PENERAPAN ALGORITMA K-MEANS UNTUK MENETAPKAN KELOMPOK MUTU KARET

Handi Kurniawan Sohdianata [1] , Sushermanto [2]
Jurusan Teknik Informatika STMIK Banjarbaru [1], Jurusan Sistem Informasi STMIK Banjarbaru [2]
Jl. Jend. Ahmad Yani Km. 33.3 Loktabat Banjarbaru[1,2]
my5tmik@yahoo.co.id [2]

ABSTRAK

Selama ini penglompokkan mutu karet dilakukan melalui pengujian lab terhadap kadar kotoran, kadar abu, PO dan PRI. Dengan perkembangan ilmu pengetahuan yang semakin pesat, pengelompokkan karet bisa digantikan oleh cara-cara/metode baru. *K-Means* adalah salah satu metode yang telah digunakan pada penelitian ini untuk mengelompokkan karet berdasarkan standar mutunya.

Algoritma *K-Means* yang digunakan mengelompokkan data karet berdasarkan kadar kotoran, kadar abu, V.M, PO, dan PRI. Penelitian ini menguji tingkat akurasi algoritma *K-Means* dalam penetapan mutu karet.

Penerapan algoritma *K-Means* dalam penetapan kelompok mutu karet pada 75 sample data karet yang diuji menunjukkan bahwa algoritma *K-Means* memiliki 35 data yang akurat (47%) dan 40 data yang tidak akurat (53%).

Kata Kunci : K-Means, pengelompokan mutu, karet.

I. PENDAHULUAN

Mutu karet biasanya dipengaruhi oleh kadar kotoran, kadar abu, VM (zat menguap), PO (daya tekan karet) dan PRI.Dengan perkembangan ilmu pengetahuan yang semakin pesat, pengelompokkan karet yang biasanya menggunakan uji laboratorium nantinya bisa digantikan oleh cara-cara baru. Pendekatan data mining dimungkinkan dapat digunakan dalam mengelompokkan karet berdasarkan standar mutunya.

Salah satu tugas dalam data mining adalah *clustering*. Tujuan utama dari *clustering* adalah pengelompokkan sejumlah data/obyek ke dalam *cluster* sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin. Salah satu metode *clustering* yang umum digunakan adalah algoritma *K-Means*.

K-Means adalah salah satu metode yang diharapkan dapat diterapkan dalam kasus pengelompokkan karet berdasarkan standar mutunya. Sistem ini akan menggunakan variabel kadar kotoran, kadar abu, V.M, PO (daya tekan karet), dan PRI. Data-data ini nantinya akan diproses dan dianalisa untuk kemudian dilakukan peng*cluster*an atau pengelompokkan mutu karet sesuai dengan standar mutu yang sudah ditetapkan. Diharapkan metode *K-Means* dapat mengelompokkan karet berdasarkan standar mutunya dengan akurat.

Adapun rumusan masalah dalam penelitian ini Penerapan Algoritma *K-Means* untuk menetapkan mutu karet berdasarkan kadar kotoran , kadar abu, VM, PO dan PRI pada PT. Insan Bonafide Banjarmasin .

II. TINJAUAN PUSTAKA DAN KERANGKA PEMIKIRAN

a. Tinjauan Pustaka

Salah satu metode yang banyak digunakan dalam melakukan clustering/pengelompokkan data adalah metode *K-Means*. Alasan kenapa orang banyak memakai metode ini adalah cukup sederhana dan cepat dalam pekerjaan pengelompokkan data.

Di dalam jurnal dengan judul Aplikasi *K-Mean*s Untuk Pengelompokkan Mahasiswa

Berdasarkan Nilai Body Mass Index (BMI) & Ukuran Kerangka (Rismawan & Kusumadewi, 2008) meneliti tentang pengelompokkan mahasiswa berdasarkan BMI dengan menggunakan algoritma *K-Means*. Pada penelitian ini dibangun sistem yang dapat digunakan untuk mengklasifikasi mahasiswa menurut BMI dan ukuran kerangkanya berdasarkan data kondisi fisik dari mahasiswa yang bersangkutan yang telah diambil terlebih dahulu. Data kondisi fisik yang digunakan adalah tinggi badan, berat badan dan lingkar lengan bawah. Diasumsikan data yang diambil adalah data mahasiswa putra. Setelah data tersebut diperoleh kemudian dilakukan perhitungan untuk mencari status gizi dan ukuran rangka dari masing-masing data yang ada. Setelah mendapatkan status gizi dan nilai rangka dari masing-masing data maka langkah selanjutnya adalah melakukan proses klasifikasi data menggunakan metode klasifikasi *K-Means*.

Di jurnal lainnya dengan judul Pengelompokkan Dokumen Menggunakan *K-Means* Dan Singular Value Decomposition : Studi Kasus Menggunakan Data Blog (Umran & Abidin, 2009). Di dalam penelitian ini *K-means* memilih secara acak k buah data sebagai *centroid*. Kemudian menempatkan data dalam cluster yang terdekat, dihitung dari titik tengah *cluster* (*centroid*). *Centroid* baru akan ditentukan bila semua data telah ditempatkan dalam *cluster* terdekat. Proses penentuan *centroid*
dan penempatan data dalam *cluster* diulangi sampai nilai *centroid konvergen*.

K-means adalah salah satu metode klasterisasi, klasteriasasi adalah salah satu fungsi di dalam data mining, yang dikembangkan oleh Mac Queen pada tahun 1967 dan kemudian dilanjutkan sama JA Hartigan dan MA Wong sekitar 1975. Metode ini banyak digunakan di berbagai bidang karena sederhana, mudah diimplementasikan, memiliki kemampuan untuk mengklaster data yang besar, mampu menangani data outlier, dan kompleksitas waktunya linear O(nKT) dengan n adalah jumlah dokumen, K adalah jumlah *cluster*, dan T adalah jumlah *iterasi*.

Dalam algoritma *K-Means*, setiap data harus termasuk ke cluster tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke *cluster* yang lain. Pada dasarnya penggunaan algoritma *K-Means* dalam melakukan proses clustering tergantung dari data yang ada dan konklusi yang ingin dicapai. Untuk itu digunakan algoritma *K-Means* yang didalamnya memuat aturan sebagai berikut:

a. Jumlah *cluster* yang perlu di inputkan
b. Hanya memiliki attribut bertipe numerik

Algoritma *K-Means* pada awalnya mengambil sebagian dari banyaknya komponen dari populasi untuk dijadikan pusat *cluster* awal. Pada step ini pusat *cluster* dipilih secara acak dari sekumpulan populasi data. Berikutnya *K-Means* menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu pusat *cluster* yang telah di definisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap pusat *cluster*. Posisi pusat *cluster* akan dihitung kembali sampai semua komponen data digolongkan kedalam tiap-tiap *cluster* dan terakhir akan terbentuk posisi *cluster* baru.

Adapun tujuan dari data *clustering* ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalisasikan variasi antar cluster. *Objective function* yang berusaha diminimalkan oleh *K-Means* adalah :

$$J(U,V) = SUM(k = 1\ to\ N)SUM(i = 1\ to\ c)(a_ik * (x_k, v_i)^2)$$

dimana :
U : matriks keanggotaan data ke masing-masing cluster yang berisikan nilai 0 dan 1
V : matriks centroid/rata-rata masing-masing cluster
N : jumlah data
c : jumlah cluster
a_ik: keanggotaan data ke-k ke cluster ke-i
x_k : data ke-k
v_i : nilai centroid cluster ke-i

Algoritma K-Means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat cluster dan proses pencarian anggota dari tiap-tiap cluster. Proses algoritma K-means adalah sebagai berikut :

1. Tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk.
2. Bangkitkan k *centroid* (titik pusat *cluster*) awal secara random. *Centroid*/rata-rata dari data didapatkan menggunakan rumus sebagai berikut:

$$v_ij = \frac{SUM(k = 0\ to\ N_i)(x_kj)}{N_i}$$

dimana:
i,k : indeks dari cluster
j : indeks dari variabel
v_ij : centroid/rata-rata cluster ke-i untuk variabel ke-j
N_i : jumlah data yang menjadi anggota cluster ke-i

3. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua objek yaitu *Euclidean Distance* dan kesamaan *Cosine*.
   Rumus Euclidean Distance $d_{v=\sqrt{\sum_{k=1}^{n}(X_{Jk}-X_{Jk})^2}}$:
4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
5. Tentukan posisi *centroid* baru (C) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama. $C_k = \sum\left(\frac{1}{n_k}\right)\sum d_i$
   Dimana *n* adalah jumlah dokumen dalam *cluster k* dan *d* adalah dokumen *cluster* k.
6. Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

Berdasarkan cara kerjanya, algoritma *K-Means* memiliki karakteristik:

1. K-Means sangat cepat dalam melakukan proses *clustering*.
2. K-Means sangat sensitif pada pembangkitan *centroid* awal secara random. Pada saat pembangkitan awal titik pusat yang random tersebut mendekati solusi akhir pusat *cluster*, *K-Means* mempunyai posibilitas yang tinggi untuk menemukan titik pusat *cluster* yang tepat. Sebaliknya, jika awal titik pusat tersebut jauh dari solusi akhir pusat *cluster*, maka besar kemungkinan ini menyebabkan hasil pengklasteran yang tidak tepat.
3. Memungkinkan suatu *cluster* tidak mempunyai anggota.
4. Hasil *clustering* dengan *K-Means* bersifat tidak unik (selalu berubah-ubah) – terkadang baik, terkadang jelek. Ini disebabkan karena *K-Means* sangat sensitif dalam menentukan *centroid* awal.

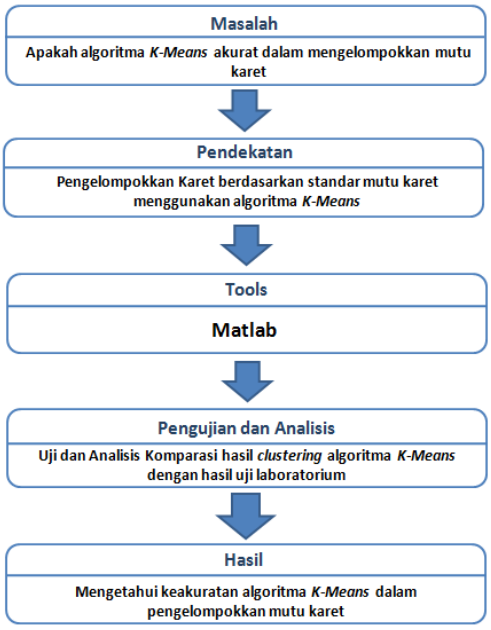
b. Standar Mutu Karet

Bokor SIR adalah karet yang berasal dari lateks kebun dari pohon karet. Bokor SIR tergolong ke dalam karet spesifikasi teknis, karena penilaian mutunya didasarkan pada sifat teknis dari parameter dan besaran nilai yang dipersyaratkan dalam penetapan mutu karet remah dan tercantum dalam skema SIR. PT. Insan Bonafide untuk memenuhi konsistensi mutu maka menetapkan mutu karetnya berdasarkan tabel berikut:

Tabel 1. Kadar Mutu Karet

No	Pelanggan	Jenis	Spesifikasi				
			Kadar Kotoran	Kadar Abu	V.M	PO	PRI
			%	%	%	-	-
1.	SNI SIR 06-1903-2000	SIR 10	Maks. 0,10	Maks. 0,75	Maks. 0,80	Min. 30	Min. 60
		SIR 20	Maks. 0,20	Maks. 1,00	Maks. 0,80	Min. 30	Min. 50
2.	Bridgestone	SIR 20	Maks. 0,20	Maks. 1,00	Maks. 0,80	32-36	Min. 50
3.	Goodyear	SIR 10	Maks. 0,10	Maks. 0,75	Maks. 0,80	32-38	Min. 60
		SIR 20	Maks. 0,20	Maks. 1,00	Maks. 0,80	32-38	Min. 60
4.	SMPT	SIR 20	Maks. 0,14	Maks. 1,00	Maks. 0,80	33-41	Min. 40
5.	Weber & Schaer	SIR 10	Maks. 0,20	Maks. 1,00	Maks. 0,80	37-43	Min. 70

c. Kerangka Pemikiran Penelitian



Gambar 1. Kerangka Penelitian

III. METODE PENELITIAN

a. Analisa Kebutuhan

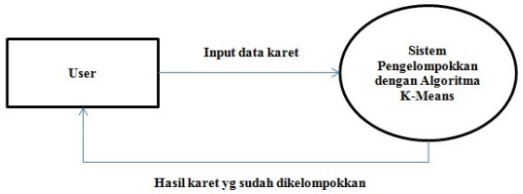
Secara umum, sistem pengelompokkan ini bertugas mengelompokkan data karet berdasarkan standar mutu yang sudah ditetapkan. Variabel yang digunakan adalah berupa kadar kotoran, kadar abu, PO dan PRI. Berdasarkan inputan yang diberikan oleh user, nantinya *K-Means* akan menciptakan output berupa hasil karet yang sudah dikelompokkan sehingga kita akan tahu karet tersebut termasuk kedalam kelompok mutu yang mana.

b. Metode Pengumpulan Sampel

Metode Pengumpulan data Observasi dan Interview.

Sampel data yang digunakan adalah data karet PT. Insan Bonafide. Variabel data yang akan digunakan dalam penelitian adalah kadar kotoran, kadar abu, V.M, PO, dan PRI

c. Perancangan Penelitian



Gambar 2. Diagram Konteks

d. Teknik Analisis Data

Data karet dikelompokkan berdasarkan mutu karet menggunakan algoritma *K-Means*. Data yang ada akan dikelompokkan menjadi 5 kelompok/*cluster*.

Selanjutnya menentukan pusat *cluster* secara acak. *Cluster* yang terbentuk secara acak dengan menggunakan Matlab adalah:

```
>>[IDX,C] = kmeans(X,5)
C=
   0,0486   0,5484   0,3096  32,8800  68,5200
   0,0479   0,5589   0,2772  30,2222  66,5778
   0,0485   0,5675   0,3008  33,3333  71,7583
   0,0528   0,5200   0,2967  37,1111  70,3889
   0,0534   0,5118   0,3118  36,8182  72,8455
```

Sehingga C1 = (0,0486; 0,5484; 0,3096; 32,8800; 68,5200), C2 = (0,0479; 0,5589; 0,2772; 30,2222; 66,5778), C3 = (0,0485; 0,5675; 0,3008; 33,3333; 71,7583), C4 = (0,0528; 0,5200; 0,2967; 37,1111; 70,3889) dan C5 = (0,0534; 0,5118; 0,3118; 36,8182; 72,8455).

Hitung jarak setiap data yang ada terhadap setiap pusat cluster dengan menggunakan rumus *Euclidean Distance*, yaitu:

$$d_{v}=\sqrt{\sum_{k=1}^{n}(X_{Jk}-X_{Jk})^{2}}$$

Misalkan untuk menghitung jarak data karet pertama dengan pusat *cluster* pertama adalah:

$d_{11}=\sqrt{(0,049-0,0486)^2+(0,53-0,5484)^2+(0,24-0,3096)^2+(30-32,8800)^2+(66,7-68,5200)^2}$ = 3,407

$d_{12}=\sqrt{(0,049-0,0479)^2+(0,53-0,5589)^2+(0,24-0,2772)^2+(30-30,2222)^2+(66,7-66,5778)^2}$ = 0,257

$d_{13}=\sqrt{(0,049-0,0485)^2+(0,53-0,5675)^2+(0,24-0,3008)^2+(30-33,3333)^2+(66,7-71,7583)^2}$ = 6,058

$d_{14}=\sqrt{(0,049-0,0528)^2+(0,53-0,5200)^2+(0,24-0,2967)^2+(30-37,1111)^2+(66,7-70,3889)^2}$ = 8,011

$d_{15}=\sqrt{(0,049-0,0534)^2+(0,53-0,5118)^2+(0,24-0,3118)^2+(30-36,8182)^2+(66,7-72,8455)^2}$ = 9,179

Dari hasil perhitungan tersebut di atas dihitung pusat *cluster* baru dengan menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama dan menggunakan rumus $C_k = \sum\left(\frac{1}{n_k}\right)\sum d_i$

Untuk *cluster* pertama ada 24 data yaitu data ke-5, 9, 10, 11, 19, 20, 21, 22, 24, 28, 31, 34, 41, 47, 49, 53, 58, 63, 67, 70, 71, 72, 73, dan 74, sehingga:

C_{11}=(0,055+0,047+0,045+0,056+0,044+0,047+0,05+0,049+0,05+ 0,055+0,047+0,044+0,048+0,044+0,051+0,047+0,046+0, 053+0,045+0,049+0,049+0,049+0,047+0,049)/24

= 0,048

C_{12}=(0,53+0,56+0,58+0,54+0,5+0,53+0,54+0,56+0,55+0,53+0,58 +0,6+0,53+0,56+0,6+0,5+0,52+0,57+0,49+0,57+0,55+0,5 3+0,56+0,53)/24

= 0,546

C_{13}=(0,32+0,25+0,27+0,3+0,29+0,31+0,33+0,35+0,33+0,34+0,34 +0,25+0,36+0,28+0,29+0,34+0,3+0,33+0,26+0,28+0,35+ 0,33+0,32+0,3)/24

= 0,309

C_{14}=(32+33+34+32+34+33+33+34+32+33+32+32+33+32+34+34 +32+33+32+30+33+33+33+34)/24

=32,791

C_{15}=(68,8+69,7+67,6+68,8+67,6+69,7+69,7+67,6+68,8+69,7+68, 8+68,8+66,7+68,6+67,7+67,6+68,8+69,7+68,8+70+66,7+ 66,7+69,7+67,6)/24

= 68,516

Untuk *cluster* kedua, ada 17 data yaitu data ke-1, 2, 4, 16, 17, 18, 33, 35, 36, 40, 46, 51, 54, 56, 57, 62, dan 65 sehingga:

C_{21}=(0,049+0,043+0,047+0,048+0,052+0,057+0,050+0,049+0,040 +0,045+0,045+0,050+0,052+0,042)/17

= 0,047

C_{22}=(0,53+0,53+0,54+0,64+0,65+0,57+0,57+0,51+0,55+0,54+0,5 7+0,55+0,52+0,59+0,54+0,53)/17

= 0,556

C_{23}=(0,24+0,32+0,3+0,35+0,36+0,24+0,29+0,26+0,28+0,27+0,24 +0,27+0,25+0,27+0,25+0,27+0,28)/17

= 0,278

C_{24}=(30+30+32+30+30+30+30+30+30+30+30+30+30+30+30+30 +32)/17

= 30,235

C_{25}=(66,7+66,7+65,6+66,7+66,7+66,7+66,7+66,7+66,7+66,7+66,7+66, 7+66,7+66,7+66,7+66,7+66,7+65,6)/17

= 66,570

Untuk *cluster* ketiga, ada 12 data yaitu data ke-6, 12, 13, 25, 30, 37, 38, 45, 52, 55, 59, dan 75 sehingga:

C_{31}=(0,049+0,042+0,046+0,057+0,052+0,045+0,054+0,053+0,044 +0,052+0,052+0,036)/12

= 0,048

C_{32}=(0,59+0,62+0,54+0,58+0,59+0,56+0,58+0,63+0,52+0,52+0,5 4+0,54)/12

= 0,567

C_{33}=(0,34+0,32+0,26+0,34+0,28+0,27+0,31+0,26+0,36+0,28+0,2 9+0,3)/12

= 0,300

C_{34}=(34+35+35+32+34+33+32+32+35+34+32+32)/12

= 33,333

C_{35}=(73,5+71,4+71,4+71,9+70,6+72,7+71,9+71,9+71,4+70,6+71, 9+71,9)/12

= 71,758

Untuk *cluster* keempat, ada 10 data yaitu data ke-8, 26, 32, 39, 42, 44, 50, 60, 61, dan 69, sehingga:

C_{41}=(0,045+0,048+0,054+0,056+0,05+0,05+0,058+0,052+0,051+ 0,059)/10

= 0,052

C_{42}=(0,52+0,6+0,54+0,54+0,53+0,49+0,56+0,46+0,5+0,54)/10

= 0,528

C_{43}=(0,31+0,32+0,3+0,35+0,3+0,32+0,3+0,25+0,23+0,31)/10

= 0,299

C_{44}=(37+35+37+37+37+38+37+37+37)/10

= 36,9

C_{45}=(70,3+68,6+70,3+70,3+70,3+70,3+71,1+70,3+70,3+70,3)

=70,2

Untuk *cluster* kelima, ada 12 data yaitu data ke-3, 7, 14, 15, 23, 27, 29, 43, 48, 64, 66, dan 68 sehingga:

C_{51}=(0,05+0,057+0,058+0,051+0,046+0,055+0,045+0,055+0,055+0,058+0,057+0,055)/12
= 0,053
C_{52}=(0,51+0,51+0,52+0,52+0,51+0,49+0,53+0,47+0,51+0,58+0,48+0,6)/12
= 0,519
C_{53}=(0,34+0,3+0,28+0,29+0,28+0,33+0,31+0,27+0,33+0,36+0,34+0,25)/12
= 0,306
C_{54}=(37+36+37+36+37+37+37+37+37+36+38+30)/12
= 36,25

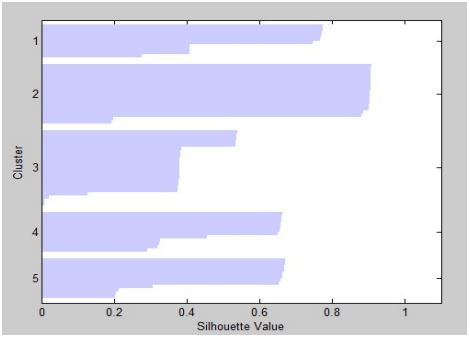
C_{55}=(73+72,2+73+72,2+73+73+73+73+73+72,2+73,7+66,7)/12
=72,333

Ulangi penghitung jarak setiap data terhadap setiap pusat *cluster* dimana pusat *cluster* di iterasi kedua adalah:
C =

0,048	0,546	0,309	32,791	68,516
0,047	0,556	0,2788	30,235	66,570
0,048	0,567	0,300	33,333	71,758
0,052	0,528	0,299	36,9	70,21
0,053	0,519	0,306	36,25	72,333

Sehingga C1 = (0,048; 0,546; 0,309; 32,791; 68,516), C2 = (0,047; 0,556; 0,2788; 30,235; 66,570), C3 = (0,048; 0,567; 0,300; 33,333; 71,758), C4 = (0,052; 0,528; 0,299; 36,9; 70,21) dan C5 = (0,053; 0,519; 0,306; 36,25; 72,333)

Adapun hasil pengelompokan dapat dilihat pada gambar berikut ini :



Gambar 3. Hasil Pengelompokan

IV. HASIL DAN PEMBAHASAN

Setelah melakukan penghitungan didapat 4 iterasi, dimana pusat *cluster* terakhirnya adalah sebagai berikut:
C =

0,048	0,548	0,309	32,88	68,52
0,047	0,556	0,2788	30,235	66,570
0,05	0,57	0,296	32,777	71,877
0,052	0,52	0,296	37,11	70,388
0,051	0,527	0,308	36	72,14

Dari informasi pusat *cluster* yang dihasilkan pada iterasi terakhir, dapat disimpulkan bahwa:

1. *Cluster* pertama memiliki pusat (0,048; 0,548; 0,309; 32,88; 68,52) yang dapat diartikan sebagai kelompok karet dengan mutu Bridgestone SIR20.
2. *Cluster* kedua memiliki pusat (0,047; 0,556; 0,2788; 30,235; 66,570) yang dapat diartikan sebagai kelompok karet dengan mutu SNI SIR 06-1903-2000 SIR20.
3. *Cluster* ketiga memiliki pusat (0,05; 0,57; 0,296; 32,777; 71,877) yang dapat diartikan sebagai kelompok karet dengan mutu Goodyear SIR20.
4. *Cluster* keempat memiliki pusat (0,052; 0,52; 0,296; 37,11; 70,388) yang dapat diartikan sebagai kelompok karet dengan mutu Weber & Scaher SIR20.
5. *Cluster* kelima memiliki pusat (0,051; 0,527; 0,308; 36; 72,14) yang dapat diartikan sebagai kelompok karet dengan mutu SMPT SIR 20.

Penetapan hasil akurasi didasarkan pada hasil uji laboratorium. Jika hasil kelompok mutu dari algoritma *K-Means* tidak sama dengan hasil uji laboratorium maka dinyatakan tidak akurat. Perbandingan akurasi yang dilakukan menggunakan algoritma *K-Means* disajikan pada tabel berikut ini:

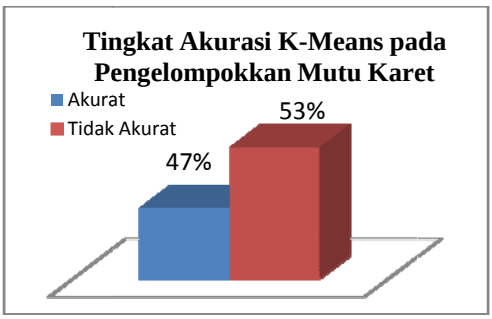
Tabel 2. Hasil Akurasi

Data Ke-	Hasil Uji Laboratorium	Hasil Algoritma *K-Means*	Ket
1	SNI SIR20	SNI SIR20	Akurat
2	SNI SIR20	SNI SIR20	Akurat
3	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
4	Bridgestone SIR20	SNI SIR20	Tidak Akurat
5	Bridgestone SIR20	Bridgestone SIR20	Akurat

6	Bridgestone SIR20	Goodyear SIR20	Tidak Akurat
7	Goodyear SIR20	SMPT SIR20	Tidak Akurat
8	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
9	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
10	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
11	Bridgestone SIR20	Bridgestone SIR20	Akurat
12	SNI SIR20	SMPT SIR20	Tidak Akurat
13	SMPT SIR20	SMPT SIR20	Akurat
14	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
15	Goodyear SIR20	SMPT SIR20	Tidak Akurat
16	SNI SIR20	SNI SIR20	Akurat
17	SNI SIR20	SNI SIR20	Akurat
18	SNI SIR20	SNI SIR20	Akurat
19	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
20	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
21	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
22	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
23	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
Data Ke-	Hasil Uji Laboratorium	Hasil Algoritma *K-Means*	Ket
24	Bridgestone SIR20	Bridgestone SIR20	Akurat
25	Bridgestone SIR20	Goodyear SIR20	Tidak Akurat
26	Bridgestone SIR20	Bridgestone SIR20	Akurat
27	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
28	SNI SIR20	Bridgestone SIR20	Tidak Akurat
29	Goodyear SIR20	SMPT SIR20	Tidak Akurat
30	Goodyear SIR20	Goodyear SIR20	Akurat
31	SNI SIR20	Bridgestone SIR20	Tidak Akurat
32	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
33	SNI SIR20	SNI SIR20	Akurat
34	SNI SIR20	Bridgestone SIR20	Tidak Akurat
35	SNI SIR20	SNI SIR20	Akurat
36	SNI SIR20	SNI SIR20	Akurat
37	Goodyear SIR20	Goodyear SIR20	Akurat
38	Bridgestone SIR20	Goodyear SIR20	Tidak Akurat
39	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
40	SNI SIR20	SNI SIR20	Akurat
41	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
42	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
43	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
44	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
45	Goodyear SIR20	Goodyear SIR20	Akurat
46	SNI SIR20	SNI SIR20	Akurat
47	SNI SIR20	Bridgestone SIR20	Tidak Akurat
48	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
49	Goodyear SIR20	Bridgestone SIR20	Tidak Akurat
50	Goodyear SIR20	Weber & Schaer SIR20	Tidak Akurat
51	SNI SIR20	SNI SIR20	Akurat
Data Ke-	Hasil Uji Laboratorium	Hasil Algoritma *K-Means*	Ket
52	SMPT SIR20	SMPT SIR20	Akurat
53	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
54	SNI SIR20	SNI SIR20	Akurat
55	Goodyear SIR20	Goodyear SIR20	Akurat
56	SNI SIR20	SNI SIR20	Akurat
57	SNI SIR20	SNI SIR20	Akurat
58	Bridgestone SIR20	Bridgestone SIR20	Akurat
59	Bridgestone SIR20	Goodyear SIR20	Tidak Akurat

60	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
61	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
62	SNI SIR20	SNI SIR20	Akurat
63	SNI SIR20	Bridgestone SIR20	Tidak Akurat
64	SNI SIR20	SMPT SIR20	Tidak Akurat
65	Bridgestone SIR20	SNI SIR20	Tidak Akurat
66	Weber & Schaer SIR20	SMPT SIR20	Tidak Akurat
67	Goodyear SIR20	Bridgestone SIR20	Tidak Akurat
68	SNI SIR20	SNI SIR20	Akurat
69	Weber & Schaer SIR20	Weber & Schaer SIR20	Akurat
70	SNI SIR20	Bridgestone SIR20	Tidak Akurat
71	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
72	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
73	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
74	SMPT SIR20	Bridgestone SIR20	Tidak Akurat
75	Bridgestone SIR20	Goodyear SIR20	Tidak Akurat

Pada tabel 2 hasil penglompokkan yang dilakukan oleh algoritma *K-Means*, sebanyak 35 dari 75 data sampel karet atau 47% yang sesuai dengan kelompok mutunya dan 40 dari 75 data sampel karet atau 53% tidak sesuai dengan kelompok mutunya. Akurasi pengelompokkan mutu karet dapat dilihat pada grafik berikut ini:



Gambar 4. Grafik Akurasi Pengelompokan

V. KESIMPULAN DAN SARAN

Dari hasil pengujian algoritma *K-Means* dalam menentukan kelompok mutu karet pada 75 sampel data karet, menunjukkan sebanyak 35 data atau 47% yang akurat dan 40 data atau 53% yang tidak akurat. Dari hasil tersebut disimpulkan bahwa algoritma K-Means tidak cocok/ akurat dalam mengelompokkan karet berdasarkan mutunya.

Untuk mendapatkan tingkat akurasi yang lebih baik dalam penerapan algoritma *K-Means* untuk menetapkan kelompok mutu karet, serta untuk memberikan kontribusi yang lebih besar di dunia riset, disarankan agar hasil penelitian ini dikembangkan dengan cara memodifikasi algoritma *K-Means* yang digunakan saat ini dengan menggabungkan algoritma *K-Means* dengan algoritma lain atau mengganti algoritma *K-Means* dengan
algoritma yang lain.

DAFTAR PUSTAKA

Fadli, A (2011, Mare 14), *Ilmukomputer.Com*, Dipetik April 17, 2011.

Rismawan, T., & Kusumadewi.S, 2008, *Aplikasi K-Means untuk Pengelompokan Mahasiswa Berdasarkan BMI & Ukuran* Kerangka, Graha Ilmu, Surabaya.

Sani Susanto & Dedy Suryadi, 2010, *Pengantar Data Mining*, Andi Offset, Yogyakarta

Penulis:
Handy Kurniawan Sohdianata, S.Kom.
Alumni Jurusan Teknik Informatika
STMIK Banjarbaru
Tahun 2011

Drs. H. Sushermanto, M.Kom.
Dosen Tetap Yayasan
Pada Jurusan Sistem Informasi
STMIK Banjarbaru