

Data Mining Pengelompokan Data Pada PT. Gatra Kasindo Jaya Menggunakan Algoritma K-Maens

Bagas Anjas Wara ^{#1}, Yohanni Syahra, S.Si., M.Kom. ^{#2}, Milfa Yetri, S.Kom., M.Kom^{#3}

^{#1} Program Studi Sistem Informasi, STMIK Triguna Dharma

^{#2,3} Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received xxxx xxth, 2020

Revised xxxx xxth, 2020

Accepted xxxx xxth, 2020

Keyword:

Data Mining

K - Maens

Trend Penjualan

ABSTRAK

Daya tarik yang ditawarkan oleh PT. Gatra Kasindo Jaya pada umumnya berkaitan dengan kualitas layanan. Di mana kualitas layanan yang dimaksud dijabarkan pada tangibles, reliability, responssiveness, assurance, dan emphaty yang dalam penerapannya berbeda antara satu perusahaan dengan perusahaan lainnya. Dalam tren penjualan dibutuhkan analisis tren penjualan jasa pada PT. Gatra Kasindo Jaya. Berdasarkan hal tersebut pengelompokan tren penjualan dapat diterapkan menggunakan konsep keilmuan Data Mining.

Penerapan Data Mining telah banyak digunakan dalam pengolahan data untuk menghasilkan pengetahuan, salah satunya pada yang menggunakan Data Mining untuk memprediksi pencapaian yang target penjualan pada tren penjualan. Algoritma yang akan digunakan untuk pengelolaan Data Mining pada kasus mengestimasi target penjualan di PT Gatra Kasindo Jaya Cab Medan menggunakan metode K - Maens

Penerapan metode regresi linear berganda telah digunakan dalam berbagai kasus mengestimasi data seperti pada digunakannya regresi linear berganda untuk mengetahui pencapaian penjualan berdasarkan tren penjualan. Selain itu dalam untuk mengetahui target penjualan di wilayah Indoenisa. Sehingga Metode Regeresi Linear Berganda sangat sesuai dan dapat mengestimasi target penjualan jasa di PT. Gatra Kasindo Jaya .

Kata Kunci: Trend Pengelompokan, Data Mining, K - Maens

Copyright © 201x STMIK Triguna Dharma.
All rights reserved.

Nama : Bagas Anjas Wara
Kator : STMIK Triguna Dharma
Program Studi : Sistem Informasi
Email : anjasbagas01@gmail.com

%1. PENDAHULUAN

PT. Gatra Kasindo Jaya merupakan perusahaan minuman di kota medan, yang harus mempertimbangkan kepuasan bagi pelanggan, maka peneliti ingin mengetahui hal apa yang harus dilakukan oleh perusahaan dan hal apa saja yang di harapkan oleh pelanggan, sehingga PT. Gatra Kasindo Jaya dapat meningkatkan kinerjanya sebaik mungkin. Daya tarik yang ditawarkan oleh PT. Gatra Kasindo Jaya pada umumnya berkaitan dengan kualitas layanan. Di mana kualitas layanan yang dimaksud dijabarkan pada *tangibles, reliability, responsiveness, assurance*, dan *emphaty* yang dalam penerapannya berbeda antara satu perusahaan dengan perusahaan lainnya. Dalam tren penjualan jasa Gatra Kasindo Jaya dibutuhkan analisis tren penjualan jasa pada PT. Gatra Kasindo Jaya . Hal tersebut tentu akan sangat membangun dan membantu bagi pihak perusahaan PT. Gatra Kasindo Jaya dalam mengestimasi pencapaian target penjualan.

Berdasarkan hal tersebut pengelompokan tren penjualan dapat diterapkan menggunakan konsep keilmuan *Data Mining*. Penerapan *Data Mining* telah banyak digunakan dalam pengolahan data untuk menghasilkan pengetahuan, salah satunya pada yang menggunakan Data Mining untuk memprediksi pencapaian yang target penjualan pada tren penjualan. Selanjutnya algoritma yang akan digunakan untuk pengelolaan *Data Mining* pada kasus mengestimasi target penjualan jasa di PT Gatra Kasindo Jaya Cab Medan menggunakan metode K - Maens [1].

Model *K-Maens* merupakan perluasan dari model *K-Maens* sederhana. Dengan memperluas model *K-Maens* dua atau tiga cluster, maka model *K-Maens* dengan cluster terikat Y dan k variable bebas $X_1, X_2, X_3, \dots, X_k$.

Penerapan metode regresi linear berganda telah digunakan dalam berbagai kasus mengestimasi data seperti pada [2] digunakannya *K-Maens* berganda untuk mengetahui pencapaian penjualan berdasarkan tren penjualan. Selain itu dalam [3] untuk mengetahui target penjualan di wilayah Indoensia. Sehingga Metode *K-Maens* sangat sesuai dan dapat mengestimasi target penjualan jasa di PT. Gatra Kasindo Jaya .

%1. **Kajian Pustaka**

%1.%2. **Data Mining**

Data Mining adalah sebuah proses menganalisa data untuk mengetahui suatu pola dari kelompok data yang tersembunyi. yang bertujuan untuk menemukan, menggali pengetahuan dari data atau informasi yang kita miliki. Data Mining berisi pencarian pola yang diinginkan pada database untuk membantu pengambil keputusan di waktu yang akan datang.

Data Mining merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar. yang tujuan utama Data Mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki[4] .

Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, data mining dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan field dari sebuah relasional database yang besar[5].

%1.%2. **K - Maens**

Metode *K - Maens* merupakan teknik analisis yang mencoba menjelaskan hubungan antara dua perubah atau lebih khususnya antara peubah- peubah yang mengandung sebab akibat yang disebut regresi linier[12]. *K - Maens* Juga merupakan sebuah analisis yang memiliki variabel yang bebas dan lebih dari satu. Analisis *K - Maens* telah lama dikembangkan untuk mempelajari pola dan mengukur hubungan statistik antara dua atau lebih perubah (*variabel*). Teknik analisis yang mencoba menjelaskan hubungan antara dua peubah atau lebih khususnya antara peubah- peubah yang mengandung sebab dan akibat disebut analisis *regresi linier*.

Dalam analisis *regresi linier* berganda terdapat beberapa uji asumsi klasik, yaitu uji heteroskedastisitas, ujitersebut bertujuan untuk menguji apakah dalam sebuah model *regresi* terjadi ketidaksamaan varians residual dari suatu pengamatan ke pengamatan yang lain tetap. Kemudian Uji Normalitas, yang bertujuan untuk menguji apakah dalam suatu model *regresi*, variabel terikat dan variabel bebas atau keduanya mempunyai distribusi normal[12]. Teknik *K - Maens* digunakan untuk mengetahui ada tidaknya pengaruh signifikan dua atau lebih variabel bebas ($X_1, X_2, X_3, \dots, X_k$) terhadap variabel terikat atau tidak bebas (Y). Model *K - Maens* untuk populasi dapat ditunjukkan sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Model *K - Maens* untuk populasi diatas dapat ditaksir dengan model *K - Maens* untuk sampel, yaitu :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$$

dengan :

Y = nilai penduga bagi variabel Y

b0=dugaan bagi parameter konstanta

b1, b2, ... bk = dugaan bagi parameter konstanta $\beta_1, \beta_2, \dots \beta_n$

X = Variabel bebas [13].

%1. Metodologi Penelitian

%1.%2. Algoritma Sistem

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan *Data mining* dalam memprediksikan jumlah kebutuhan target trend penjualan ke malaysia berdasarkan variabel dengan menggunakan metode *Regresi Linear Bergandan*. Hal ini dilakukan untuk meningkat kinerja pada PT. Gatra Kasindo Jaya .

Tahap ini dilakukan penerapan algoritma *k-means* dengan rumus :

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

Penetapan jumlah *cluster* (K) yaitu 3 *cluster*.

Setelah menetapkan jumlah *cluster*, tentukan titik pusat awal *cluster* (*centroid*). Berikut ini titik *centroid* yang telah dipilih :

Tabel 3.1 Tabel Titik Pusat Awal *Cluster*

No	Centroid	Paket		
		Eksklusif	Kombinas i	Ekonomi
1.	m1	25	21	20
2.	m2	15	16	17
3.	m3	15	19	15

%1. Menghitung jarak antara variabel dari setiap sampel data dengan *centroid* nya.

%1. Jarak antara bulan Januari dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(25 - 25)^2 + (21 - 21)^2 + (20 - 20)^2}$$

$$= 0$$

%1. Jarak antara bulan Februari dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(26 - 25)^2 + (22 - 21)^2 + (21 - 20)^2}$$

$$= 1.73$$

%1. Jarak antara bulan Maret dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(20 - 25)^2 + (21 - 21)^2 + (25 - 20)^2}$$

$$= 7.07$$

%1. Jarak antara bulan Januari dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(25 - 15)^2 + (21 - 16)^2 + (20 - 17)^2}$$

$$= 11.58$$

%1. Jarak antara bulan Februari dengan titik m2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(26 - 15)^2 + (22 - 16)^2 + (21 - 17)^2} \\
 &= 13.15
 \end{aligned}$$

%1. Jarak antara bulan Maret dengan titik m2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(20 - 15)^2 + (21 - 16)^2 + (25 - 17)^2} \\
 &= 10.68
 \end{aligned}$$

%1. Jarak antara bulan Januari dengan titik m3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(25 - 15)^2 + (21 - 19)^2 + (20 - 15)^2} \\
 &= 11.36
 \end{aligned}$$

%1. Jarak antara bulan Februari dengan titik m3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(26 - 15)^2 + (22 - 19)^2 + (21 - 15)^2} \\
 &= 12.88
 \end{aligned}$$

%1. Jarak antara bulan Maret dengan titik m3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(20 - 15)^2 + (21 - 19)^2 + (25 - 15)^2} \\
 &= 11.36
 \end{aligned}$$

%1. Mencari jarak terdekat setiap *cluster*

Tabel 3.2 Tabel Hasil Perhitungan Iterasi Ke 1

No	Inisial	Jarak ke titik m1	Jarak ke titik m2	Jarak ke titik m3	Jarak terdekat	Keterangan
1.	Januari	0.00	11.58	11.36	0.00	Cluster 1
2.	Februari	1.73	13.15	12.88	1.73	Cluster 1
3.	Maret	7.07	10.68	11.36	7.07	Cluster 1
4.	April	3.32	9.00	9.27	3.32	Cluster 1
5.	Mei	8.12	6.78	4.36	4.36	Cluster 3
6.	Juni	8.54	4.12	5.48	4.12	Cluster 2
7.	Juli	9.17	2.45	4.12	2.45	Cluster 2
8.	Agustus	8.06	5.20	7.07	5.20	Cluster 2
9.	September	7.28	6.40	8.25	6.40	Cluster 2
10.	Oktober	10.05	3.00	6.00	3.00	Cluster 2
11.	November	10.44	4.12	2.45	2.45	Cluster 3
12.	Desember	9.11	3.00	4.69	3.00	Cluster 2

Setelah mengetahui jarak terdekat setiap cluster. Urutkan sesuai *clusternya* masing-masing.

Tabel 3.3 Tabel Data Yang Telah Diurutkan Sesuai *Cluster*

No	Inisial	m1	m2	m3	Jarak terdekat	Keterangan
1.	Januari	0.00	11.58	11.36	0.00	Cluster 1
2.	Februari	1.73	13.15	12.88	1.73	Cluster 1
3.	Maret	7.07	10.68	11.36	7.07	Cluster 1
4.	April	3.32	9.00	9.27	3.32	Cluster 1
5.	Mei	8.54	4.12	5.48	4.12	Cluster 2
6.	Juni	9.17	2.45	4.12	2.45	Cluster 2
7.	Juli	8.06	5.20	7.07	5.20	Cluster 2
8.	Agustus	7.28	6.40	8.25	6.40	Cluster 2
9.	September	10.05	3.00	6.00	3.00	Cluster 2
10.	Oktober	9.11	3.00	4.69	3.00	Cluster 2
11.	November	10.05	1.73	3.74	1.73	Cluster 2
12.	Desember	11.36	3.61	0.00	0.00	Cluster 3

%1. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 0^2 + 1.73^2 + 7.07^2 + \dots + 7.68^2$$

$$WCV = 301$$

%1. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} \%1. \quad d(m1, m2) &= \sqrt{(m1 - m2)^2} \\ &= \sqrt{(25 - 15)^2 + (21 - 16)^2 + (20 - 17)^2} \\ &= 11.58 \end{aligned}$$

$$\begin{aligned} \%1. \quad d(m1, m3) &= \sqrt{(m1 - m3)^2} \\ &= \sqrt{(25 - 15)^2 + (21 - 19)^2 + (20 - 15)^2} \\ &= 11.36 \end{aligned}$$

$$\begin{aligned} \%1. \quad d(m2, m3) &= \sqrt{(m2 - m3)^2} \\ &= \sqrt{(15 - 15)^2 + (16 - 19)^2 + (17 - 15)^2} \\ &= 3.61 \end{aligned}$$

$$\begin{aligned} \text{Nilai BCV} &= d(m1, m2) + d(m1, m3) + d(m2, m3) \\ &= 11.58 + 11.36 + 3.61 \\ &= 26.54 \end{aligned}$$

%1. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$\begin{aligned} BCV/WCV &= 26.54 / 301 \\ &= 0.09 \end{aligned}$$

Kemudian langkah berikutnya pada iterasi ke 2 adalah :

%1. Menghitung kembali pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk.

%1. Titik Pusat Awal *Cluster* 1 (m1) yaitu : 23.25, 21, 21.75

%1. Titik pusat awal *cluster* 2 (m2) yaitu :

$$\begin{aligned} &= \text{rata-rata (m2)} \\ &= (17.33, 16.08, 18) \end{aligned}$$

%1. Titik pusat awal *cluster* 3 (m3) yaitu :
= rata-rata (m3)

$$= (18, 19.25, 14.25)$$

%1. Menghitung kembali jarak antara variabel dari setiap sampel data dengan *centroid* nya.

%1. Jarak antara bulan Januari dengan titik m1

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(25 - 23.25)^2 + (21 - 21)^2 + (20 - 21.75)^2} \\ &= 2.47 \end{aligned}$$

%1. Jarak antara bulan Februari dengan titik m1

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(26 - 23.25)^2 + (22 - 21)^2 + (21 - 21.75)^2} \\ &= 3.02 \end{aligned}$$

%1. Jarak antara bulan Maret dengan titik m1

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(20 - 23.25)^2 + (21 - 21)^2 + (25 - 21.75)^2} \\ &= 4.60 \end{aligned}$$

%1. Jarak antara bulan Januari dengan titik m2

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(25 - 17.33)^2 + (21 - 16.08)^2 + (20 - 18)^2} \\ &= 9.32 \end{aligned}$$

%1. Jarak antara bulan Februari dengan titik m2

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(26 - 17.33)^2 + (22 - 16.08)^2 + (21 - 18)^2} \\ &= 10.91 \end{aligned}$$

%1. Jarak antara bulan Maret dengan titik m2

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(20 - 17.33)^2 + (21 - 16.08)^2 + (25 - 18)^2} \\ &= 8.96 \end{aligned}$$

%1. Jarak antara bulan Januari dengan titik m3

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(25 - 18)^2 + (21 - 19.25)^2 + (20 - 14.25)^2} \\ &= 9.23 \end{aligned}$$

%1. Jarak antara bulan Februari dengan titik m3

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(26 - 18)^2 + (22 - 19.25)^2 + (21 - 14.25)^2} \\ &= 10.82 \end{aligned}$$

%1. Jarak antara bulan Maret dengan titik m3

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= \sqrt{(20 - 18)^2 + (21 - 19.25)^2 + (25 - 14.25)^2} \\ &= 11.07 \end{aligned}$$

Tabel 3.4 Tabel Hasil Perhitungan Iterasi Ke 2

No	Inisial	Jarak ke titik m1	Jarak ke titik m2	Jarak ke titik m3	Jarak terdekat	Keterangan
1.	Januari	2.47	9.32	9.23	2.47	Cluster 1
2.	Februari	3.02	10.91	10.82	3.02	Cluster 1
3.	Maret	4.60	8.96	11.07	4.60	Cluster 1
4.	April	1.77	6.79	7.88	1.77	Cluster 1

Tabel 3.4 Tabel Hasil Perhitungan Iterasi Ke 2 (Lanjutan)

No	Inisial	Jarak ke titik m1	Jarak ke titik m2	Jarak ke titik m3	Jarak terdekat	Keterangan
5.	Mei	7.15	2.79	5.97	2.79	Cluster 2
6.	Juni	8.31	0.98	4.49	0.98	Cluster 2
7.	Juli	7.79	2.88	6.01	2.88	Cluster 2
8.	Agustus	6.97	3.95	7.04	3.95	Cluster 2
9.	September	9.09	1.51	6.45	1.51	Cluster 2
10.	Oktober	8.67	1.20	4.26	1.20	Cluster 2
11.	November	9.09	1.62	4.81	1.62	Cluster 2
12.	Desember	10.28	7.84	4.76	4.76	Cluster 3

%1. Hitung kembali nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 2.47^2 + 3.02^2 + 4.60^2 + \dots + 4.76^2$$

$$WCV = 160.58$$

%1. Hitung kembali nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\%1. d(m1, m2) = \sqrt{(23.25 - 17.33)^2 + (21 - 16.08)^2 + (21.75 - 18)^2} = 8.56$$

$$\%1. d(m1, m3) = \sqrt{(23.25 - 18)^2 + (21 - 19.25)^2 + (21.75 - 14.25)^2} = 9.32$$

$$\%1. d(m2, m3) = \sqrt{(17.33 - 18)^2 + (16.08 - 19.25)^2 + (18 - 14.25)^2} = 4.95$$

$$\text{Nilai BCV} = 8.56 + 9.32 + 4.95 = 22.83$$

%1. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV.

$$BCV/WCV = 22.83/160.58 = 0.14$$

Pada iterasi ke 3 langkah-langkah yang harus dilakukan adalah sebagai berikut:

%1. Menghitung kembali pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk.

%1. Titik Pusat Awal *Cluster* 1 (m1) yaitu : 23.25, 21, 21.75

%1. Titik pusat awal *cluster* 2 (m2) yaitu :

$$= \text{rata-rata (m2)} = (17.33, 16.08, 18)$$

%1. Titik pusat awal *cluster* 3 (m3) yaitu :

$$= \text{rata-rata (m3)} = (18, 19.25, 14.25)$$

%1. Menghitung kembali jarak antara variabel dari setiap sampel data dengan *centroid* nya.

%1. Jarak antara bulan Januari dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} = \sqrt{(25 - 23.25)^2 + (21 - 21)^2 + (20 - 21.75)^2} = 2.47$$

%1. Jarak antara bulan Februari dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} = \sqrt{(26 - 23.25)^2 + (22 - 21)^2 + (21 - 21.75)^2} = 3.02$$

%1. Jarak antara bulan Maret dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(20 - 23.25)^2 + (21 - 21)^2 + (25 - 21.75)^2}$$

$$= 4.60$$

%1. Jarak antara bulan Januari dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(25 - 17.33)^2 + (21 - 16.08)^2 + (20 - 18)^2}$$

$$= 9.32$$

%1. Jarak antara bulan Februari dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(26 - 17.33)^2 + (22 - 16.08)^2 + (21 - 18)^2}$$

$$= 10.91$$

%1. Jarak antara bulan Maret dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(20 - 17.33)^2 + (21 - 16.08)^2 + (25 - 18)^2}$$

$$= 8.96$$

%1. Jarak antara bulan Januari dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(25 - 18)^2 + (21 - 19.25)^2 + (20 - 14.25)^2}$$

$$= 9.23$$

%1. Jarak antara bulan Februari dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(26 - 18)^2 + (22 - 19.25)^2 + (21 - 14.25)^2}$$

$$= 10.82$$

%1. Jarak antara bulan Maret dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(20 - 18)^2 + (21 - 19.25)^2 + (25 - 14.25)^2}$$

$$= 11.07$$

Tabel 3.5 Tabel Hasil Perhitungan Iterasi Ke 3

No	Inisial	Jarak ke titik m1	Jarak ke titik m2	Jarak ke titik m3	Jarak terdekat	Keterangan
1.	Januari	2.47	9.32	9.23	2.47	Cluster 1
2.	Februari	3.02	10.91	10.82	3.02	Cluster 1
3.	Maret	4.60	8.96	11.07	4.60	Cluster 1
4.	April	1.77	6.79	7.88	1.77	Cluster 1
5.	Mei	7.15	2.79	5.97	2.79	Cluster 2
6.	Juni	8.31	0.98	4.49	0.98	Cluster 2
7.	Juli	7.79	2.88	6.01	2.88	Cluster 2
8.	Agustus	6.97	3.95	7.04	3.95	Cluster 2
9.	September	9.09	1.51	6.45	1.51	Cluster 2
10.	Oktober	8.67	1.20	4.26	1.20	Cluster 2
11.	November	9.09	1.62	4.81	1.62	Cluster 2

Tabel 3.5 Tabel Hasil Perhitungan Iterasi Ke 3 (Lanjutan)

No	Inisial	Jarak ke titik m1	Jarak ke titik m2	Jarak ke titik m3	Jarak terdekat	Keterangan
12.	Desember	10.28	7.84	4.76	4.76	Cluster 3

%1. Hitung kembali nilai *WCV* (*Within Cluster Variation*) dengan cara mengangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai *WCV*.

$$WCV = 2.47^2 + 3.02^2 + 4.60^2 + \dots + 4.76^2$$

$$WCV = 160.58$$

%1. Hitung kembali nilai *BCV* (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} \%1. \quad d(m1, m2) &= \sqrt{(23.25 - 17.33)^2 + (21 - 16.08)^2 + (21.75 - 18)^2} \\ &= 8.56 \end{aligned}$$

$$\begin{aligned} \%1. \quad d(m1, m3) &= \sqrt{(23.25 - 18)^2 + (21 - 19.25)^2 + (21.75 - 14.25)^2} \\ &= 9.32 \end{aligned}$$

$$\begin{aligned} \%1. \quad d(m2, m3) &= \sqrt{(17.33 - 18)^2 + (16.08 - 19.25)^2 + (18 - 14.25)^2} \\ &= 4.95 \end{aligned}$$

$$\begin{aligned} \text{Nilai } BCV &= 8.56 + 9.32 + 4.95 \\ &= 22.83 \end{aligned}$$

%1. Menghitung kembali nilai rasio dengan membandingkan nilai *BCV* dan *WCV*.

$$\begin{aligned} BCV/WCV &= 22.83/160.58 \\ &= 0.14 \end{aligned}$$

Dikarenakan nilai rasio *BCV* dan *WCV* pada iterasi kedua dan ketiga sama, maka perhitungan dihentikan pada iterasi ke 3 dengan nilai rasio 0.14.

Pada tahap ini dapat diketahui hasil klasterisasi menggunakan algoritma *K-Means Clustering* untuk segmentasi pemetaan pasar. Mengetahui potensial Penjualan Deli Serdang 2019 minuman lasegar. Membantu perusahaan dalam mengetahui pasar-pasar yang perlu dilakukan promosi sehingga dapat meningkatkan persediaan Penjualan Deli Serdang 2019 dan perusahaan.

Tabel 3.10 Pengelompokkan Hasil Cluster

Cluster	Bulan Periode Penjualan
Cluster 1	1,2,3
Cluster 2	3,4,5,6,7,8,9,10,11
Cluster 3	12

%1. Pengujian dan implementasi

Implementasi sistem merupakan kegiatan akhir dari proses penerapan sistem baru dimana proses yang baru ini akan dioperasikan secara menyeluruh. Sebelum sistem benar-benar bias digunakan dengan baik, sistem harus melalui tahap pengujian terlebih dahulu untuk menjamin tidak ada kendala yang muncul pada saat sistem digunakan. Implementasi yang dilakukan terdapat beberapa tahap prosedur untuk menyelesaikan analisa yaitu aplikasi yang disetujui, melakukan penginstalan, pengujian data, dan mulai menggunakan sistem yang diperbaiki atau system baru. Berikut merupakan implementasi dari system.

1. Tampilan *Form Login*

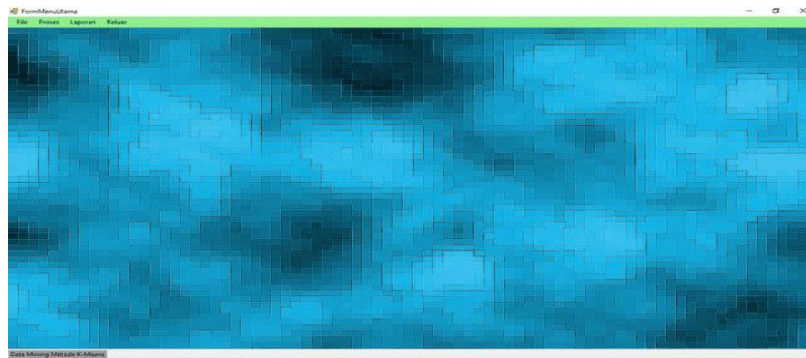
Berikut ini merupakan tampilan dari *form login* :

Gambar 5.1 Tampilan *Form Login*

Adapun *form login* berfungsi untuk melakukan proses validasi *username* dan *password* pengguna ketika sebelum masuk ke dalam tampilan *form* menu utama.

2. Tampilan *Form* Menu Utama

Berikut ini merupakan tampilan *form* menu utama setelah pengguna berhasil *login*, Adapun tampilan dari *form* menu utama yaitu :

Gambar 5.2 Tampilan *Form* Menu Utama

3. Tampilan *Form* Data Tren Penjualan

Form Data Tren Penjualan adalah *Form* pengolahan data tren penjualan pada jasa PT. GATRA KASINDO JAYA dalam penginputan data, ubah data dan penghapusan data tren penjualan. Adapun *Form* data tren penjualan adalah sebagai berikut:

Nilai Centroid

Centroid - 1

Centroid 1

Centroid 2

Centroid 3

	25	21	20
Centroid - 1	25	21	20
Centroid - 2	15	16	17
Centroid - 3	15	19	15

Gambar 5.3 Tampilan *Form* Data centroid

4. Tampilan *Form* Proses *K-Means*

Form Proses Regresi adalah *form* data yang digunakan untuk memprediksi trend penjualan pada jasa PT. GATRA KASINDO JAYA Tampilan *form* proses *K-Means* dapat dilihat sebagai berikut :

Nilai Centroid

No	Bulan Penjualan	Kode	C1	C2	C3
1	Januari	1	25	21	20
2	April	4	22	20	21
3	Desember	12	15	16	17
4	Februari	2	25	22	21
5	Maret	3	20	21	25
6	Mei	5	15	22	15
7	Juni	6	17	15	20

Nilai Centroid

	25	21	20
Centroid - 1	25	21	20
Centroid - 2	15	16	17
Centroid - 3	15	19	15

Penghitungan Jarak Centroid

No	Bulan Penjualan	Kode	C1	C2	C3

Cluster baru

No	Bulan Penjualan	C1	C2	C3	Cluster

Pilih Data Cluster yang ingin di cetak

Gambar 5.4 Tampilan *Form* Proses *Regresi*

5.3 Pengujian

Setelah melakukan proses implementasi, proses selanjutnya adalah melakukan pengujian terhadap sistem yang telah di bangun. Pengujian ini untuk melihat bahwa hasil perancangan dan perhitungan yang ada di bab III sesuai dengan hasil yang ditampilkan pada sistem. Keluaran yang dihasilkan oleh sistem akan disesuaikan dengan hasil perhitungan.

%1. Kesimpulan dan Saran

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang di bahas tentang mengetahui prediksi baju dengan menerapkan metode *Regresi* terhadap sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan sebagai berikut :

- %1. Untuk menganalisa masalah dalam memprediksi tren penjualan dengan mengambil data tren penjualan untuk melakukan pengujian dalam data mining dengan menggunakan metode *regresi*.
- %1. Berdasarkan hasil analisa perancangan sistem ini, yang dimulai pertama kali dalam program ini adalah rancangan *database* rancangan program seperti tampilan *form login* tampilan menu utama tampilan input data tren penjualan lalu tampilan *form* proses metode *regresi* dan yang terakhir adalah hasil dari data.
- %1. Berdasarkan hasil analisa dalam mengimplementasikan sistem ini diisi kedalam sebuah komputer baik digunakan oleh *user* dengan cara *login* kemudian memasukan data-data yang kemudian akan memproses hasil dari data tersebut dan menerima hasil laporannya.

Dari hasil penelitian ini, dapat disampaikan beberapa saran sebagai berikut:

- %1. Agar lebih efektif aplikasi itu bisa dikembangkan lagi sehingga bisa mengolah data yang sifatnya lebih banyak lagi hingga ribuan sehingga bisa menghasilkan hasil yang lebih optimal dan akurat.
- %1. Di harapkan aplikasi ini bisa di pakai untuk perusahaan lain yang produknya sejenis atau menggunakan konsep produk yang sama sehingga bisa lebih berkembang lebih besar lagi untuk masyarakat luas.
- %1. Apabila PT. GATRA KASINDO JAYA nanti berkembang maka sistem ini diharapkan pengembangannya bisa multi *user* kemudian bisa diakses dari berbagai tempat atau cabang dari perusahaan tersebut sehingga data lebih terpusat dan hasilnya bisa lebih optimal.

REFERENSI

- [1] D. Yuliasari and F. Rahmawati, "HUBUNGAN POLYHIDRAMNION DAN PRESENTASI JANIN DENGAN KEJADIAN KETUBAN PECAH DINI DI RS PURI BETIK HATI PROVINSI LAMPUNG TAHUN 2016," 2017.
- [2] M. Puji Sari Ramadhan and M. Usti Fatimah S. Pane, Judul : Mengenal Metode Sistem Pakar, Cetakan Pertama ed., Fungy, Ed., 2018.
- [3] N. Budi Riyanto and O. Suria, "Sistem Pakar Diagnosa Penyakit Pencernaan Menggunakan Metode Teorema Bayes 7".
- [4] M. J. Effendi, M. Triawan and S. Musirawas Lubuklinggau, "SISTEM PAKAR UNTUK MENDIAGNOSA PENYAKIT TANAMAN KOPI BERBASIS WEB," 2019.
- [9] M. Zulfian Azmi, ST., M.Kom. dan Verdi Yasin, S.Kom., Pengantar Sistem Pakar dan Metode (Introduction of Expert System and Methods), Jakarta: Mitra Wacana Media, 2019, pp. 11-17.
- [10] Chairun Na s, "SISTEM PAKAR DIAGNOSA PENYAKIT TIROID MENGGUNAKAN METOD E DEMPSTERSHAPE R," JURNALTEKNOLOGIDANOPENSOURCE, vol. VOL.2No.1, 2019.

-
- [11] N. Sari Br Sembiring and M. Dayan Sinaga, "Penerapan Metode Dempster Shafer Untuk Mendiagnosa Penyakit Dari Akibat Bakteri Treponema Pallidum Application Of Dempster Shafer Method For Diagnosing Diseases Due To Treponema Pallidum Bacteria," 180. CSRID Journal, vol. 9, no. 3, 2017.
- [12] P. Metode, D. Shafer, U. Mendiagnosa, P. Dari, A. Bakteri, S. Mikha, D. Sinaga, N. Sari and B. Sembiring, "Penerapan Metode Dempster Shafer... □ 94".

UCAPAN TERIMA KASIH

Terima kasih diucapkan kepada pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal selanjutnya.

BIOGRAFI PENULIS

	Nama : Bagas Anjas Wara Tempat Tanggal Lahir : Siantar 22 February 1998 Alamat : Jl. Pringgan Dusun III Agama : Islam Email : anjasbagas01@gmail.com
	Nama Pembimbing I : Yohanni Syahra, S.Si., M.Kom Jabatan : Ka. BAAK
	Nama Pembimbing II : Milfa Yetri, S.Kom., M.Kom Jabatan : Sekretaris Prodi MI