Pengelompokan Pengiriman Hasil Kelapa Sawit Berdasarkan Tonase dan Kwalitas Menggunakan Metode *Clustering* (Studi Kasus : KUD Bumi Pusaka)

Beri Lesmana
STMIK Kaputama Binjai
E-mail: **berris144@gmail.com**

Abstrak

Prospek pasar bagi olahan kelapa sawit cukup menjanjikan, karena permintaan dari tahun ke tahun mengalami peningkatan yang cukup besar, tidak hanya didalam negeri, tetapi juga di luar negeri. Sebagai negara tropis yang masih memiliki lahan yang cukup luas, Indonesia berpeluang besar untuk mengembangkan pertanian kelapa sawit. Untuk mengetahui data pengiriman hasil kelapa sawit berdasarkan pabrik, tonase dan kwalitas, maka perlu adanyadilakukan perancangan sistem data pengiriman hasil kelapa sawit untuk mengetahui hasil tonase dan kwalitas kelapa sawit yang sebenarnya. Salah satucarauntukmengetahui pengiriman hasilkelapa sawit adalah dengan mengelompokan data-data pengiriman kelapa sawit ke setiap pabrik. *Cluster analysis* adalah pekerjaan mengelompokan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya. Dari 500 data pengiriman hasil kelapa sawit diperoleh 3 *Group*, di mana *Group* 1 berjumlah 174 data, *Group* 2 berjumlah 246 data dan *Group* 3 berjumlah 80 data. Dapat diketahui pada *cluster* 1 terdapat pada pabrik APL (Anugrah Putra Langkat) berat tonase Max 10 ton memiliki kwalitas Tenera lebih tebal pada batoknya, pada *cluster* 2 diketahui terdapat pada pabrik JPN (Jaya Palma Nusantara) berat tonase Min 2 ton memiliki kwalitas Tidak restan sedangkan pada *cluster* 3 terdapat pada pabrik MPA (Mega Pusaka Andalas) berat tonase Min 2 ton memiliki kwalitas Tenera lebih tebal pada batoknya.

Kata Kunci :*Data Mining, Clustering,* kelapa sawit.

Abstract

The market prospect for processed palm oil is promising, because the demand from year to year has increased quite significantly, not only domestically, but also abroad. As a tropical country which still has a large area of land, Indonesia has a great opportunity to develop oil palm agriculture. To find out the data on the delivery of oil palm products based on mill, tonnage and quality, it is necessary to design a data system for the delivery of oil palm products to determine the actual tonnage yield and quality of oil palm. One way to find out the shipments of palm oil products is by grouping data on palm oil shipments to each mill. Cluster analysis is the job of grouping data (objects) based solely on the information found in the data that describes these objects and the relationships between them. From 500 data on oil palm product shipments, 3 groups were obtained, of which Group 1 had 174 data, Group 2 had 246 data and Group 3 had 80 data. It can be seen that in cluster 1, it is found in the APL (Anugrah Putra Langkat) factory, the max tonnage weight of 10 tons has the quality of the Tenera is thicker on the shell, in cluster 2 it is known that it is found in the JPN (Jaya Palma Nusantara) factory. in cluster 3 there is an MPA (Mega Pusaka Andalas) factory with a tonnage weight of Min 2 tons which has a thicker Tenera quality on the shell

Keywords : Data Mining, *Clusering*, *Oil Palm*.

1. Pendahuluan

Kelapa sawit merupakan salah satu jenis tanaman perkebunan yang menduduki posisi terpenting di sektor pertanian, hal ini di karenakan kelapa sawit mampu menghasilkan nilai ekonomi terbesar perhektarnya jika dibandingkan dengan tanaman penghasil minyak atau lemak lainnya. Selain itu kelapa sawit juga memiliki banyak manfaat yaitu sebagai bahan bakar alternatif Biodisel, bahan pupuk komos, bahan dasar industi lainnya seperti industri kosmetik, industri makanan, dan sebagai obat. Prospek pasar bagi olahan kelapa sawit cukup menjanjikan, karena permintaan dari tahun ke tahun mengalami peningkatan yang cukup besar, tidak hanya didalam negeri, tetapi juga di luar negeri. Sebagai negara tropis yang masih memiliki lahan yang cukup luas, Indonesia berpeluang besar untuk mengembangkan pertanian kelapa sawit.

Dengan demikian untuk mengetahui pengiriman kelapa sawit berdasarkan tonase dan kwalitasnya yang terdapat di arsip data, maka pihak KUD Bumi Pusaka perlu mempunyai sistem data pengelompokan pengiriman hasil kelapa sawit yaitu yang memiliki prosedur yang terstruktur jelas yang sesuai dengan visi, misi dan strategi. Untuk mengetahui data pengiriman hasil kelapa sawit berdasarkan pabrik, tonase dan kwalitas, maka perlu adanyadilakukan perancangan sistem data pengiriman hasil kelapa sawit untuk mengetahui hasil tonase dan kwalitas kelapa sawit yang sebenarnya.

Analisa dilakukan dengan *Metode Clustering* yang menggunakan metode *K-Means* yang kemudian diterjemahkan dalam sebuah perangkat lunak. Perangakat lunak ini yang digunakan untuk pengelompokkan data.

Salah satu cara untuk mengetahui pengiriman hasil kelapa sawit adalah dengan mengelompokan data-data pengiriman kelapa sawit ke setiap pabrik .*Cluster analysis* adalah pekerjaan mengelompokan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya. Objek-objek yang bergabung dalam sebuah kelompok merupakan objek-objek yang mirip (atau berhubungan) satu sama lain dan berbeda (atau tidak berhubungan) dengan objek dalam kelompok yang lain. Salah satu metode yang paling banyak dilakukan pada metode *clustering* adalah dengan algoritma *K-Mean*.

2. Metode Penelitian

2.1. Pengertian *Data Mining*

Data Mining sering juga disebut *Knowledge Discovery in Database*, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Berikut ini pengertian menurut para ahli:

"Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar/Data *Warehouse*"[1].

Data mining sering dikenal dengan istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi didalam database. Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan dan pembelajaran mesin (*machine learning*) mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang terkait[2].

Dalam bukunya yang berjudul "*Data Mining Konsep dan Aplikasi Menggunakan Matlab*" menjelaskan bahwa : "*Data mining* adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar".

Secara umum definisi data mining dapat diartikan sebagai berikut:

1. Proses penemuan pola yang menarik dari data yang tersimpan dalam jumlah besar.
2. Ekstrasi dari suatu informasi yang berguna atau menarik (non-trivial, implicit, sebelumnya belum diketahui potensial kegunaannya) pola atau pengetahuan dari data yang disimpan dalam jumlah besar.

3. Eksplorasi dari analisa secara otomatis atau semiotomatis terhadap data-data dalam jumlah besar untuk mencari pola dan aturan yang berarti.

2.2. Langkah-langkah *Data Mining*

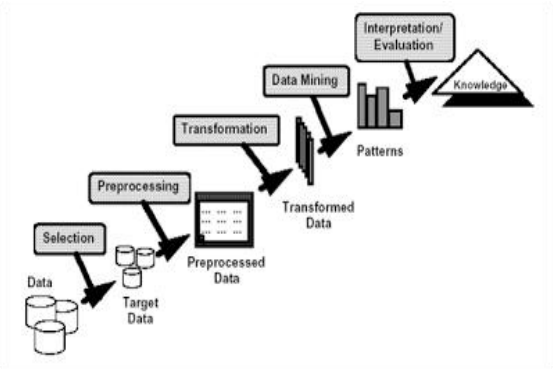
Berikut adalah langkah yang diperlukan dalam proses Data Mining [3]:

1. *Identity the Business Problem*
   Hal utama adalah mengetahui masalah bisnis yang dihadapi . Karena data tidak dapat diolah jika belum mengetahui apa masalah yang sedang hadapi. Maka harus dipahami masalah-masalah apa saja yang sedang dihadapu di bisnis tersebut. Dengan mengetahu masalah yang dihadapi kemudian dapat ditentukan data – data mana saja yang dibutuhkan untuk dapat dilakukan tahap analisa.
2. *Mine the Data for Actionable Information*
   Setelah mengetahui identifikasi masalah untuk memperoleh data-data mana saja yang diperlukan guna kepentingan analisa. Barulah dilakukan tahap analisa terhadap data –data tersebut. Dan dari analisa tersebut akan dapat memperoleh sebuah *knowledge* baru untuk kemudian dapat diambil suatu keputusan/kebijakan.
3. *Take the Action*
   Dan dari keputusan/kebijakan yang diperoleh dari proses data mining itu barulah diterapkan aksi berupa tindakan-tindakan yang kongkrit/nyata dalam proses bisnis.
4. *Measure Result*
   Selanjutnya adalah memonitor hasil dari langkah-langkah yang sudah dijalankan. Apakah sudah sesuai dengan target yang diharapkan dan apakah sudah bisa mengatasi masalah-masalah yang ada.

2.3. *Data Mining* Sebagai Proses Dalam *Knowledge Discovery In* Data (KDD)

Knowledge Discovery in Database (KDD) adalah keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, di mana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dimengerti[4].

KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data.



Gambar 1. Proses KDD *Data Mining*

Proses/langka-langkah dalam KDD Data Mining adalah sebagai berikut [5]:

1. *Data Selection*
   Menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada *subset variable* atau sampel data, di mana penemuan (discovery) akan dilakukan. Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining* disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-procesing/Cleaning*
   Pemrosesan pendahuluan dan pembersihan data merupakan operasi dasar seperti penghapusan *noise* dilakukan. Sebelum proses *data mining*dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi focus KDD. Proses cleaning mencakup antara lain membuang duplikasi data seperti kesalahan cetak (tipografi).
3. *Transformation*
   Pencarian fitur-fitur yang berguna untuk mempresentasikan data bergantung kepada *goal* yang ingin dicapai. Merupakan proses transformasi pada data yang telah dipilih sehingga data tersebut sesuai untuk proses *data mining.*Proses ini merupakan proses kreatif dan sangat tergantung pada jenins atau pola informasi yang akan dicari dalam basis data.
4. *Data Mining*
   Pemilihan tugas *data mining* : pemilihan goal dari proses KDD misalnya klasifikasi, regresi, *clustering*, dan lain-lain. Pemilihan algoritma *data mining*untuk pencarian (searching). Proses *Data Mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.
5. *Interpretation Evaluation*
   Penerjemahan pola-pola yang dihasilkan dari data mining. Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari Proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.4. Tugas-tugas *Data Mining*

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu[6]:

1. Deskripsi
   Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan. Sebagai contoh, petugas pengumpul suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup professional akan sedikit didukung dalam pemilihan presiden.
2. Estimasi
   Estimasi hamper sama dengan klasifikasi, kecuali variable target lebih kea rah numeric dari pada kea rah kea rah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi.
3. Prediksi
   Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan datang di masa mendatang.
4. Klasifikasi
   Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
5. Pengklusteran
   Pengklusteran merupakan peneglompokan *record*, pengamatan ataupun memperhatikan dan membentuk kelas objek-objek yang dimiliki kemiripan. *Cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record-record* dalam *cluster*lain.
6. Asosiasi
   Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.5. Algoritma K-*Means*

Algoritma *K-Means* merupakan algoritma yang relative sederhana untuk mengklasifikasikan atau mengelompokan sejumlah besar obyek dengan atribut tertentu ke dalam kelompok-kelompok (*cluster*) sebanyak K. Pada algoritma K-Means, jumlah *cluster* K sudah ditentukan lebih dahulu[7].

"*K-Means* adalah salah satu metode pengelompokan non hirarki (sekatan) yang berusaha mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama akan dimasukan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain".

Data *clustering* menggunakan metode *K-Means* ini secara umum dilakukan dengan memerlukan tiga komponen yaitu:

1. Jumlah *Cluster*
   Seperti yang telah dijelaskan sebelumnya, *K-Means*merupakan bagian dari metode non-hirarki sehingga dalam metode ini jumlah K harus ditentukan terlebih dahulu . Jumlah *cluster* K dapat ditentukan melalui pendekatan metode hirarki.
2. *Cluster* Awal
   Cluster awal yang dipilih berkaitan dengan penentuan pusat *cluster* awal.
3. Ukuran Jarak
   Dalam hal ini, ukuran jarak digunakan untuk menempatkan observasi ke dalam *cluster* berdasarkan *centroid* terdekat. Ukuran jarak yang digunakan dalam metode K-Means adalah [d]*Euclidean.*

Adapun algoritma *K-Means* dalam pembentukan klaster sebagai berikut:

1. Misalkan diberikan matriks X = {X_{ij} } data berukuran *n x p* dengan *i=1,2,...n, j=1,2,....p* dan asumsi jumlah klaster awal K.
2. Tentukan *centroid*
3. Hitung jarak setiap objek ke setiap *centroid* dengan menggunakan jarak *Euclid* atau dapat ditulis sebagai berikut:
   J(x,c) = $\sqrt{(x_1 - c_2)^2}$……...……(1)
4. Setiap objek disusun ke *centroid* terdekat dan kumpulkan objek tersebut akan membentuk klaster.
5. Tentukan *centroid* baru dari *cluster* yang akan dibentuk, diobjek yang mana *centroid* baru diperoleh dari rata-rata setiap yang terletak pada klaster yang sama.
6. Ulangi langkah 3, jika *centroid* awal dan baru tidak sama.

Karakteristik *K-Means* dapat diringkas menjadi seperti berikut:

1. *K-Means* merupakan metode pengelompokan yang sederhana dan dapat digunakan dengan mudah.
2. Pada jenis set data tertentu, *K-Means* tidak dapat melakukan segmentasi data dengan baik dimana hasil segmentasinya tidak dapat memberikan pola kelompok yang mewakili karakteristik bentuk alami data.
3. *K-Means* bisa mengalami masalah ketika mengelompokkan data yang mengandung *outlier.*

3. Hasil dan Pembahasan

3.1 Data Pedukung Penelitian

Dengan menganalisa data-data pengiriman hasil kelapa sawit yang menumpuk dan data hanya diproses di Micosoft Excel, maka dapat dilihat permasalah selama ini yaitu menurunnya tonase setiap pengririman dikarenakan kwalitas hasil sawit yang melibatkan pegiriman ke setiap pabrik sehingga mengakibatkan setiap pegiriman mengalami jumlah penurunan tonase. Penyeleksian data – data yang ada hanya secara manual sehingga untuk mengetahui pegiriman hasil kelapa sawit mana yang paling banyak.

Dari banyaknya data yang telah terkumpul belum diterapkan pengelompokan pabrik, tonase dan kwalitas dari data pengiriman. Sehingga data yang sudah ada menjadi kurang

informative. Data pegiriman yang telah ada selama ini hanya di simpan oleh pihak KUD sebagai bentuk dokumentasi.

Dengan menganalisa data pegiriman hasil kelapa sawit yang menumpuk dan data hanya diproses di Arsip, maka dapat dilihat permasalah selama ini yaitu menumpuknya data hasil kelapa sawit tanpa mengetahui data yang sebenarnya yang terdapat di Arsip tersebut. Penyeleksian data – data yang ada hanya secara manual sehingga untuk mengetahui data pengelompokan hasil kelapa sawit. Data yang dilampirkan yaitu:

1. Nama Pabrik
   a. JPN (Jaya Palma Nusantara)
   b. APL (Anugrah Putra Langkat
   c. YAP (Yapuyra Alpa Palmindo)
   d. MPA (Megah Pusaka Andalas)
2. Tonase
   a. Mix 2 Ton
   b. Max 10 Ton
3. Kwalitas
   a. Tenera lebih tebal dari pada batoknya
   b. Tidak restan
   c. Komidel janjang di atas 8 kg
4. Kecamatan
   a.Gebang
   b Batang Serangan

Tabel 1. Data Analisa Metode *Clustering*

No	Nama pabrik	Tonase	Kwalitas	Kecamatan
1	JPN (Jaya Palma Nusantara)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Gebang
2	APL (Anugrah Putra Langkat)	Min 2 ton	Tidak restan	Batang Serangan
3	JPN (Jaya Palma Nusantara)	Max 10 ton	Komidel janjang di atas 8 kg	Gebang
4	APL (Anugrah Putra Langkat)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan
5	YAP (Yapuyra Alpa Palmindo)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan
6	JPN (Jaya Palma Nusantara)	Min 2 ton	Tidak restan	Gebang
7	APL (Anugrah Putra Langkat)	Max 10 ton	Komidel janjang di atas 8 kg	Batang Serangan
8	YAP (Yapuyra Alpa Palmindo)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan
9	MPA (Megah Pusaka Andalas)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan
10	JPN (Jaya Palma Nusantara)	Max 10 ton	Tidak restan	Gebang
11	APL (Anugrah Putra Langkat)	Min 2 ton	Komidel janjang di atas 8 kg	Batang Serangan
12	JPN (Jaya Palma Nusantara)	Max 10 ton	Tenera lebih tebal dari pada batoknya	Gebang
13	APL (Anugrah Putra Langkat)	Min 2 ton	Tidak restan	Batang Serangan
14	JPN (Jaya Palma Nusantara)	Max 10 ton	Komidel janjang di atas 8 kg	Gebang
15	APL (Anugrah Putra Langkat)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan

16	JPN (Jaya Palma Nusantara)	Max 10 ton	Tidak restan	Gebang
17	APL (Anugrah Putra Langkat)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Batang Serangan
18	YAP (Yapuyra Alpa Palmindo)	Min 2 ton	Tidak restan	Batang Serangan
19	MPA (Megah Pusaka Andalas)	Max 10 ton	Komidel janjang di atas 8 kg	Batang Serangan
20	JPN (Jaya Palma Nusantara)	Min 2 ton	Tenera lebih tebal dari pada batoknya	Gebang

1. Inisialisasi Data

Dari data yang ada maka dapat dilakukan inisialisasi data sesuai dengan kebutuhan variabel sebagai berikut:

a. Inisialisasi Kriteria Pabrik

Dari data pengiriman hasil kelapa sawit yang ada di KUD BumiPusaka terdapat beberapa nama pabrik. Berikut ini adalah tabel nama pabrik:

Tabel 2.Inisialisasi Kriteria Pabrik

Kode	Pabrik
1	JPN (Jaya Palma Nusantara)
2	APL (Anugrah Putra Langkat)
3	YAP (YapuyraAlpaPalmindo)
4	MPA (MegahPusakaAndalas)

b. Inisialisasi Kriteria Tonase

Dari data pengiriman hasil kelapa sawit yang ada di KUD BumiPusaka terdapat beberapa tonase. Berikut ini adalah tabel tonase:

Tabel 3. Inisialisasi Kriteria Tonase

Kode	Tonase
1	Min 2 ton
2	Max 10 ton

c. Inisialisasi Kriteria Kwalitas

Dari data pengiriman hasil kelapa sawit yang ada di KUD Bumi Pusaka terdapat beberapa kwalitas. Berikut ini adalah tabel kwalitas:

Tabel 4.Inisialisasi Kriteria Kwalitas

Kode	Kwalitas
1	Tenera lebih tebal dari pada batoknya
2	Tidak restan
3	Komidel janjang di atas 8 kg

d. Inisialisasi Kriteria Kecamatan

Dari data pengiriman hasil kelapa sawit yang ada di KUD BumiPusaka terdapat beberapa kecamatan. Berikut ini adalah tabel kecamatan:

Tabel 5. Inisialisasi Kriteria Kecamatan

Kode	Kecamatan
1	Gebang
2	Batang Serangan

3.2. Pembahasan

Langkah-langkah yang dilakukan untuk perhitungan data pengiriman hasil kelapa sawit dengan algoritma *K-Means* ini, agar dapat dihasilkan sebuah pengetahuan baru, mengenai berapa banyak kelompok data pengiriman hasil kelapa sawit yang terjadi yaitu kelompok pabrik, tonase dan kwalitas. Sehingga dapat diketahui hubungan terdekat antara kelompok.

3.2.1. Penentuan Jarak Pada Pengelompokan (*Clustering*)

Untuk menentukan *group* dari satu objek, pertama yang harus dilakukan adalah mengukur jarak ^{d}*Euclidean* antara dua titik atau objek X dan Y yang didefinisikan sebagai berikut:

$$dEuclidean(X,Y)\sqrt{\sum_i (X1 - Y2)^2}$$

Dengan rumus diatas maka dapat dilakukan perhitungan agar dapat menentukan jarak pada pengelompokan pengiriman hasil kelapa sawit.

3.2.2. Hasil Pembahasan Perhitungan *Group*

Setelah data di*import* ke dalam *MatLab* dengan *syntax* yang sudah di tentukan berdasarkan jarak *centriod* terkecil berada. Hasil dari perhitungan *group* dibagi menjadi 3 *group* dengan data pabrik (X), tonase (Y) dan kwalitas (Z) yaitu dapat dilihat pada tabel berikut:

Dari hasil tabel diatas dapat dijumlahkan masing-masing *group* yaitu sebagai berikut:

1. *Group* 1 sebanyak174 data
2. *Group* 2 sebanyak246 data
3. *Group* 3 sebanyak80 data

Jadi total data = 500

3.2.3. Perhitungan Jarak Objek Ke *Centroid*

Dari hasil analisa diatas proses *Replicate* ditentukan sebanyak 5 kali perulangan dimana *cluster* ditentukan sebanyak 3 (X, Y dan Z) maka total iterasi sebanyak 18 kali, hal ini menunjukan bahwa proses iterasi berhenti jika total jarak dengan iterasi sebelumnya sampai pada *group* yang tidak berubah lagi. Hasil iterasi yang diperoleh dari perhitungan jarak objek ke *centroid* menggunakan pemerograman *MatLab* adalah sebagai berikut:

3 iterations, total sum of distances = 593

3.2.4. Perhitungan *Centroid*

Adapun penentuan hasil jumlah *centroid* untuk setiap *group* adalah sebagai berikut::

Centroid 1 = total *group* 1/banyak*group* 1
C1 = 348/ 174= 2
C2 = 348 / 174 = 2
C3 = 696/ 174 = 1

Centroid 2 = total *group* 2/banyak*group* 2
C1 = 246/ 246 = 1
C2 = 246/ 246 = 1
C3 = 492/ 246= 2

Centroid 3 = total *group* 3/banyak*group* 3
C1 = 320 / 80 = 4
C2 = 80 / 80 = 1

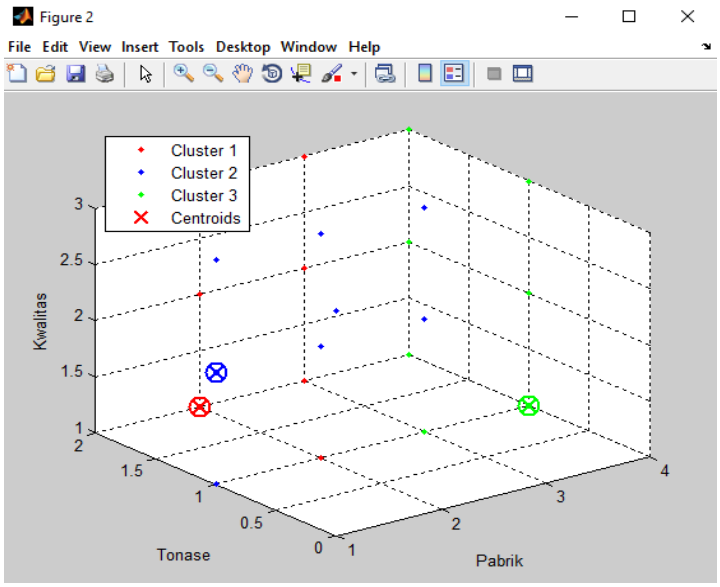
C3 = 80 / 80 = 1

Dari hasil perhitungan *centroid* diatas dapat dilihat pada tabel berikut:

Tabel 6.Perhitungan *Centroid*

Pabrik (X)	Tonase (Y)	Kwalitas (Z)	Keterangan
2	2	1	*Centroid* 1
1	1	2	*Centroid* 2
4	1	1	*Centroid* 3

Berikut hasil *cluster* 1, *cluster* 2 dan *cluster* 3 dapat dilihat pada grafik dibawah in:



Gambar 2. Grafik *Clusteing*

Pusatnya =

Group 1 = (2 2 1)

Group 2 = (1 1 2)

Group 3 = (4 1 1)

Keterangan :

1. 2 2 1

Dapat diketahui pada *cluster* 1 terdapat pada pabrik APL (Anugrah Putra Langkat) berat tonase Max 10 ton memiliki kwalitas Tenera lebih tebal pada batoknya.

2. 1 1 2

Dapat diketahui pada *cluster* 2 terdapat pada pabrik JPN (Jaya Palma Nusantara) berat tonase Min 2 ton memiliki kwalitas Tidak restan.

3. 4 1 1

Dapat diketahui pada *cluster* 3 terdapat pada pabrik MPA (Mega Pusaka Andalas) berat tonase Min 2 ton memiliki kwalitas Tenera lebih tebal pada batoknya.

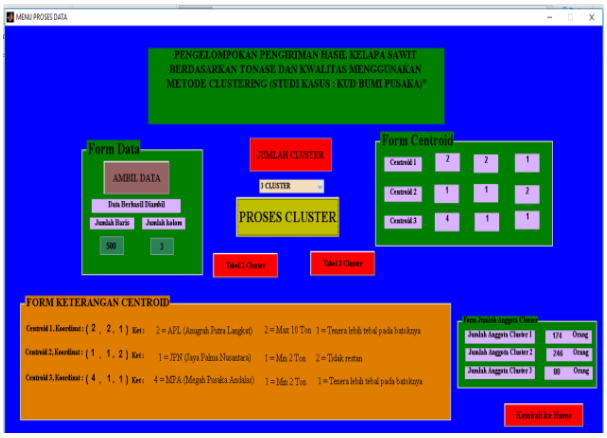
3.3. Pembahasan Antarmuka (*Interface*)

Dalam pembahasan antar muka ini akan dijelaskan mengenai hasil perancangan program yang menggunakan *GUIDE Matlab*, dan dapat dilihat sebagai berikut:

1. Menu *Home* disini menampilkan *interface* awal yang berisi menu *Proses clustering* dan Informasi. Berikut gambar *interface* Menu utama.:

Gambar 3. Halaman Utama

2. Menu Proses *Clustering*, di sini akan terlihat keseluruhan proses *data mining* sampai pada pemunculan grafik dan keterangan *centroid* sebagai hasil dari perhitungan dengan metode *clustering* menggunakan algortima *k-means*. Untuk lebih jelasnya dapat dilihat pada gamabar berikut:



Gambar 4. Proses *Clustering*

3. Pada Menu Informasi akan muncul gambar data pabrik, tonase dan kwalitas adalah sebagai berikut:

Gambar 5. Hasil *Clustring*

4. Kesimpulan

Berdasarkan *clustering* pengiriman hasil kelapa sawit di dapat, maka dapat diambil suatu kesimpulan yaitu sebagai berikut:

1 Dari pengujian yang dilakukan menggunakan metode *clustering* dengan algoritma *K-Means* ini, dapat diketahui kelompok pabrik, kelompok tonase dan kelompok kwalitas mana saja yang memiliki kelompok paling tinggi dan paling sering muncul pada pengelompokkan pengiriman hasil kelapa sawit.

2. Dari 500 data pengiriman hasil kelapa sawit diperoleh 3 *Group*, di mana *Group 1* berjumlah 174 data, *Group* 2 berjumlah 246 data dan *Group* 3 berjumlah 80 data
   1. Dapat diketahui pada *cluster* 1 (2 2 1) terdapat pada pabrik APL (Anugrah Putra Langkat) berat tonase Max 10 ton memiliki kwalitas Tenera lebih tebal pada batoknya.
   2. Dapat diketahui pada *cluster* 2 (1 1 2) terdapat pada pabrik JPN (Jaya Palma Nusantara) berat tonase Min 2 ton memiliki kwalitas Tidak restan.
   3. Dapat diketahui pada *cluster* 3 (4 1 1) terdapat pada pabrik MPA (Mega Pusaka Andalas) berat tonase Min 2 ton memiliki kwalitas Tenera lebih tebal pada batoknya.

Daftar Pustaka

[1] Isy Karima Fauzia, Budi Arif Dermawan, and Tesa Nur Padilah, "Penerapan K-Means Clustering pada Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kabupaten Karawang," *J. Sist. dan Inform.*, vol. 15, no. 1, 2020, doi: 10.30864/jsi.v15i1.350.

[2] M. F. I. Al-Rizki, I. Widaningrum, and G. A. Buntoro, "Prediksi Penyebaran Penyakit TBC dengan Metode K-Means Clustering Menggunakan Aplikasi Rapidminer," *JTERA (Jurnal Teknol. Rekayasa)*, vol. 5, no. 1, 2020, doi: 10.31544/jtera.v5.i1.2019.1-10.

[3] J. Wandana, S. Defit, and S. Sumijan, "Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means," *J. Inf. dan Teknol.*, 2020, doi: 10.37034/jidt.v2i4.73.

[4] W. Lestari, F. Fatoni, and H. Hutrianto, "IMPLEMENTASI DATA MINING UNTUK KARTU INDONESIA SEHAT BAGI MASYARAKAT KURANG MAMPU MENGGUNAKAN METODE CLUSTERING PADA DINAS SOSIAL KOTA PALEMBANG," *J. Nas. Ilmu Komput.*, vol. 1, no. 4, 2020, doi: 10.47747/jurnalnik.v1i4.163.

[5] S. Butsianto and N. T. Mayangwulan, "Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 3, no. 3, 2020, doi: 10.32672/jnkti.v3i3.2428.

[6] R. Ordila, R. Wahyuni, Y. Irawan, and M. Yulia Sari, "PENERAPAN DATA MINING UNTUK PENGELOMPOKAN DATA REKAM MEDIS PASIEN BERDASARKAN

JENIS PENYAKIT DENGAN ALGORITMA CLUSTERING (Studi Kasus : Poli Klinik PT.Inecda)," *J. Ilmu Komput.*, vol. 9, no. 2, 2020, doi: 10.33060/jik/2020/vol9.iss2.181.

[7] A. Wibowo and A. R. Handoko, "Segmentasi Pelanggan Ritel Produk Farmasi Obat Menggunakan Metode Data Mining Klasterisasi Dengan Analisis Recency Frequency Monetary (RFM) Termodifikasi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, 2020, doi: 10.25126/jtiik.2020702925.