terdekat

Dari tabel 3 di atas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 1 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 = C_0 112.3 116.5 129 125.5 102

Pusat cluster 1 = C_1 91.3 125.7 95.7 127.0 135.7

Pusat cluster 2 = C_2 92.5 67.5 119 134.5 125.5

Pusat cluster 3 = C_3 97.14 120.29 99.86 97.86 92.86

Pusat cluster 4 = C_4 102.5 73.5 120.5 79.5 85.5

5. Kembali ke langkah 3, jika masih ada data yang berpindah cluster atau jika nilai centroid diatas nilai ambang, atau jika nilai pada fungsi obyektif yang digunakan masih diatas ambang

Tabel 3.4. Hasil Perhitungan Pusat Cluster Baru Iterasi 1

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	345.3	2389.0	2686.3	3141.3	5436.5	345.3
KD_2	2367.8	259.0	3051.3	2541.0	6381.5	259.0
KD_3	2413.8	516.7	5114.3	1480.9	7110.5	516.7
KD_4	2223.8	2290.7	5907.3	201.4	4235.5	201.4
KD_5	4810.3	6813.3	9672.3	1121.7	3613.5	1121.7
KD_6	2989.3	5424.0	458.3	6788.6	5328.5	458.3
KD_7	2179.8	5660.7	7846.3	852.0	2858.5	852.0
KD_8	6365.3	10440.7	5690.3	4402.9	425.5	425.5
KD_9	2133.3	6044.7	7196.3	1471.7	2281.5	1471.7
KD_10	3912.8	3980.0	5930.3	916.0	2156.5	916.0
KD_11	450.3	3092.0	4030.3	825.4	2597.5	450.3
KD_12	2257.8	265.0	5517.3	2957.6	9202.5	265.0
KD_13	4195.8	3667.3	458.3	5250.0	4914.5	458.3
KD_14	3203.8	282.7	4500.3	3498.0	9565.5	282.7
KD_15	1766.3	5530.0	3249.3	1794.3	650.5	650.5
KD_16	1132.8	5633.7	6366.3	2908.4	5904.5	1132.8
KD_17	5213.8	7691.7	6411.3	2534.7	543.5	543.5
KD_18	5926.8	10815.7	5336.3	5326.7	414.5	414.5
KD_19	676.3	855.0	2098.3	2605.0	5923.5	676.3
KD_20	1416.8	854.0	4320.3	1021.7	5868.5	854.0

Dari tabel 4 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 4 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 =C_0 112.3 116.5 129.0 125.5 102.0

Pusat cluster 1 = C_1	88	125.8	92.8	125.6	126.2
Pusat cluster 2 = C_2	92.5	67.5	119	134.5	125.5
Pusat cluster 3 = C_3	102.8	118	104.4	87.6	85.2
Pusat cluster 4 = C_4	95	78	124	79.3	89

Tabel 3.5. Hasil Perhitungan Pusat Cluster Baru Iterasi 2

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	345.3	2395.9	2686.3	3949.6	4790.8	345.3
KD_2	2367.8	419.5	3051.3	3814.8	5842.8	419.5
KD_3	2413.8	184.9	5114.3	2720.6	6706.4	184.9
KD_4	2223.8	1500.1	5907.3	681.0	4005.8	681.0
KD_5	4810.3	5426.5	9672.3	746.2	3645.1	746.2
KD_6	2989.3	5584.5	458.3	7688.2	5049.1	458.3
KD_7	2179.8	4666.5	7846.3	437.0	2498.4	437.0
KD_8	6365.3	9363.3	5690.3	3579.4	746.4	746.4
KD_9	2133.3	5286.9	7196.3	914.6	1834.4	914.6
KD_10	3912.8	3343.3	5930.3	841.4	1965.1	841.4
KD_11	450.3	2562.1	4030.3	963.0	2173.1	450.3
KD_12	2257.8	390.3	5517.3	4423.6	8434.8	390.3
KD_13	4195.8	3662.1	458.3	6402.6	4889.1	458.3
KD_14	3203.8	299.7	4500.3	5281.4	9105.8	299.7
KD_15	1766.3	4817.3	3249.3	1527.0	529.8	529.8
KD_16	1132.8	4739.9	6366.3	3226.4	5519.4	1132.8
KD_17	5213.8	6902.7	6411.3	1775.2	486.1	486.1
KD_18	5926.8	10087.9	5336.3	4373.2	450.8	450.8
KD_19	676.3	885.5	2098.3	3788.8	5386.1	676.3
KD_20	1416.8	387.3	4320.3	2069.8	5508.4	387.3

Dari tabel 5 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 5 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru

Pusat cluster 0 = C_0	114.0	105.5	134.3	127.0	101.8
-----------------------	-------	-------	-------	-------	-------

Pusat cluster 1 = C ₁	90	97.7	105.3	126.7	130.3
Pusat cluster 2 = C ₂	87	128.8	91.3	127.8	123
Pusat cluster 3 = C ₃	98.8	117.8	97	88.8	85.3
Pusat cluster 4 = C ₄	105.8	82.6	123.2	80.2	85.4

Tabel 3.6. Hasil Perhitungan Pusat Cluster Baru Iterasi 3

KODE	C ₀	C ₁	C ₂	C ₃	C ₄	Min
KD_1	331.4	1607.4	2562.2	4421.4	4623.8	331.4
KD_2	2824.4	477.4	654.2	3711.9	5812.6	477.4
KD_3	3309.9	1577.8	128.2	2393.4	6356.4	128.2
KD_4	3007.4	2737.8	1411.2	451.9	3570.0	451.9
KD_5	5547.4	6707.1	5267.7	541.9	3065.6	541.9
KD_6	1958.4	2193.1	6011.7	8113.9	5322.0	1958.4
KD_7	2699.9	5376.4	4619.7	735.9	1997.2	735.9
KD_8	5593.9	6715.1	9700.2	3711.9	812.4	812.4
KD_9	2403.9	5267.8	5401.2	1423.9	1462.8	1423.9
KD_10	4340.4	3417.8	3569.2	742.9	1834.4	742.9
KD_11	653.4	2378.4	2646.7	1297.9	1864.8	653.4
KD_12	3182.9	1747.4	426.7	4386.9	8138.2	426.7
KD_13	3551.9	1086.4	4051.7	6246.9	5198.8	1086.4
KD_14	3971.9	1301.1	300.2	4921.9	8858.4	300.2
KD_15	1391.9	3203.8	5033.7	1865.4	417.2	417.2
KD_16	1367.4	5038.1	4510.7	3599.4	4932.6	1367.4
KD_17	5042.4	5525.4	7277.2	1961.9	502.6	502.6
KD_18	4939.4	6784.8	10592.7	4878.4	664.2	664.2
KD_19	852.4	500.8	1020.2	3939.4	5238.2	500.8
KD_20	2116.4	1409.8	303.2	1867.9	5132.4	303.2

Dari tabel 6 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 6 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru

Pusat cluster 0 = C ₀	118.2	108.0	124.6	118.6	94.6
----------------------------------	-------	-------	-------	-------	------

Pusat cluster 1 = C_1	90	97.7	105.3	126.7	130.3
Pusat cluster 2 = C_2	87	128.8	91.3	127.8	123
Pusat cluster 3 = C_3	102.8	118	104.4	87.6	85.2
Pusat cluster 4 = C_4	102.5	73.5	120.5	79.5	85.5

Tabel 3.7. Hasil Perhitungan Pusat Cluster Baru Iterasi 4

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	857.5	1607.4	2562.2	3949.6	5436.5	857.5
KD_2	2691.5	477.4	654.2	3814.8	6381.5	477.4
KD_3	2586.9	1577.8	128.2	2720.6	7110.5	128.2
KD_4	1772.1	2737.8	1411.2	681.0	4235.5	681.0
KD_5	3553.7	6707.1	5267.7	746.2	3613.5	746.2
KD_6	2954.9	2193.1	6011.7	7688.2	5328.5	2193.1
KD_7	1565.3	5376.4	4619.7	437.0	2858.5	437.0
KD_8	4398.1	6715.1	9700.2	3579.4	425.5	425.5
KD_9	1569.7	5267.8	5401.2	914.6	2281.5	914.6
KD_10	2961.7	3417.8	3569.2	841.4	2156.5	841.4
KD_11	228.5	2378.4	2646.7	963.0	2597.5	228.5
KD_12	3115.5	1747.4	426.7	4423.6	9202.5	426.7
KD_13	3720.1	1086.4	4051.7	6402.6	4914.5	1086.4
KD_14	3766.5	1301.1	300.2	5281.4	9565.5	300.2
KD_15	855.7	3203.8	5033.7	1527.0	650.5	650.5
KD_16	1153.9	5038.1	4510.7	3226.4	5904.5	1153.9
KD_17	3737.9	5525.4	7277.2	1775.2	543.5	543.5
KD_18	4317.1	6784.8	10592.7	4373.2	414.5	414.5
KD_19	1143.5	500.8	1020.2	3788.8	5923.5	500.8
KD_20	1494.5	1409.8	303.2	2069.8	5868.5	303.2

Dari tabel 7 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 7 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru

Pusat cluster 0 = C_0	111.3	117.7	132.7	123.0	94.7
-----------------------	-------	-------	-------	-------	------

Pusat cluster 1 = C ₁	98	90.5	113.8	129.8	128.5
Pusat cluster 2 = C ₂	87	128.8	91.3	127.8	123
Pusat cluster 3 = C ₃	102.8	118	104.4	87.6	85.2
Pusat cluster 4 = C ₄	82	58.8	96.4	63.6	68.4

Tabel 3.8. Hasil Perhitungan Pusat Cluster Baru Iterasi 5

KODE	C ₀	C ₁	C ₂	C ₃	C ₄	Min
KD_1	591.0	1338.6	2562.2	3949.6	11532.9	591.0
KD_2	3167.7	1048.6	654.2	3814.8	10899.7	654.2
KD_3	2993.3	2488.6	128.2	2720.6	10725.9	128.2
KD_4	2363.3	3524.1	1411.2	681.0	6575.5	681.0
KD_5	4467.3	7472.1	5267.7	746.2	4129.5	746.2
KD_6	3478.0	1239.6	6011.7	7688.2	10625.1	1239.6
KD_7	1728.7	5680.6	4619.7	437.0	5695.5	437.0
KD_8	6088.7	6308.6	9700.2	3579.4	541.9	541.9
KD_9	1704.7	5321.1	5401.2	914.6	5726.7	914.6
KD_10	4074.0	4093.1	3569.2	841.4	3644.3	841.4
KD_11	366.0	2407.6	2646.7	963.0	6647.1	366.0
KD_12	2949.0	2526.6	426.7	4423.6	14826.1	426.7
KD_13	5011.3	853.1	4051.7	6402.6	8038.7	853.1
KD_14	4107.3	2104.6	300.2	5281.4	14138.3	300.2
KD_15	1578.0	2856.1	5033.7	1527.0	3375.9	1527.0
KD_16	747.7	4830.1	4510.7	3226.4	10556.5	747.7
KD_17	5053.7	5641.1	7277.2	1775.2	1508.9	1508.9
KD_18	5613.7	6062.6	10592.7	4373.2	1975.7	1975.7
KD_19	1201.0	572.1	1020.2	3788.8	11344.9	572.1
KD_20	1808.0	2062.6	303.2	2069.8	9531.5	303.2

Dari tabel 8 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 8 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 = C ₀	111.3	117.7	132.7	123.0	94.7
----------------------------------	-------	-------	-------	-------	------

Pusat cluster 1 = C_1	100	82.7	118.7	134	125
Pusat cluster 2 = C_2	88	125.8	92.8	125.6	126.2
Pusat cluster 3 = C_3	107.3	113.0	108.0	89.5	85.5
Pusat cluster 4 = C_4	93.3	68.7	118.7	73.0	85.0

Tabel 3.9. Hasil Perhitungan Pusat Cluster Baru Iterasi 6

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	591.0	1498.2	2395.9	3521.2	6649.2	591.0
KD_2	3167.7	1861.6	419.5	3715.8	7295.2	419.5
KD_3	2993.3	3471.9	184.9	2863.8	8147.6	184.9
KD_4	2363.3	4373.2	1500.1	823.8	5000.2	823.8
KD_5	4467.3	8232.6	5426.5	937.2	4047.6	937.2
KD_6	3478.0	651.2	5584.5	6768.2	6242.2	651.2
KD_7	1728.7	6246.6	4666.5	433.8	3532.2	433.8
KD_8	6088.7	6029.9	9363.3	3031.2	330.9	330.9
KD_9	1704.7	5775.2	5286.9	766.2	2894.2	766.2
KD_10	4074.0	4934.6	3343.3	871.8	2394.2	871.8
KD_11	366.0	2759.2	2562.1	738.2	3452.2	366.0
KD_12	2949.0	3577.6	390.3	4506.8	10489.2	390.3
KD_13	5011.3	726.6	3662.1	5780.8	5522.2	726.6
KD_14	4107.3	2997.2	299.7	5305.8	10806.2	299.7
KD_15	1578.0	2781.2	4817.3	1064.2	1144.2	1064.2
KD_16	747.7	4844.9	4739.9	2958.8	7192.6	747.7
KD_17	5053.7	6010.2	6902.7	1528.8	485.9	485.9
KD_18	5613.7	5684.2	10087.9	3675.8	349.9	349.9
KD_19	1201.0	937.6	885.5	3482.2	7108.6	885.5
KD_20	1808.0	2799.2	387.3	2102.8	6928.9	387.3

Dari tabel 9 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 9 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 = C_0	111.3	117.7	132.7	123.0	94.7
-----------------------	-------	-------	-------	-------	------

Pusat cluster 1 = C₁ 92.5 67.5 119 134.5 125.5

Pusat cluster 2 = C₂ 92.5 123.7 97.0 126.8 125.8

Pusat cluster 3 = C₃ 107.3 113.0 108.0 89.5 85.5

Pusat cluster 4 = C₄ 93.3 68.7 118.7 73.0 85.0

Tabel 3.10. Hasil Perhitungan Pusat Cluster Baru Iterasi 7

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	591.0	2686.3	1961.7	3521.2	6649.2	591.0
KD_2	3167.7	3051.3	368.0	3715.8	7295.2	368.0
KD_3	2993.3	5114.3	289.7	2863.8	8147.6	289.7
KD_4	2363.3	5907.3	1575.0	823.8	5000.2	823.8
KD_5	4467.3	9672.3	5525.7	937.2	4047.6	937.2
KD_6	3478.0	458.3	4938.0	6768.2	6242.2	458.3
KD_7	1728.7	7846.3	4508.0	433.8	3532.2	433.8
KD_8	6088.7	5690.3	9032.3	3031.2	330.9	330.9
KD_9	1704.7	7196.3	5006.0	766.2	2894.2	766.2
KD_10	4074.0	5930.3	3388.0	871.8	2394.2	871.8
KD_11	366.0	4030.3	2282.7	738.2	3452.2	366.0
KD_12	2949.0	5517.3	386.3	4506.8	10489.2	386.3
KD_13	5011.3	458.3	3373.7	5780.8	5522.2	458.3
KD_14	4107.3	4500.3	359.7	5305.8	10806.2	359.7
KD_15	1578.0	3249.3	4433.3	1064.2	1144.2	1064.2
KD_16	747.7	6366.3	4361.7	2958.8	7192.6	747.7
KD_17	5053.7	6411.3	6731.7	1528.8	485.9	485.9
KD_18	5613.7	5336.3	9581.3	3675.8	349.9	349.9
KD_19	1201.0	2098.3	618.7	3482.2	7108.6	618.7
KD_20	1808.0	4320.3	393.7	2102.8	6928.9	393.7

Dari tabel 10 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui dengan cara menggunakan rumus rata-rata pada data awal di tabel 10 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 = C_0	111.3	117.7	132.7	123.0	94.7
Pusat cluster 1 = C_1	92.5	67.5	119	134.5	125.5
Pusat cluster 2 = C_2	92.5	123.7	97	126.8	125.83
Pusat cluster 3 = C_3	107.3	113	108	89.5	85.5
Pusat cluster 4 = C_4	93.33	68.67	118.7	73	85

Tabel 3.11. Hasil Perhitungan Pusat Cluster Baru Iterasi 8

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	591.0	2686.3	1961.7	3521.2	6649.2	591.0
KD_2	3167.7	3051.3	368.0	3715.8	7295.2	368.0
KD_3	2993.3	5114.3	289.7	2863.8	8147.6	289.7
KD_4	2363.3	5907.3	1575.0	823.8	5000.2	823.8
KD_5	4467.3	9672.3	5525.7	937.2	4047.6	937.2
KD_6	3478.0	458.3	4938.0	6768.2	6242.2	458.3
KD_7	1728.7	7846.3	4508.0	433.8	3532.2	433.8
KD_8	6088.7	5690.3	9032.3	3031.2	330.9	330.9
KD_9	1704.7	7196.3	5006.0	766.2	2894.2	766.2
KD_10	4074.0	5930.3	3388.0	871.8	2394.2	871.8
KD_11	366.0	4030.3	2282.7	738.2	3452.2	366.0
KD_12	2949.0	5517.3	386.3	4506.8	10489.2	386.3
KD_13	5011.3	458.3	3373.7	5780.8	5522.2	458.3
KD_14	4107.3	4500.3	359.7	5305.8	10806.2	359.7
KD_15	1578.0	3249.3	4433.3	1064.2	1144.2	1064.2
KD_16	747.7	6366.3	4361.7	2958.8	7192.6	747.7
KD_17	5053.7	6411.3	6731.7	1528.8	485.9	485.9
KD_18	5613.7	5336.3	9581.3	3675.8	349.9	349.9
KD_19	1201.0	2098.3	618.7	3482.2	7108.6	618.7
KD_20	1808.0	4320.3	393.7	2102.8	6928.9	393.7

Dari tabel 11 diatas dapat dilihat alokasi nilai data ke centroid terdekat dengan menentukan nilai minimum dari masing-masing baris. Setelah diketahui nilai ke centroid terdekat, maka nilai centroid baru dapat diketahui

dengan cara menggunakan rumus rata-rata pada data awal di tabel 11 menggunakan nilai centroid terdekat di dapat nilai pusat cluster baru.

Pusat cluster 0 = C_0 111.3 117.7 132.7 123.0 94.7

Pusat cluster 1 = C_1 92.5 67.5 119 134.5 125.5

Pusat cluster 2 = C_2 92.5 123.7 97 126.8 125.83

Pusat cluster 3 = C_3 107.3 113 108 89.5 85.5

Pusat cluster 4 = C_4 93.33 68.67 118.7 73 85

Tabel 3.12. Hasil Perhitungan Pusat Cluster Baru Iterasi 9

KODE	C_0	C_1	C_2	C_3	C_4	Min
KD_1	591.0	2686.3	1961.7	3521.2	6649.2	591.0
KD_2	3167.7	3051.3	368.0	3715.8	7295.2	368.0
KD_3	2993.3	5114.3	289.7	2863.8	8147.6	289.7
KD_4	2363.3	5907.3	1575.0	823.8	5000.2	823.8
KD_5	4467.3	9672.3	5525.7	937.2	4047.6	937.2
KD_6	3478.0	458.3	4938.0	6768.2	6242.2	458.3
KD_7	1728.7	7846.3	4508.0	433.8	3532.2	433.8
KD_8	6088.7	5690.3	9032.3	3031.2	330.9	330.9
KD_9	1704.7	7196.3	5006.0	766.2	2894.2	766.2
KD_10	4074.0	5930.3	3388.0	871.8	2394.2	871.8
KD_11	366.0	4030.3	2282.7	738.2	3452.2	366.0
KD_12	2949.0	5517.3	386.3	4506.8	10489.2	386.3
KD_13	5011.3	458.3	3373.7	5780.8	5522.2	458.3
KD_14	4107.3	4500.3	359.7	5305.8	10806.2	359.7
KD_15	1578.0	3249.3	4433.3	1064.2	1144.2	1064.2
KD_16	747.7	6366.3	4361.7	2958.8	7192.6	747.7
KD_17	5053.7	6411.3	6731.7	1528.8	485.9	485.9
KD_18	5613.7	5336.3	9581.3	3675.8	349.9	349.9
KD_19	1201.0	2098.3	618.7	3482.2	7108.6	618.7
KD_20	1808.0	4320.3	393.7	2102.8	6928.9	393.7

Mengingat langkah ini telah mendapatkan cluster tetap yang konsisten maka iterasi dihentikan. Dari hasil iterasi ke 9 cluster terbagi menjadi 5 bagian.. berikut ini merupakan data hasil analisa perhitungan K-means cluster sehingga dapat diambil kesimpulan bahwa PT. Cipta Mebelindo Lestari sebaiknya lebih banyak menjual barang mebel perlengkapan kantor. berikut ini merupakan cluster data berdasarkan nama barang dan kode barang.

Tabel 3.13. Hasil Cluster Data Penjualan Barang

Nama cluster	Kode	Nama barang
Cluster 0	KD_1	Lemari Pakaian Mewah Racoco Flower
	KD_11	Meja UNO Modern Series
	KD_16	Kitchen Set Mutiara S Series
Cluster 1	KD_6	Ranjang Besi Levander + Kasur
	KD_13	PARTISI 4 STAFF VARIASI G
Cluster 2	KD_2	AIRLAND 505 ESSENTIAL 27 cm (Medium firm)
	KD_3	INOAC YUKATA (BEST SELLER)
	KD_12	Meja Rapat Euro (kaki besi)
	KD_14	SOFA Kennedy 321 tanpa meja
	KD_19	SD 1512 TECIDO
	KD_20	Lunar SR 9390
Cluster 3	KD_5	Meja Rias Volga - Kaca Sliding
	KD_6	Ranjang Besi Levander + Kasur
	KD_7	VIP FILING CABINET V-302
	KD_9	DIRBT-MDT
	KD_10	Rak Buku MPR 199
Cluster 4	KD_8	SOFA BED TURTLE
	KD_17	MEJA MAKAN GRANIT 4 KURSI
	KD_18	Kursi Taman Celebrity 08

3.3 Analisis Kebutuhan Perangkat

Analisis kebutuhan perangkat keras bertujuan untuk mengetahui secara tepat perangkat keras yang dibutuhkan. Adapun perangkat keras (*hardware*) yang dibutuhkan untuk pembuatan dan penerapan aplikasi yaitu:

3.3.1 Analisis Kebutuhan Perangkat Keras

Spesifikasi perangkat keras yang digunakan untuk pembuatan aplikasi yaitu:

- a. Komputer dengan processor minimal Intel Pentium IV2,3Ghz.
- b. Memory Minimal 2Gbyte
- c. Harddisk dari 100 GB ke atas
- d. Monitor SVGA.
- e. Printer dengan inkjet.
- f. Mouse dan Keyboard.

3.3.2 Analisis Kebutuhan Perangkat Lunak

Spesifikasi perangkat lunak yang digunakan untuk pembuatan aplikasi ini yaitu:

- 1) Sistem operasi windows 7/windows 8/windows 10
- 2) *RapidMiner Studio Free 7.3.0*

BAB IV

IMPLEMENTASI SISTEM

4.1 Implementasi Aplikasi

Pada bab ini akan dibahas tentang aplikasi data mining dengan K-Means Cluster untuk memprediksi produk potensial dan penentuan persediaan produk. Implementasi ini dilakukan hingga analisa hasil menggunakan aplikasi *RapidMiner*. Implementasi yang terjadi pada penggunaan aplikasi ini adalah sebagai berikut:

- a. Proses instalasi aplikasi pada komputer.
- b. Proses retrieve data analisa data penjualan.
- c. Menampilkan hasil analisa di aplikasi *RapidMiner*.

4.2 Pengujian Aplikasi

Pengujian aplikasi ini dijalankan sistem operasi *Windows 7*. Adapun spesifikasi yang dibutuhkan guna menjalankan aplikasi data mining dengan K-Means Cluster untuk memprediksi produk potensial dan penentuan persediaan produk adalah sebagai berikut:

4.2.1 Perangkat Keras (*Hardware*)

Perangkat keras yang dibutuhkan guna menjalankan aplikasi ini antara lain adalah:

1. *Processor* pentium IV 2,0 Ghz

2. *Memory* 2 GB
3. *Harddisk* 500 GB
4. *Monitor/ LCD*
5. *Printer*

4.2.2 Perangkat Lunak (*Software*)

Perangkat Lunak yang dibutuhkan guna menjalankan aplikasi ini antara lain adalah *RapidMiner Studio Free 7.3.0*

4.3 Instalasi Aplikasi

Proses instalasi aplikasi *RapidMiner Studio* dilakukan dengan beberapa langkah. Berikut ini merupakan proses instalasi *RapidMiner Studio*:

- a. Klik dua kali file yang diunduh (misalnya, *rapidminer-studio-<version>-win64-install.exe*).

Gambar 4.1. Instal aplikasi RapidMiner Studio



- b. Jika diminta, izinkan program membuat perubahan pada komputer. Wizard Pengaturan Studio RapidMiner muncul. Klik Berikutnya untuk melanjutkan.
- c. Baca ketentuan perjanjian lisensi dan klik Saya Setuju untuk melanjutkan.

- d. Pilih folder tujuan (atau biarkan default). Pastikan jalur folder tidak berisi +atau %karakter. Dengan mengklik Instal , wisaya mengekstrak dan menginstal RapidMiner Studio. Ketika instalasi selesai, klik Berikutnya dan kemudian Selesai untuk menutup wizard dan memulai RapidMiner Studio.
- e. Baca ketentuan perjanjian lisensi dan klik Saya Setuju untuk melanjutkan.

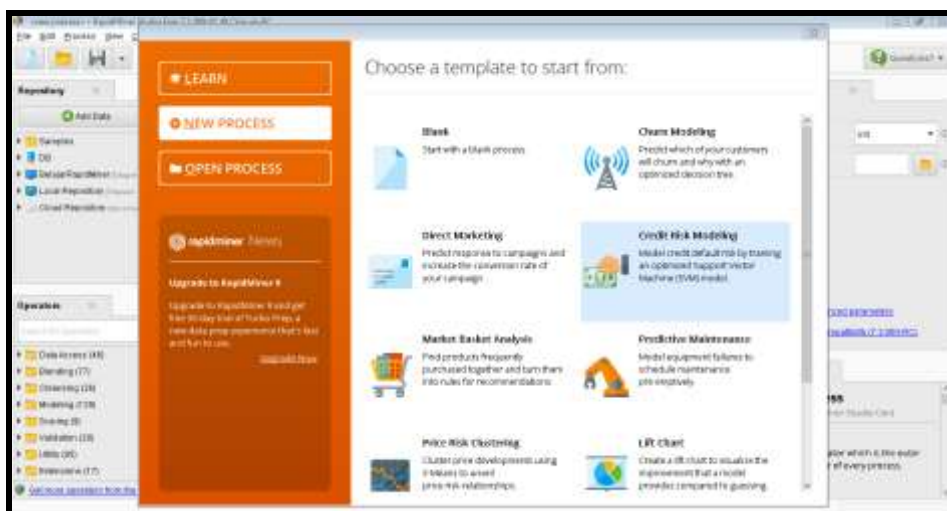
4.4 Pengujian Data Dengan Aplikasi RapidMiner

Pengujian data dilakukan dengan menggunakan aplikasi *RapidMiner Studio*. Berikut ini merupakan penjelasan langkah-langkah pengujian yang dilakukan dengan menggunakan aplikasi.

4.4.1 Tampilan Utama

Tampilan utama pada saat pertama sekali aplikasi *RapidMiner* dijalankan pada komputer dapat dilihat pada gambar berikut ini.

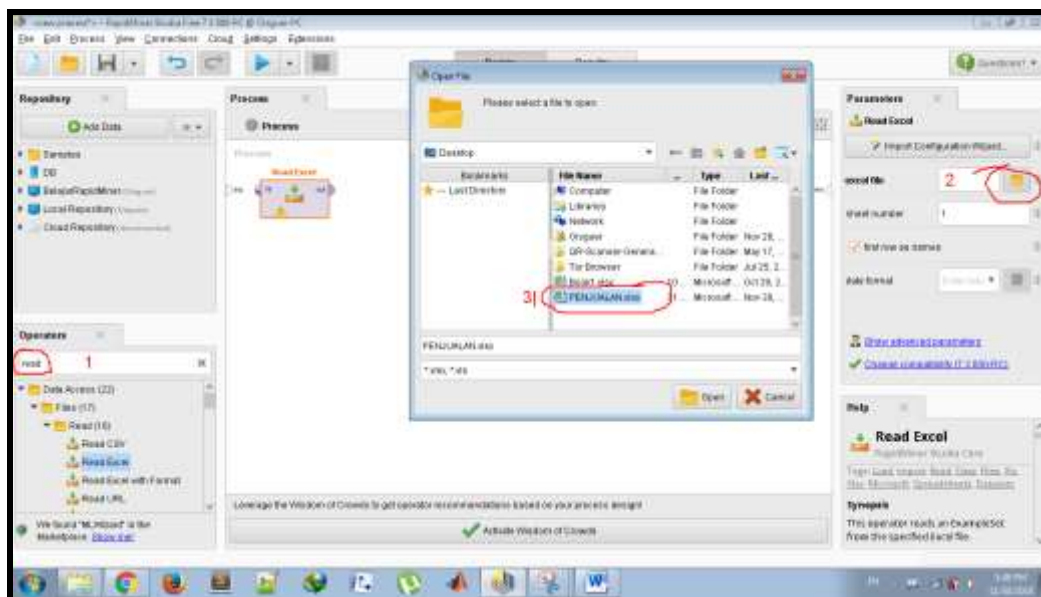
Gambar 4.2.Tampilan Awal *RapidMinerStudio 7.3*



4.4.2 Upload Data Penjualan

Untuk melakukan *upload* data penjualan yang akan dianalisa dengan cara ketikkan pada bagian operators “read” klik dan drag read excel pada aplikasi.

Gambar 4.3.Tampilan Awal *RapidMinerStudio 7.3*



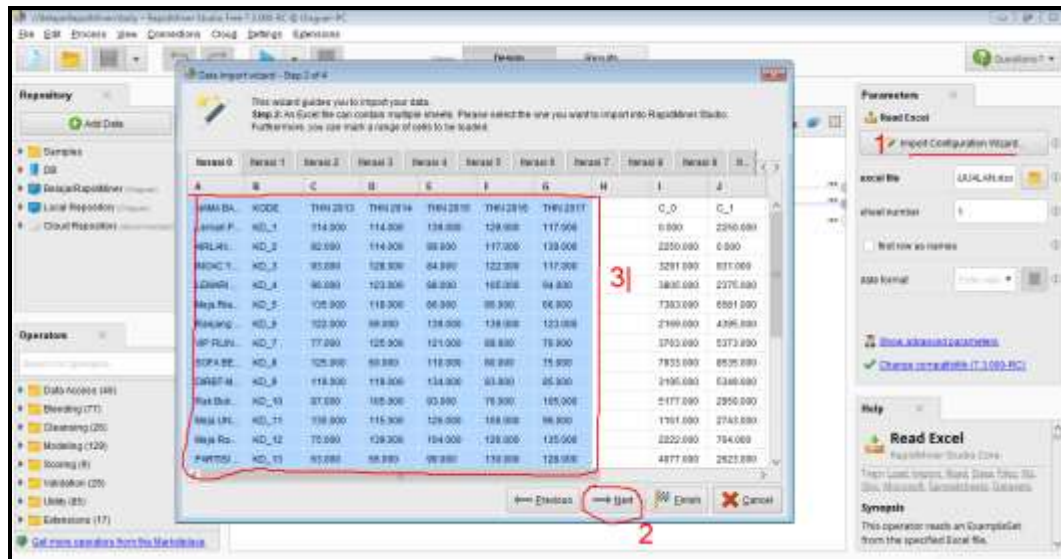
Keterangan:

- 1) Melakukan pencarian pada operator untuk membaca *file excel*.
- 2) Memilih data excel yang akan digunakan untuk pengujian proses perhitungan di aplikasi.
- 3) Setelah data dipilih, klik tombol open untuk load data excel yang akan digunakan di aplikasi.

4.4.3 *Import Configuration Wizard*

Langkah berikutnya adalah menentukan data yang akan digunakan untuk proses perhitungan pada aplikasi dengan menggunakan metode K-Means.

Gambar 4.4. Import Configuration Wizard



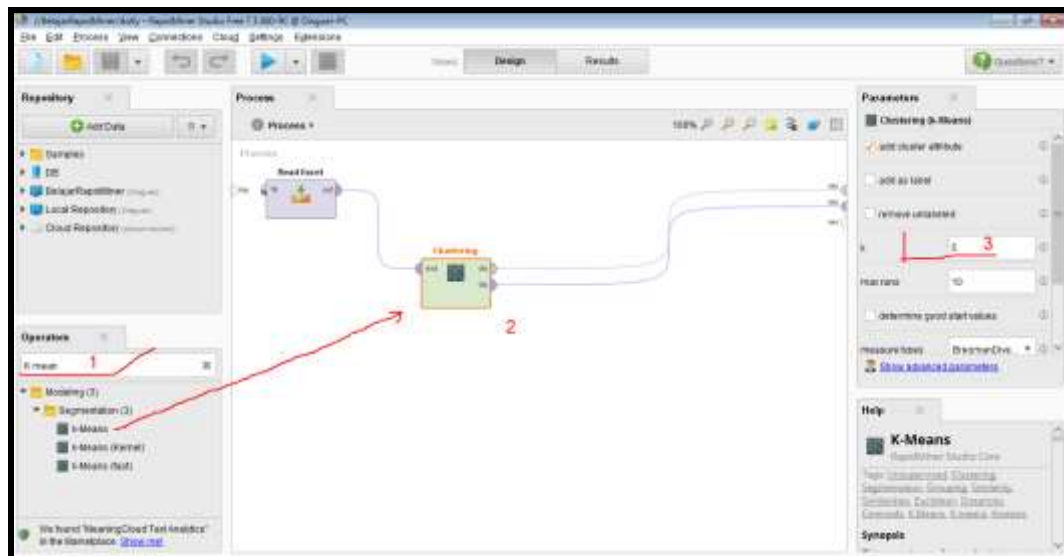
Keterangan:

- 1) Pilih *Import Configuration wizard* kemudian pilih data yang ingin diuji dalam bentuk file excel dilanjutkan dengan klik sheet yang akan dipilih datanya untuk dianalisa.\
- 2) klik next setelah data sheet yang akan digunakan untuk pengujian di aplikasi.
- 3) Blok data yang akan diuji diaplikasi. Pada atribut kode diubah menjadi ID pada aplikasi.

4.4.4 Proses Metode K-Means

Untuk melakukan proses *K-Means* pada aplikasi bisa dilakukan dengan pencarian operator dengan kata kunci K-mean seperti terlihat pada gambar 4 berikut ini.

Gambar 4.5. Proses Metode K-Means



Keterangan:

- 1) Melakukan proses pencarian operator k-means
- 2) Klik dan *drag* *K-Means* kedalam proses pada lembar kerja di aplikasi, lakukan hubungan data pada read excel ke *K-means Cluster*.
- 3) Tentukan jumlah cluster sesuai dengan data yang dibutuhkan

4.4.5 Hasil Perhitungan

Untuk melakukan proses perhitungan klik *Icon Play* berwarna biru pada menu bar di aplikasi untuk melakukan proses dan mendapatkan hasil perhitungan dengan metode *K-means cluster*. Berikut ini merupakan hasil perhitungan.

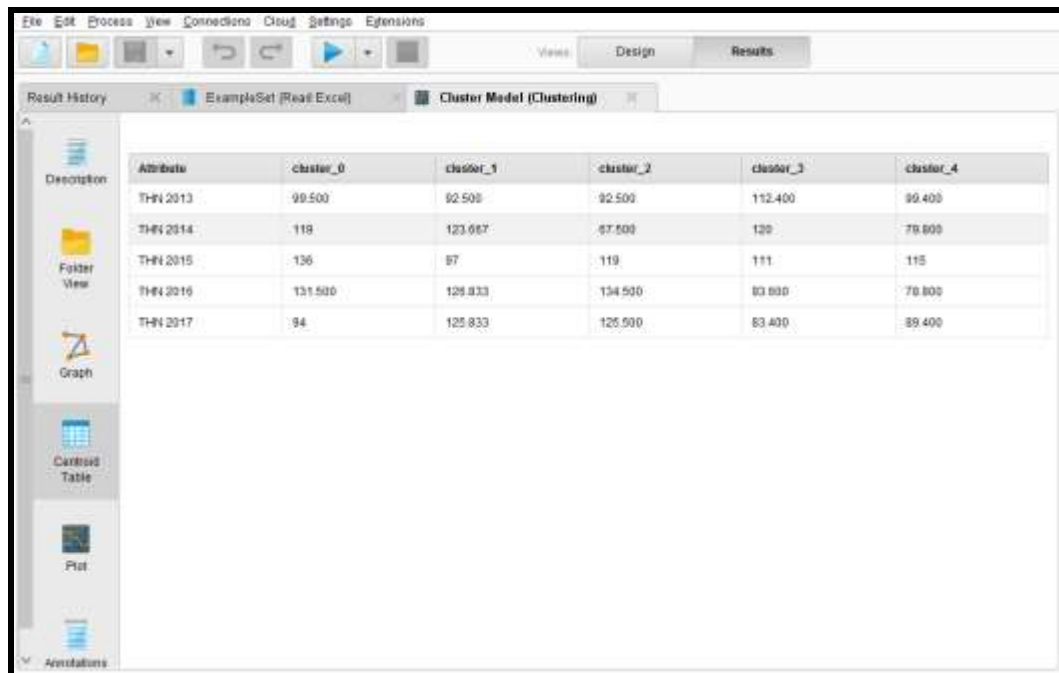
Gambar 4.6. Hasil Perhitungan Aplikasi

Item No.	Model Name	Code	Cluster	2013	2014	2015	2016	2017
1	Lamar Paka	KD_1	cluster_3	114	118	128	120	112
2	ARLAND 505	KD_2	cluster_1	92	114	99	117	128
3	INOAC YUKA	KD_3	cluster_1	63	126	84	102	117
4	LEMAR MIRAC	KD_4	cluster_3	96	122	86	105	94
5	Maja Rize Volg	KD_5	cluster_3	126	116	89	89	86
6	Ranjang Besi	KD_6	cluster_2	122	89	128	120	123
7	VP FILING CAB	KD_7	cluster_3	77	126	121	89	76
8	SOFA BED TURT	KD_8	cluster_4	125	83	118	83	75
9	DIREKT-MOT	KD_9	cluster_3	119	119	124	83	85
10	Rak Buku MPR	KD_10	cluster_4	67	105	92	74	102
11	Maja UNO Mod	KD_11	cluster_3	126	115	128	100	86
12	Maja Rajal Euro	KD_12	cluster_1	76	126	104	120	126
13	PARTISI 4 ST	KD_13	cluster_2	92	89	99	120	128
14	SOFA Kennedy	KD_14	cluster_1	107	124	84	120	123
15	Maja Tema B	KD_15	cluster_4	128	89	128	89	87
16	Kitchen Set M	KD_16	cluster_3	86	124	121	124	77
17	MEJA BUKU	KD_17	cluster_4	76	85	111	84	82

Gambar 4.7. Hasil Cluster Aplikasi

- root
 - cluster_0
 - Lamar Pakaian Mewah Racoco Flower
 - Kitchen Set Mutara S Series
 - cluster_1
 - ARLAND 505 ESSENTIAL 27 cm (Medium firm)
 - INOAC YUKATA (BEST SELLER)
 - Maja Rajal Euro (kaki besi)
 - SOFA Kennedy 321 tanpa meja
 - SD 1512 TECIDO
 - Lunar SR 9390
 - cluster_2
 - Ranjang Besi Levandar + Kasur
 - PARTISI 4 STAFF VARIASI G
 - cluster_3
 - LEMAR MIRACLE 2 pintu XL
 - Maja Rize Volga - Kaca Sliding
 - VP FILING CABINET V-302
 - DIREKT-MOT
 - Maja UNO Modern Series
 - cluster_4
 - SOFA BED TURTLE
 - Rak Buku MPR 199

Gambar 4.8. Pusat Centroid Yang Digunakan Di Aplikasi



The screenshot shows the 'Results' tab in RapidMiner. The main window displays a table with the following data:

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
THN 2013	99.500	92.500	92.500	112.400	90.400
THN 2014	118	123.667	67.500	120	70.800
THN 2015	136	87	119	111	115
THN 2016	131.500	128.833	134.500	80.800	70.800
THN 2017	94	125.833	125.500	83.400	89.400

Berdasarkan hasil pengujian pada aplikasi maka dapat diketahui kelebihan dan kekurangan aplikasi. Adapun kelebihan dari aplikasi diantaranya adalah:

- a. Aplikasi memberikan hasil perhitungan yang lebih akurat.
- b. Hasil perhitungan aplikasi ditampilkan dalam bentuk tabel, graph dan clustering data.

Sedangkan untuk kekurangan pada aplikasi *RapidMiner studio* diantaranya adalah:

- a. Aplikasi tidak menampilkan langkah-langkah perhitungan dari proses metode K-Means Cluster.
- b. Hasil perhitungan sulit dipahami bagi orang awam yang belum biasa menggunakan aplikasi ini.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil uraian analisa dan pengujian perancangan pada bab sebelumnya, kesimpulan yang dapat diambil adalah sebagai berikut:

- a. Hasil dari perhitungan pada aplikasi dapat memberikan pengetahuan baru bagi perusahaan yang selama ini belum dapat memastikan apa saja barang yang banyak diminati oleh konsumen pada umumnya.
- b. Model perhitungan dengan algoritma K-means cluster pada bab 3 dan pada pemodelan perhitungan di aplikasi mengalami selisih data pada proses perhitungan manual dan pada sistem.
- c. Menerapkan algoritma K-means clustering diawali dengan melakukan asumsi untuk tiap-tiap variabel pada data penjualan yang menjadi tolak ukur sebagai penyelesaian masalah, kemudian menormalisasikan setiap data penjualan menggunakan asumsi, selanjutnya menentukan titik pusat cluster awal untuk melakukan iterasi dan cluster, melakukan perulangan sehingga cluster barang yang terbentuk tidak berubah polanya dan dilakukan pembentukan kesimpulan pengetahuan baru bagi perusahaan dalam bentuk tabel dengan cara mengurutkan data barang yang paling diminati oleh para konsumen .

5.2 Saran

Saran yang dibutuhkan untuk pengembangan lebih lanjut terhadap aplikasi ini antara lain adalah:

- a. Bagi mahasiswa diharapkan agar Algoritma K-means dapat diterapkan pada pemodelan lain selain melakukan pengelompokan data penjualan.
- b. Bagi pembacayang bersedia dan mampu menyempurnakan pemodelan perhitungan disarankan untuk menggunakan model perhitungan lain pada bidang ilmu data mining sehingga bisa dilakukan perbandingan hasil perhitungannya.
- c. Bagi pembacayang bersedia dan mampu untuk menyempurnakan hendaknya dilakukan pengembangan dengan menggunakan aplikasi selain *RapidMiner Studio*.

DAFTAR PUSTAKA

335, 2014.

Andrian, Yudhi, and Purwa Hasan Putra. "Analisis Penambahan Momentum Pada Proses Prediksi Curah Hujan Kota Medan Menggunakan Metode Backpropagation Neural Network." Seminar Nasional Informatika (SNIf). Vol. 1. No. 1. 2017.

Anonim, E. H. Rachmawanto and C. A. Sari, "Keamanan File Menggunakan Teknik Kriptografi Shift Cipher," Jurnal Techno. Com, vol. 14, no. 2, pp. 329-

Aryza, S., Irwanto, M., Lubis, Z., Siahaan, A. P. U., Rahim, R., & Furqan, M. (2018). A Novelty Design Of Minimization Of Electrical Losses In A Vector Controlled Induction Machine Drive. In IOP Conference Series: Materials Science and Engineering (Vol. 300, No. 1, p. 012067). IOP Publishing.

Bishop, Rosdiana, "Sekuritas Sistem Dengan Kriptografi," in Prosiding Sendi_U 2013, Semarang, 2013.

Fresly, Faizal Zuli1, Ari Irawan, "Implementasi Kriptografi Dengan Algoritma Blowfish dan Riverst Shamir Adleman (RSA) Untuk Proteksi File," Jurnal Format Volume 6 nomor 2 Tahun 2016.

- Gede Angga Pradipta " Penerepan Kombinasi metode Enkripsi Vigenere Cipher Dan Trasposisi Pada Aplikasi Client Server Chatting, " Jurnal Sistem Dan Informatika Vol. 10, Nomor 2, 2016.
- Ginting, G., Fadlina, M., Siahaan, A. P. U., & Rahim, R. (2017). Technical approach of TOPSIS in decision making. *Int. J. Recent Trends Eng. Res*, 3(8), 58-64.
- Hafni, Layla, And Rismawati Rismawati. "Analisis Faktor-Faktor Internal Yang Mempengaruhi Nilai Perusahaan Pada Perusahaan Manufaktur Yang Terdaftar Di Bei 2011-2015." *Bilancia: Jurnal Ilmiah Akuntansi* 1.3 (2017): 371-382.
- Hamdi, Muhammad Nurul, Evi Nurjanah, And Latifah Safitri Handayani. "Community Development Based On Ibnu Khaldun Thought, Sebuah Interpretasi Program Pemberdayaan Umkm Di Bank Zakat El-Zawa." *El Muhasaba: Jurnal Akuntansi (E-Journal)* 5.2 (2014): 158-180.
- Indra Permana, Aminuddin "Sistem Pakar Mendeteksi Hama Dan Penyakit Tanaman Kelapa Sawit Pada Pt. Moeis Kebun Sipare-Pare Kabupaten Batubara." (2013).
- Mayasari, Nova. "Comparison of Support Vector Machine and Decision Tree in Predicting On-Time Graduation (Case Study: Universitas Pembangunan Panca Budi)." *Int. J. Recent Trends Eng. Res* 2.12 (2016): 140-151.
- Nandar Pabokory, Indah Fitri Astuti, Awang Harsa Kridalaksana, " Implementasi Kriptografi Pengamanan Data Pada Pesan Teks, Isi File Dokumen, Dan File Dokumen Menggunakan Algoritma Advanced Encryption Standard," *Jurnal Informatika Mulawarman* Vol. 10. Nomor 1, 2015.
- Permana, A. I., and Z. Tulus. "Combination of One Time Pad Cryptography Algorithm with Generate Random Keys and Vigenere Cipher with EM2B KEY." (2020).

- Permana, Aminuddin Indra. "Kombinasi Algoritma Kriptografi One Time Pad dengan Generate Random Keys dan Vigenere Cipher dengan Kunci EM2B." (2019).
- Puspita, Khairani, and Purwa Hasan Putra. "Penerapan Metode Simple Additive Weighting (SAW) Dalam Menentukan Pendirian Lokasi Gamedia Di Sumatera Utara." Seminar Nasional Teknologi Informasi Dan Multimedia, ISSN. 2015.
- Putera, A., Siahaan, U., & Rahim, R. (2016). Dynamic key matrix of hill cipher using genetic algorithm. *Int. J. Secur. Its Appl*, 10(8), 173-180.
- Ramadhan, A., & Mohd. Awal Hakimi. (2006). *Pemrograman Web Database dengan PHP dan MySQL*. Synergy Media.
- Ramadhan, M., & Nugroho, N. B. (2009). Desain web dengan php. *Jurnal Saintikom*, 6(1).
- Renddy, Teady Matius, Surya Mulyana, Fresly, " Steganografi Dengan Deret Untuk Mengacak Pola Penempatan Pada Rgb," *Jurnal Teknologi Informasi*, 2015.
- Rhee, C. A. Sari, E. H. Rachmawanto, Y. P. Astuti and L. Umaroh, "Optimasi Penyandian File Kriptografi Shift Cipher," in *Prosiding Sendi_U 2013*, Semarang, 2013.
- Suriski Sitinjak, Yuli Fauziah, Juwairiah, " Aplikasi Kriptografi File Menggunakan Algoritma Blowfish," *Jurnal Informatika Mulawarman Vol. 10. Nomor 1*, 2015.
- Syahputra, Rizki, And Hafni Hafni. "Analisis Kinerja Jaringan Switching Clos Tanpa Buffer." *Journal Of Science And Social Research* 1.2 (2018): 109-115.
- Wahyuni, Sri. "Implementasi Rapidminer Dalam Menganalisa Data Mahasiswa Drop Out." *Jurnal Abdi Ilmu* 10.2 (2018): 1899-1902.