

Implementasi K-Medoids Clustering Dalam Pengelompokan Harga 8 Jenis Minyak Goreng

Fina Nasari^{1*}, Andri Nofiar.Am²

¹²Teknik Informatika, Politeknik Kampar

Jl. Tengku Muhammad KM 2. Batu Belah, Kec. Bangkinang, Kab. Kampar, Indonesia

e-mail: finanasari@gmail.com¹, andrinofiar90@gmail.com²

Received : September, 2023

Accepted : November, 2023

Published : December, 2023

Abstract

In everyday life, cooking oil has become a necessity with international sales prices varying depending on the quality and type. The types of cooking oil sold in the international market are very diverse, including coconut, olive, palm kernel oil, palm oil, peanuts, rapeseed, soybeans and sunflower. Therefore, it is necessary to group data on selling prices of cooking oil on the international market to get the best grouping of cooking oil. The data used in this research is historical price data of 8 edible oils kaggle August 1992 to July 2022. Data grouping in this research uses the k-medoids algorithm. The k-medoids algorithm produces a more balanced group, better performance and accuracy than other algorithms. The aim of this research is that the k-medoids algorithm is able to group cooking oil price data into 4 group models, namely group models 2, 3, 4 and 5 and obtain the best group model based on the dbi value. The research results showed that the cooking oil price data was successfully grouped into group 2, 3, 4 and 5 models with the best group based on the lowest dbi performance value being the group 2 model with a dbi value of 0,000 and olive oil being the cooking oil with the highest price in the world while 7 types other cooking oils have more or less the same price (in the same price group).

Keywords: K-Medoids, Clustering, Edible Oils

Abstrak

Dalam kehidupan sehari-hari, minyak goreng menjadi kebutuhan dengan harga penjualan internasional yang beragam dipengaruhi oleh kualitas dan jenisnya. Jenis minyak goreng yang dijual dipasar internasional sangat beragam diantaranya minyak kelapa, zaitun, inti kelapa sawit, kelapa sawit, kacang tanah, rapeseed, kedelai hingga bunga matahari. Oleh karena itu perlu adanya pengelompokan data harga penjualan minyak goreng pada pasar internasional untuk mendapatkan pengelompokan minyak goreng terbaik. Data yang digunakan dalam penelitian ini adalah *data historical price data of 8 edible oils kaggle* agustus 1992 s.d juli 2022. Pengelompokan Data pada penelitian ini menggunakan algoritma k-medoids. Algoritma *k-medoids* menghasilkan kelompok yang lebih seimbang, *performance*, akurasi yang lebih baik dari algoritma yang lain. Tujuan penelitian ini, algoritma *k-medoids* mampu mengelompokkan data harga minyak goreng dalam 4 model kelompok yaitu model kelompok 2, 3, 4 dan 5 dan mendapatkan model kelompok terbaik berdasarkan nilai dbi. Hasil penelitian diperoleh bahwa data harga minyak goreng berhasil di kelompokkan dalam model kelompok 2, 3, 4 dan 5 dengan kelompok terbaik berdasarkan nilai *performance* dbi terendah adalah model kelompok 2 dengan nilai dbi 0.000 dan *olive oil* menjadi minyak goreng dengan harga tertinggi dunia sedangkan 7 jenis minyak goreng lainnya memiliki harga yang kurang lebih sama (dalam kelompok harga yang sama).

Kata Kunci: K-Medoids, Clustering, Minyak Goreng

1. PENDAHULUAN

Minyak goreng merupakan salah satu kebutuhan pokok dalam kehidupan sehari-hari. Penggunaan minyak goreng baik untuk kebutuhan usaha maupun kebutuhan pribadi. Banyak produk turunan dari minyak goreng diantaranya produk makanan, kesehatan, kecantikan hingga kebersihan. Terdapat banyak jenis minyak goreng yang beredar dipasar internasional diantaranya minyak kelapa, minyak zaitun, minyak inti kelapa sawit, minyak kelapa sawit, minyak kacang tanah, minyak *repeseed*, minyak kacang kedelai dan minyak biji bunga matahari. Harga minyak goreng beragam berdasarkan sumber bahan utama minyak hingga kualitas.

Indonesia merupakan salah satu negara eksportir minyak goreng. Minyak goreng hasil produksi Indonesia diantaranya minyak kelapa, minyak inti kelapa dan minyak kelapa sawit. Indonesia dan philipina menjadi produsen kelapa terbesar didunia tahun 2014-2018 berdasarkan data FAO, jumlah pangsa pasar sebesar 76,86%. Pada tahun 2014 pangsa pasar Indonesia sebesar 35,31%, 41,55% untuk Philippina, 8,58% untuk Netherland (Belanda), Malaysia sebesar 8,59%; 1,44% untuk Amerika Serikat serta sisanya sebesar 4,53% negara lainnya. Selain pada tahun 2018 produsen kelapa sawit terbesar di dunia adalah indonesia [1] yang memiliki jutaan hektar yang tersebar diseluruh indonesia[2]. Berdasarkan data tersebut Indonesia berpotensi menjadi produsen terbesar minyak kelapa, inti kelapa dan kelapa sawit.

Untuk melihat *cluster* harga minyak goreng dunia khususnya minyak goreng hasil produksi Indonesia, perlu dibuat sebuah pengelompokan data. Hasil pengelompokan tersebut untuk melihat posisi harga minyak goreng yang diproduksi oleh Indonesia. Berdasarkan penelitian sebelumnya *Clustering* memiliki kemampuan dalam pengelompokan data sejenis dikelompokkan pada satu kelompok, kemudian data yang berbeda pada kelompok lainnya[3]. Salah satu algoritma clustering adalah Algoritma *k-medoids* yang mampu mengelompokkan data berdasarkan objek sebagai perwakilan (*medois*)[4][5].

Pada penelitian sebelumnya, penulis melakukan pengelompokan produksi kelapa sawit di provinsi riau. Pengelompokan yang

dilakukan dilakukan dengan menggunakan algoritma *k-means*, *k-medoids* dan membandingkan algoritma *k-meands* dan *k-medoids*. Hasil dari penelitian tersebut, algoritma *k-means* dan *k-medoids* mampu mengelompokkan data sebanyak jumlah *cluster* yang akan dibangun. Sedangkan hasil perbandingan *performance* dari kedua algoritma tersebut diperoleh *performance* algoritma *k-medoids* lebih baik dari *k-means* [6][7][8].

Pada penelitian untuk topik pengelompokan performa keuangan, kelompok yang dihasilkan dengan menggunakan algoritma *k-means* dan *k-medois* berbeda dimana hasil yang diperoleh pada algoritma *k-means* menghasilkan kelompok dengan variasi yang sangat besar dan hasil yang diperoleh setelah pengelompokan lebih besar, sedangkan dengan penggunaan algoritma *k-medoids* hasil pengelompokan lebih seimbang[9]. Jika dibandingkan dengan algoritma *k-means*, algoritma *k-medoids* memiliki unjuk kerja yang lebih baik [8]. Selain itu algoritma *k-medoids* adalah algoritma yang paling menonjol sebagai penyempurna algoritma sebelumnya, yaitu *k-means*[10]. Jika dilihat berdasarkan hasil *performance* nilai *Devies Building Indeks* (DBI) algoritman *k-medoids* lebih baik dari algoritma *k-means*[11]. Algoritma *k-means* dan *k-medoids* memiliki keunggulan masing-masing terhadap karakteristik pengunjung objek wisata, namun hasil akurasi terbaik pada karakteristik gender dengan nilai akurasi 90.48% pada algoritma *k-medoids*[12].

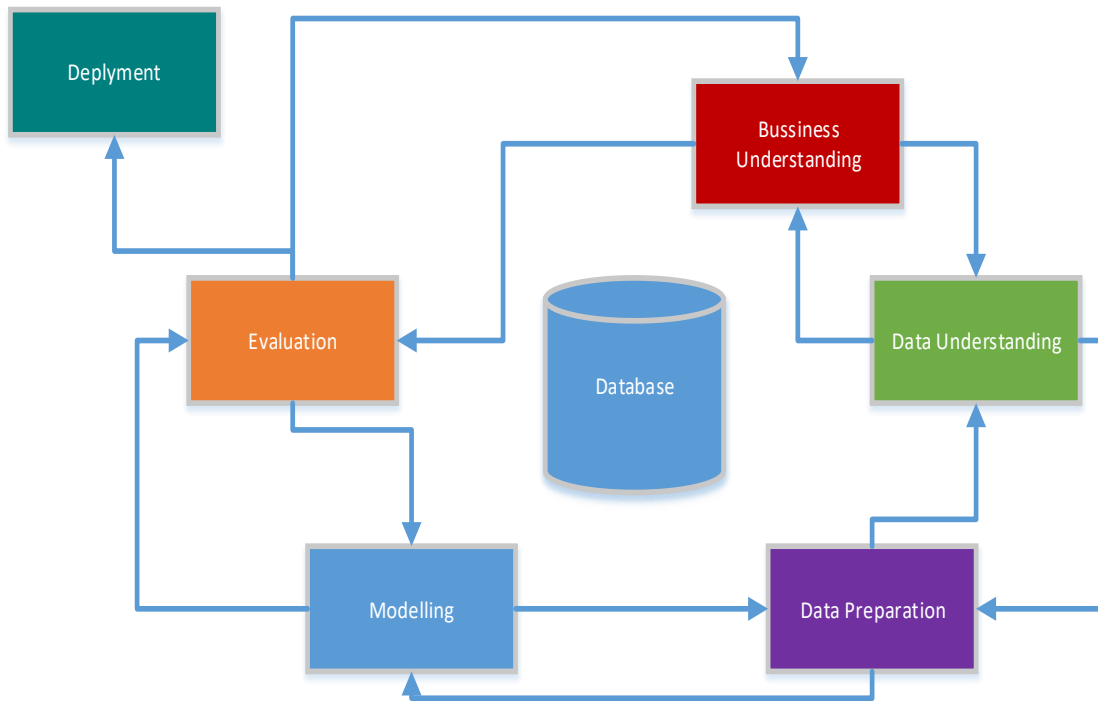
Berdasarkan hal tersebut diatas, Pengelompokan harga minyak goreng yang akan dilakukan dengan menggunakan algoritma *k-medoids*. Tujuan dari penelitian ini mengelompokkan harga 8 jenis minyak goreng kedalam beberapa kelompok dan memilih kelompok terbaik berdasarkan hasil nilai *Devies Building Indeks* (DBI). Hasil pengelompokan tersebut menjadi dasar segmentasi harga minyak goreng dunia.

2. METODE PENELITIAN

Pada penelitian ini menggunakan model *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Terdapat 6 fase dalam Model CRISP-DM yang dimulai dari pengumpulan Data

yang akan diolah, pemilihan atribut atau variable yang dibutuhkan, pembuatan model yang akan digunakan, evaluasi dan *deployment*

[13]. model CRISP-DM terlihat pada gambar 1 berikut ini:



Gambar 1. Tahapan Model CRISP-DM

Enam tahapan metode CRISP-DM diantaranya pada tahap pertama pemahaman bisnis, pada tahap kedua pemahaman data yang akan digunakan, pada tahap ketiga persiapan data yang akan digunakan, pada tahap ke empat pemodelan, pada tahap kelima proses evaluasi, dan tahapan terakhir adalah penyebaran[14].

2.1 Business Understanding

Dimulai dari pemahaman bisnis (*business understanding*) pada tahap ini minyak menjadi salah satu bahan yang sangat dibutuhkan baik untuk memenuhi kebutuhan pribadi maupun industri (usaha), banyak produk turunan minyak yang bisa dihasilkan baik produk kecantikan, kesehatan, makanan dan lain-lain. Indonesia merupakan salah satu produsen terbesar penghasil kelapa[15], inti kelapa sawit dan kelapa sawit[1]. Harga minyak terus meningkat dan beragam, oleh karena itu perlu adanya pengelompokan data minyak untuk mengetahui jenis minyak goreng dengan harga tinggi, sedang dan rendah. Hasil pengelompokan tersebut untuk melihat potensi harga jual dari produk Indonesia yaitu kelapa, inti kelapa sawit dan kelapa sawit dan

menjadikan Indonesia berpeluang untuk meningkatkan ekspor hasil memproduksi minyak kelapa, inti kelapa sawit dan kelapa sawit.

2.2 Data Understanding

Tahap berikutnya pemahaman data (*Data Understanding*) proses yang dilakukan dengan mengumpulkan data-data yang dibutuhkan dimulai dari data primer berupa data yang akan diolah berupa *data historical price data of 8 edible oils(commudity) kaggle* yang diperoleh dari data public *kaggle*. Data primer berikutnya berupa jurnal yang digunakan sebagai rujukan atau bahan bacaan dan acuan pengerjaan penelitian ini, jurnal yang digunakan adalah jurnal yang dipublikasi 8 Tahun terakhir. Selanjutnya data sekunder yang didapatkan dari website resmi atau buku-buku.

Penelitian ini menggunakan data *data historical price data of 8 edible oils(commudity)* selama 30 Tahun terakhir dimulai dari data bulan agustus tahun 1992 s.d juli 2022 sebanyak 2880 Data. Data tersebut pernah digunakan pada penelitian peramalan harga minyak goreng.

Data yang digunakan pada penelitian sebelumnya adalah data Januari 1987 s.d february 2017 dengan hasil *regresi vector* memiliki nilai yang lebih tinggi jika dibandingkan dengan *multy layer perceptron* dan *Holt Winter exponential smoothing*. Saran

untuk penelitian selanjutnya penambahan variabel pemilihan fitur untuk meningkatkan nilai akurasi[16]. Data *data historical price data of 8 edible oils(commudity)* setelah dilakukan *transpose* kolom seperti terlihat pada tabel 1 berikut:

Tabel 1. *Data historical price data of 8 edible oils (commudity)*

[Sumber: <https://www.kaggle.com/datasets/chinmayshanbhag/historical-price-data-of-8-edible-oils-commodity>]

Month - Edible Oils	August-92	September-92	October-92	June-22	July-22
Coconut Oil Price	494.00	501.00	493.00	1,700.50	1,540.53
Coconut Oil Price Change	-	1.42%	-1.60%	-6.22%	-9.41%
Olive Oil Price	3,490.25	3,342.78	2,986.24	4,030.65	4,065.02
Olive Oil Price Change	-	-4.23%	-10.67%	-1.66%	0.85%
Palm Kernel Oil Price				1,554.50	1,301.05
Palm Kernel Oil Price Change				-14.17%	-16.30%
Palm Oil Price	382.00	391.00	396.00	1,501.10	1,056.64
Palm Oil Price Change	-	2.36%	1.28%	-12.57%	-29.61%
Peanut Oil Price	600.00	562.00	565.00	2,146.19	2,146.19
Peanut Oil Price Change	-	-6.33%	0.53%	0.00%	0.00%
Rapeseed Oil Price				2,000.53	1,737.34
Rapeseed Oil Price Change				-9.82%	-13.16%
Soybean Oil Price	409.00	421.00	418.00	1,751.76	1,533.40
Soybean Oil Price Change	-	2.93%	-0.71%	-10.76%	-12.47%
Sunflower Oil Price				1,884.59	1,556.87
Sunflower Oil Price Change				-9.36%	-17.39%

2.3 Data Preparation

Tahap berikutnya persiapan Data (*Data Preparation*) proses yang dilakukan adalah dengan mempersiapkan data untuk dapat diproses dengan metode *k-medoids* yaitu dengan melakukan pemilihan atribut yang akan dijadikan variable penelitian, menghilangkan

missing data dan mengganti data yang kosong dengan nilai 0. Hal ini dilakukan karena algoritma *k-medoids* kurang sensitive terhadap *noises* dan *outliers*[10]. atribut yang akan digunakan pada penelitian ini terlihat pada tabel 2 berikut:

Tabel 2. Variabel yang dipilih untuk proses pengolahan algoritma *k-medoids*

Month - Edible Oils	Aug-92	Sep-92	Oct-92	Jun-22	Jul-22
Coconut Oil Price	494.00	501.00	493.00	1,700.50	1,540.53
Olive Oil Price	3,490.25	3,342.78	2,986.24	4,030.65	4,065.02
Palm Kernel Oil Price				1,554.50	1,301.05
Palm Oil Price	382.00	391.00	396.00	1,501.10	1,056.64
Peanut Oil Price	600.00	562.00	565.00	2,146.19	2,146.19
Rapeseed Oil Price				2,000.53	1,737.34

Soybean Oil Price	409.00	421.00	418.00	1,751.76	1,533.40
Sunflower Oil Price				1,884.59	1,556.87

Berdasarkan tabel 2 variabel yang dipilih ada sebanyak 8 variabel jenis minyak goreng diantaranya *coconut oil price, olive oil price, palm kernel oil price, palm oil price, peanut oil price, repeseed oil price, soybean oil price* dan *sunflower oil price*. Untuk variable yang tidak digunakan ada 8 variabel juga yang berisi perubahan harga untuk setiap jenis minyak goreng. Sehingga variabel yang dipilih adalah

harga minyak goreng tidak termasuk harga perubahannya untuk setiap bulan. Untuk tahap selanjutnya adalah menghilangkan *missing* dan mengganti *value* kosong dengan angka *null* agar data dapat diproses dengan algoritma *k-medoids*. Data hasil *preprocessing* dan siap untuk dianalisis dengan algoritma *k-medoids* terlihat pada tabel 3 berikut:

Tabel 3. Data hasil *prapocessing*

Month - Edible Oils	Aug-92	Sep-92	Oct-92	Jun-22	Jul-22
Coconut Oil Price	494	501	493	1700,50	1540,53
Olive Oil Price	3490,25	3342,78	2986,24	4030,65	4065,02
Palm Kernel Oil Price	0	0	0	1554,50	1301,05
Palm Oil Price	382	391	396	1501,10	1056,64
Peanut Oil Price	600	562	565	2146,19	2146,19
Rapeseed Oil Price	0	0	0	2000,53	1737,34
Soybean Oil Price	409	421	418	1751,76	1533,4
Sunflower Oil Price	0	0	0	1884,59	1556,87

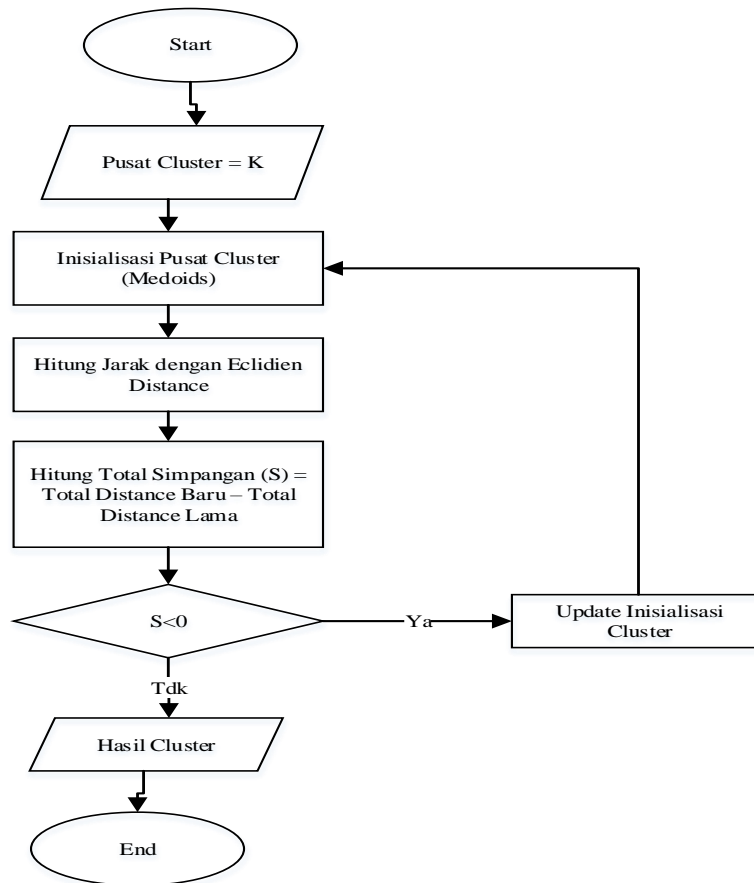
Total *value* bernilai *null* yang diganti dengan nilai 0 sebanyak 274 diantaranya 41 *palm kernel oil*, 114 *repeseed oil price* dan *sunflower oil price* sebanyak 119.

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \dots \dots \dots (1)$$

2.4 Modelling

Tahap selanjutnya Pemodelan (*Modelling*). Pemodelan jumlah *cluster* yang akan diujicoba harus dilakukan diawal sebelum dilakukan pemrosesan menggunakan algoritma *k-medoids*, Pemodelan tersebut nantinya akan dievaluasi mana model kelompok terbaik[17]. Pemodelan yang akan diujicoba yaitu model 2 kolompok, 3 kelompok, 4 kelompok dan 5 kelompok. Tahapan algoritma *k-medoids* dimulai dengan proses Inisialisasi jumlah kluster *k* dan masing-masing *cluster* ditentukan pusat *clusternya* (*medoids*). Proses dilanjutkan dengan mengumpulkan setiap data (objek) kelompok terdekat(*cluster* terdekat) berdasarkan hasil perhitungan jarak *euclidian distance* berikut ini:

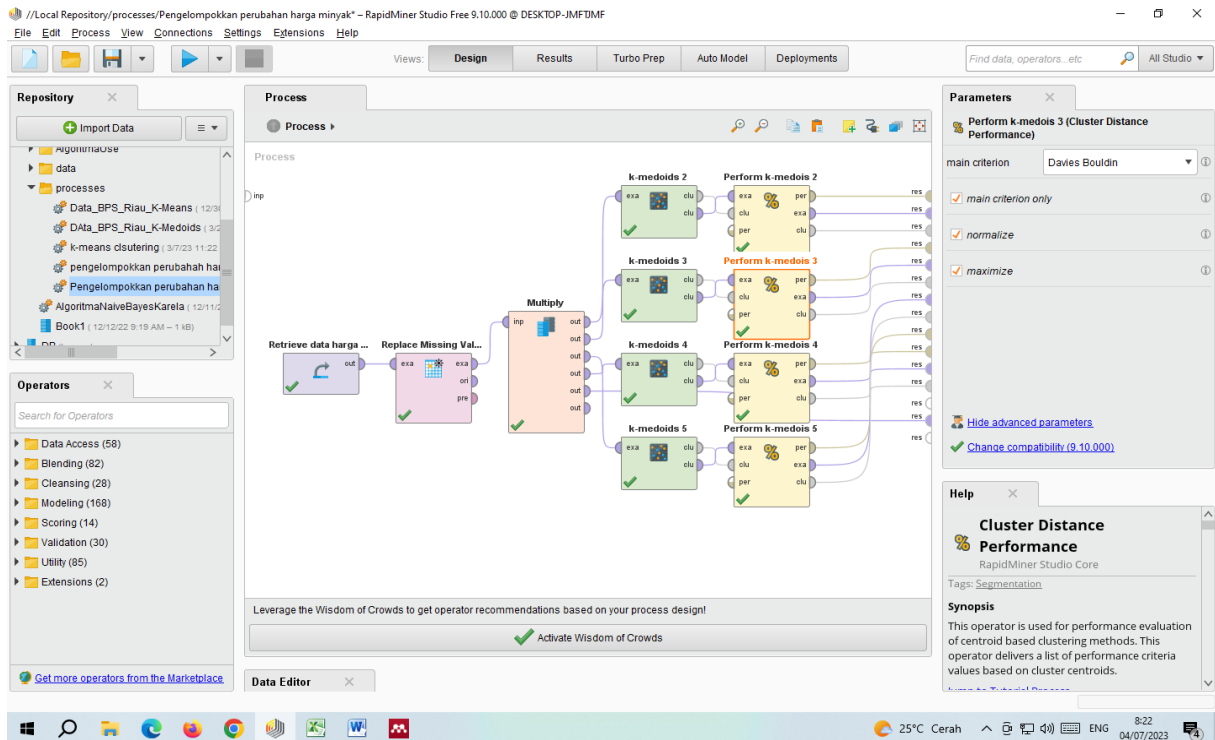
tahap algoritma *k-medoids* dilanjutkan dengan memilih acak objek pada setiap *cluster* untuk dijadikan sebagai calon *medoids* baru dan dihitung rentang(jarak) pada setiap data di setiap cluster dengan menggunakan *medoids* baru. Selanjutnya dihitung nilai simpangan (*S*) berdasarkan selisih total *distance* baru dengan total *distance* lama. apabila hasil nilai *S* kurang dari 0, lakukan penukaran objek dengan data *cluster* dalam penentuan sekumpulan nilai *k* untuk objek baru sebagai *medoids*. Lakukan proses perulangan sampai tidak ditemukan lagi perubahan *medois*, tahapan perulangan mulai dari Langkah ke tiga sampai Langkah ke lima, hasil perhitungan terakhir akan menjadi *cluster* yang telah dilengkapi anggota setiap *cluster*[8]. Bentuk step-step pengerjaan algoritma *k-medoid* yang akan dilakukan seperti terlihat pada gambar 2:



Gambar 2. Tahapan Algoritma *k-medoids*[8]

Terdapat kesamaan antara Algoritma *k-medoids* dengan algoritma *k-means* yang membedakannya adalah penggunaan objek untuk sebagai *medoids*[4][18]. Proses penerapan algoritma *k-medoids* menggunakan *tools rapid miner*. *Rapid miner* dapat membantu dalam menyelesaikan banyak masalah terkait dengan pengukuran dan memiliki keunggulan dalam membandingkan hasil perhitungan manual dan menggunakan alat[19]. Proses yang dilakukan dimulai dari ekspor data selanjutnya *replace missing value* untuk membuang data-data yang kosong selanjutnya dihubungkan dengan *multiply* agar data dapat digunakan untuk beberapa kelompok secara bersamaan, dilanjutkan dengan membuat model kelompok yang akan dibangun yaitu model dua kelompok (*cluster*),

model tiga kelompok (*cluster*), model empat kelompok (*cluster*) dan model lima kelompok (*cluster*). Untuk masing-masing kelompok *max runs* 10 iterasi dan *max optimization steps* 100. Kelompok yang dihasilkan pada proses sebelumnya kemudian dihitung nilai *performance* setiap kelompok dengan *cluster distance performa* dengan kriteria utama nilai DBI (*Devies Building Indeks*) berdasarkan *main criterion only, normalize dan maximize* untuk mendapatkan nilai DBI positif. Bentuk pemodelan algoritma *k-medoids* dalam *tools rapid miner* seperti terlihat pada gambar 3 berikut:



Gambar 3. Tahapan Algoritma *k-medoids* pada tools *rapid miner*

2.5

Tahap selanjutnya *Evaluasi*, proses yang dilakukan dengan mengevaluasi hasil pengolahan data berdasarkan pemodelan yang dilakukan yaitu model dua kelompok, model tiga kelompok, model empat kelompok dan model 5 kelompok dengan *performance Davies Building Indeks* (DBI). Tujuannya melakukan pengukuran menggunakan *Davies Building Indeks* (DBI) untuk memaksimalkan jarak antar *cluster*[20]. Hasil *Davies Building Indeks* (DBI) yang diperoleh memiliki nilai yang kecil(rendah), maka pengelompokan (*Cluster*) yang diperoleh semakin baik[21]. Tahapan perhitungan nilai DBI dimulai dari menghitung menghitung nilai *Sum of square within cluster* (SSM) untuk mendapatkan kohesi dari *cluster* ke-*l*, tahapan dilanjutkan dengan menghitung nilai *sum of square between cluster*(SSB) untuk mengetahui separasi antar *cluster* dilanjutkan dengan mencari *rasio* dan terakhir menghitung nilai *Davies Building Indeks* (DBI)[21]. Nilai DBI mengevaluasi *cluster* dengan melihat mutu dan range jarak antar anggota *cluster*[22]. Semakin rendah hasil nilai dbi yang didapatkan (non-negatif ≥ 0), *cluster* yang dihasilkan akan semakin baik[23].

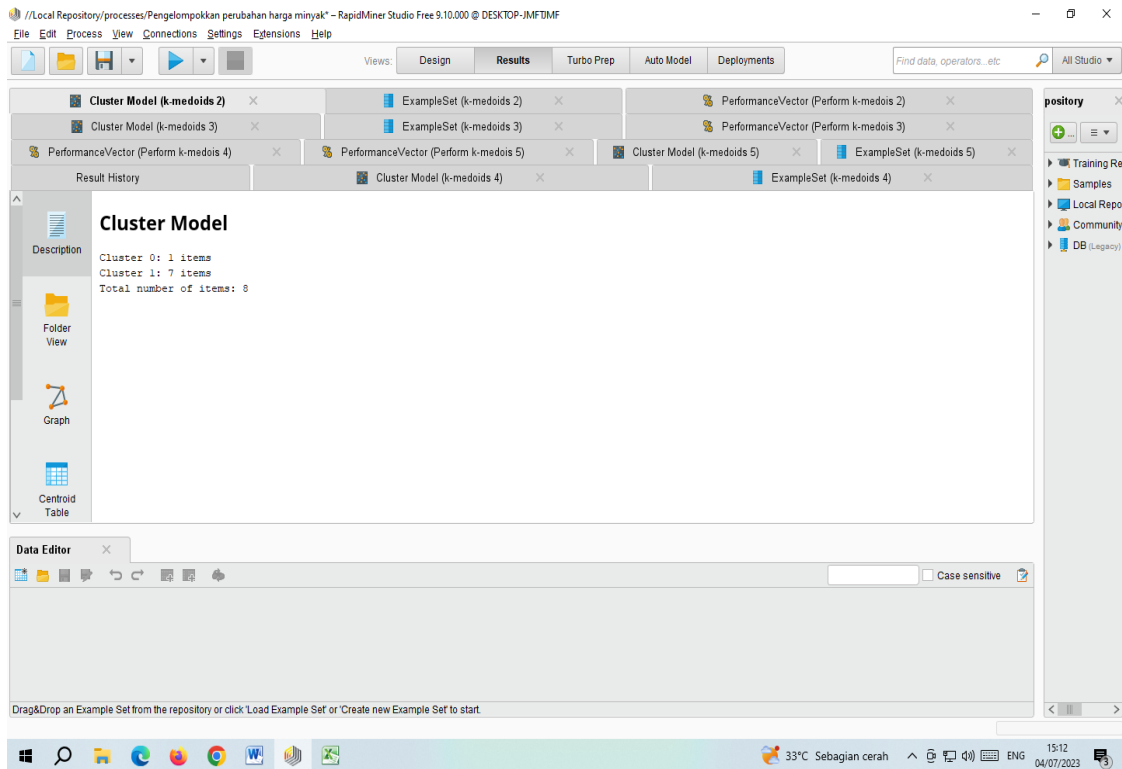
Evaluasi

2.6 Penyebaran(*Deployment*)

Tahap terakhir yaitu penyebaran (*Deployment*), proses yang dilakukan adalah melakukan publikasi hasil penelitian pada jurnal yang bereputasi/terakreditasi *sinta*, agar hasil penelitian terkait dengan segmentasi(pengelompokan) harga minyak goreng dunia dapat digunakan menjadi referensi pada penelitian selanjutnya atau referensi dalam pembuatan kebijakan atau ekspor minyak goreng khususnya minyak goreng yang diproduksi di Indonesia.

3. HASIL DAN PEMBAHASAN

Algoritma *k-medoids* digunakan pada Proses pengolahan data dan pengujiannya menggunakan *aplikasi open sources rapid miner* dengan jumlah *cluster* yang akan dibuat adalah model dua kelompok, model tiga kelompok, model empat kelompok dan model lima kelompok. Setelah dilakukan pengelompokan data berdasarkan model yang ada, selanjutnya menghitung *performance* nilai dbi. Hasil *cluster* model dua kelompok terlihat pada gambar 4 berikut:



Gambar 4. Hasil *cluster* untuk model 2 Kelompok

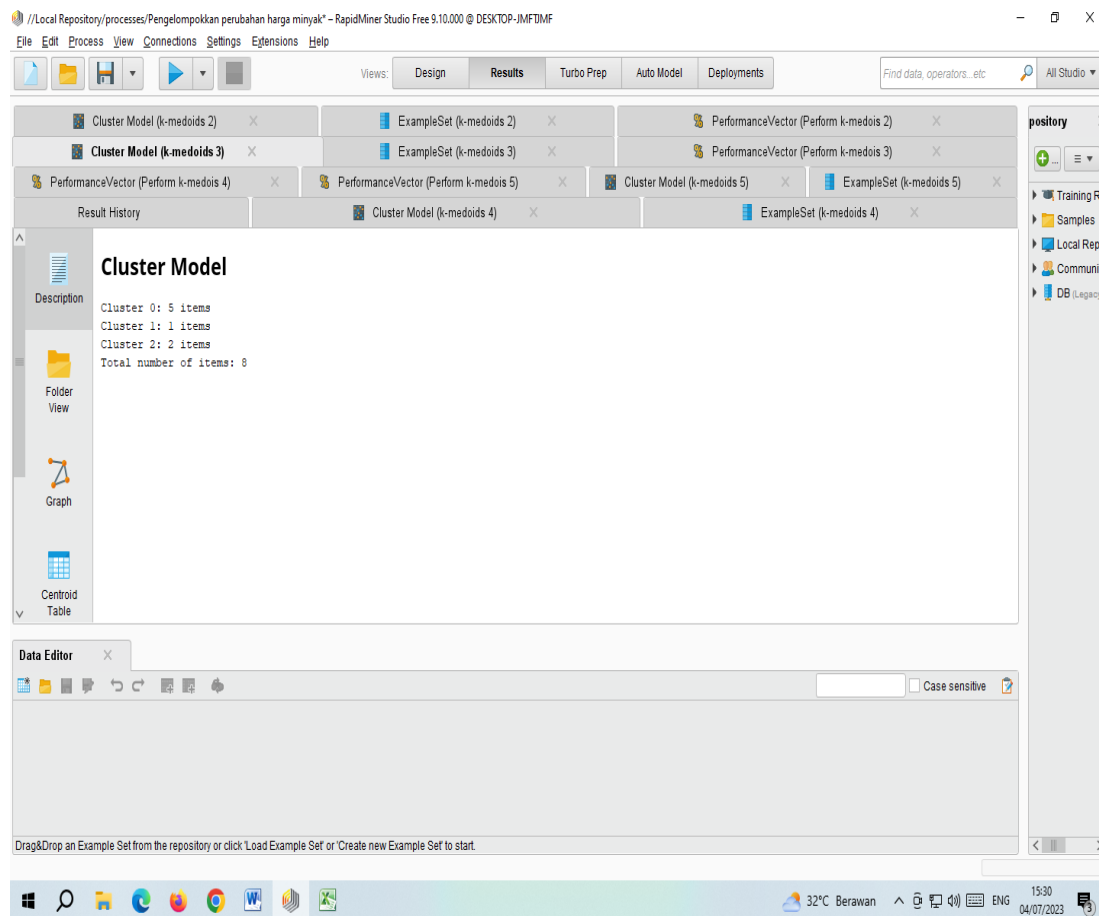
Berdasarkan *cluster* yang dihasilkan dari model kelompok dua terdapat 1 item pada *cluster* pertama (0) dan 7 item pada *cluster* kedua(1). sebaran data pada setiap kelompok dapat dilihat pada tabel 4 berikut:

Tabel 4. Penyebaran Data pada model kelompok 2

Jenis Minyak	Kelompok
Coconut Oil Price	cluster_1
Olive Oil Price	cluster_0
Palm Kernel Oil Price	cluster_1
Palm Oil Price	cluster_1

Jenis Minyak	Kelompok
Peanut Oil Price	cluster_1
Rapeseed Oil Price	cluster_1
Soybean Oil Price	cluster_1
Sunflower Oil Price	cluster_1

Cluster yang dihasilkan dari model tiga kelompok yaitu 5 item pada *cluster pertama* (0), 1 items pada *cluster kedua* (1) dan 2 items pada *cluster ketiga* (2). Gambar hasil *cluster* model tiga kolompok dapat dilihat pada gambar 5 berikut:



Gambar 5. Hasil cluster untuk model 3 Kelompok

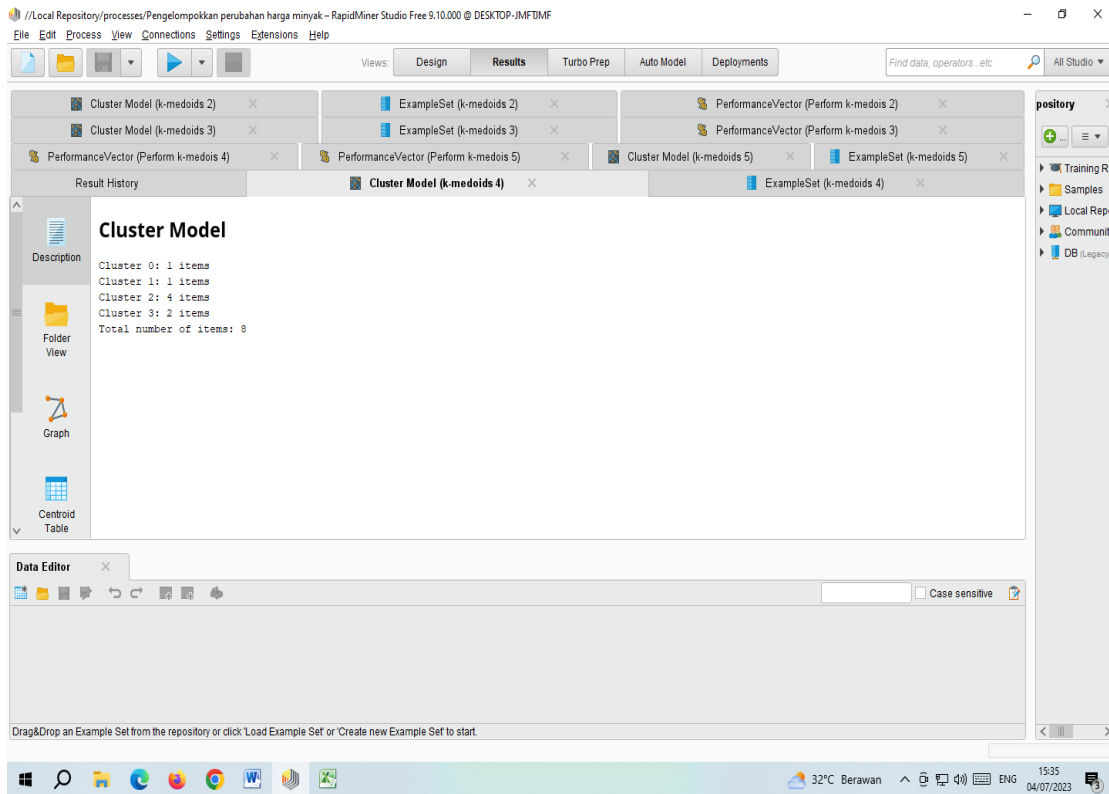
Penyebaran data model kelompok 3 pada masing-masing cluster terlihat pada tabel 5 berikut ini:

Tabel 5. Penyebaran data pada model kelompok 3

Jenis Minyak	Cluster
Coconut Oil Price	cluster_0
Olive Oil Price	cluster_1
Palm Kernel Oil Price	cluster_0
Palm Oil Price	cluster_0
Peanut Oil Price	cluster_0

Jenis Minyak	Cluster
Rapeseed Oil Price	cluster_2
Soybean Oil Price	cluster_0
Sunflower Oil Price	cluster_2

Model empat kelompok menghasilkan cluster diantaranya 1 items pada cluster pertama (0), 1 items pada cluster kedua (1), 4 items pada cluster ketiga (2) dan 2 items pada cluster keempat(3). Gambar hasil cluster model empat kelompok dapat dilihat pada gambar 6 berikut:



Gambar 6. Hasil *cluster* untuk model 4 Kelompok

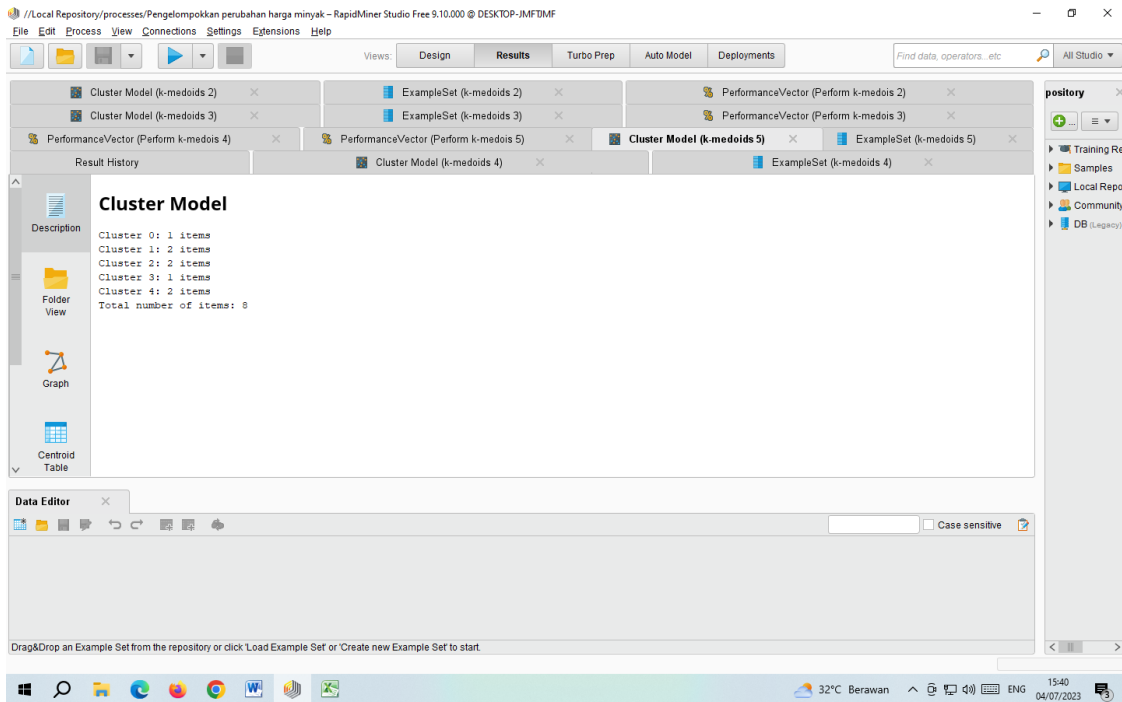
Penyebaran data model empat kelompok pada setiap kelompok terlihat pada tabel 6 berikut ini:

Jenis Minyak	Cluster
Rapeseed Oil Price	cluster_3
Soybean Oil Price	cluster_2
Sunflower Oil Price	cluster_3

Tabel 6. Penyebaran data pada model kelompok 4

Jenis Minyak	Cluster
Coconut Oil Price	cluster_2
Olive Oil Price	cluster_1
Palm Kernel Oil Price	cluster_2
Palm Oil Price	cluster_2
Peanut Oil Price	cluster_0

Cluster yang dihasilkan dari model lima kelompok yaitu 1 item pada *cluster* pertama (0), 2 items pada *cluster* kedua (1), 2 items pada *cluster* ketiga (2), 1 items pada *cluster* keempat (3) dan 2 items pada *cluster* kelima (4). Gambar hasil *kelompok* model lima kelompok terlihat pada gambar 7 berikut:



Gambar 7. Hasil *cluster* untuk model 5 Kelompok

Penyebaran data model lima kelompok pada setiap kelompok (*cluster*) terlihat pada tabel 7 berikut ini:

Tabel 7. Penyebaran data pada model kelompok 5

Jenis Minyak	Cluster
Coconut Oil Price	cluster_1
Olive Oil Price	cluster_0
Palm Kernel Oil Price	cluster_1
Palm Oil Price	cluster_2
Peanut Oil Price	cluster_3
Rapeseed Oil Price	cluster_4
Soybean Oil Price	cluster_2
Sunflower Oil Price	cluster_4

Hasil *performance* untuk setiap kelompok dilihat dari nilai DBI terlihat pada tabel 8 berikut:

Tabel 8. Nilai DBI(*davies boulding index*) pada masing-masing model kelompok

Model Kelompok	Nilai DBI
Medol Kelompok 2	0,000

Model Kelompok	Nilai DBI
Medol Kelompok 3	0,002
Medol Kelompok 4	0,001
Medol Kelompok 5	0,001

Nilai DBI yang diperoleh pada model kelompok 2 adalah 0.000, pada model kelompok 3 adalah 0.002 pada model kelompok 4 dan 5 memiliki nilai dbi yang sama yaitu 0.001. Nilai DBI yang semakin rendah akan menghasilkan kelompok dengan unjuk kerja semakin bagus [16]. Semakin rendah hasil nilai dbi yang didapatkan ($\text{non-negatif} \geq 0$), *cluster* yang dihasilkan akan semakin baik[23]. Berdasarkan penelitian komparasi k-means dan k-medoids dalam pengelompokkan susu segar Indonesia semakin rendah hasil nilai DBI maka kelompok yang dihasilkan akan semakin bagus [19]. Berdasarkan hal tersebut, maka model kelompok 2 adalah model kelompok terbaik dengan nilai dbi 0.000 dan memenuhi syarat nilai dbi terbaik yaitu bernilai positif dan ≥ 0 .

4. KESIMPULAN

Kesimpulan dari penelitian yang dilakukan adalah bahwa metode pengelompokkan data dengan menggunakan algoritma *k-medoids* berhasil dilakukan pada data *data historical price data of 8 edible oils (community)*, dalam

empat model kelompok yaitu model dua kelompok model tiga kelompok model empat kelompok dan model lima kelompok. Dilihat dari hasil *performance* nilai dbi diperoleh model kelompok terbaik adalah model kelompok 2 dengan nilai dbi terendah dan bernilai positif (non-negatif)[23], yaitu 0.000. *Cluster* yang dihasilkan dari model kelompok 2 pada *cluster* 0 dengan *olive oil price* (minyak zaitun) sebagai anggota kelompok dengan harga Tinggi, sedangkan pada *cluster* 1 terdapat *coconut oil price* (minyak kelapa), *palm kernel oil price* (minyak inti kelapa sawit), *palm oil price* (minyak kelapa sawit), *peanut oil price* (minyak kacang tanah), *soybean oil price* (minyak kacang kedelai) *repeseed oil price* dan *sunflower oil price* (minyak biji bunga matahari). Berdasarkan hasil kelompok tersebut *olive oil* (minyak zaitun) menjadi minyak dengan harga tertinggi dunia, sedangkan jenis minyak lainnya dengan harga yang kurang lebih sama.

Saran untuk penelitian ini adalah dengan menerapkan model *cluster* lainnya sebagai pertimbangan untuk hasil pengelompokan yang lebih baik. Untuk model *cluster* bisa dengan menggunakan algoritma *clustering* lainnya dengan atribut analisa yang lainnya, berdasarkan waktu dan jenis minyak.

PERNYATAAN PENGHARGAAN

Penulis mengucapkan terima kasih kepada *Kaggle* yang telah menyediakan data penelitian secara publik, sehingga penulis dapat melakukan penelitian dengan menggunakan data yang dibagikan.

Penulis mengucapkan terima kasih kepada Program Studi D3 Teknik Informatika yang telah memberikan kesempatan dalam pelaksanaan penelitian pengelompokan data harga minyak goreng dan P3M Politeknik Kampar yang telah memberikan dana dalam program hibah internal dalam pengerjaan penelitian ini.

DAFTAR PUSTAKA

- [1] D. Haryadi, "Penerapan Algoritma K-Means Clustering Pada Produksi Perkebunan Kelapa Sawit Menurut Provinsi," *J. Informatics Commun. Technol.*, vol. 3, no. 1, pp. 50–64, 2021, doi: 10.52661/j_ict.v3i1.71.
- [2] A. T. dan E. M. Faiq Husain Pratama, "Data mining k-medoids dan k-means untuk pengelompokan potensi produksi

kelapa sawit di indonesia," vol. 07, pp. 1294–1310, 2022.

- [3] A. Nofiar, S. Defit, and Sumijan, "Penentuan Mutu Kelapa Sawit Menggunakan Metode K-Means Clustering," *J. KomtekInfo*, vol. 5, no. 3, pp. 1–9, 2019, doi: 10.35134/komtekinfo.v5i3.26.
- [4] A. upi Fitriyadi, "Analisis Algoritma K-Means dan K-Medoids Untuk Clustering Data Kinerja Karyawan Pada Perusahaan Perumahan Nasional," *Kilat*, vol. 10, no. 1, pp. 157–168, 2021, doi: 10.33322/kilat.v10i1.1174.
- [5] R. Gustrianda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [6] F. Nasari, D. H. Tanjung, and F. Handayani, "Optimasi Metode K-Means dan K-Medoids Berdasarkan Jumlah Cluster dan Nilai DBI Dalam Pengelompokan Produksi Kelapa Sawit Di Provinsi Riau," vol. 7, no. 2, pp. 129–141, 2023, [Online]. Available: <https://www.doi.org/10.22303/infosys.7.2.2023.129-141>
- [7] F. Nasari, "Pengelompokan Daerah Produksi Kelapa Sawit Menggunakan Algoritma K-Means Clustering," vol. 4, no. 2, pp. 34–38, 2021.
- [8] F. Nasari, "Optimalisasi Pengelompokan Data Produksi Kelapa Sawit Menggunakan Algoritma K-Medoids," vol. 5, no. 1, 2022.
- [9] E. Herman, K. E. Zsido, and V. Fenyves, "Cluster Analysis with K-Mean versus K-Medoid in Financial Performance Evaluation," *Appl. Sci.*, vol. 12, no. 16, 2022, doi: 10.3390/app12167985.
- [10] A. S. Sunge, Y. Heryadi, Y. Religia, and Lukas, "Comparison of Distance Function to Performance of K-Medoids Algorithm for Clustering," *Proceeding - ICoSTA 2020 2020 Int. Conf. Smart Technol. Appl. Empower. Ind. IoT by Implement. Green Technol. Sustain. Dev.*, 2020, doi: 10.1109/ICoSTA48221.2020.1570615793.
- [11] A. S. Sunge, N. T. Kurniadi, and E. Widodo, "1 st Pelita International

- Conference Comparison of K-Means and K-Medoid Algorithms in Classifying Village Status (Case Study : Gorontalo Province) 1 st Pelita International Conference 1 st Pelita International Conference 1 st Pelita International Conference,” vol. 01, no. 01, pp. 1–8, 2023.
- [12] A. Jauhari, M. R. Rahabillah, D. R. Anamisa, A. F. Haq, F. A. Mufarroha, and A. A. Purnama, “Comparison of K-means and K-medoids in Tourist Attraction Clustering based on Visitor Characteristics,” in *2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, Nov. 2022, pp. 161–166. doi: 10.1109/ICSINTESA56431.2022.10041666.
- [13] E. Muningsih, I. Maryani, and V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, [Online]. Available: www.bps.go.id
- [14] D. Astuti, “Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering,” *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 60–72, 2019, doi: 10.20895/inista.v1i2.71.
- [15] B. Palupi, I. Rahmawati, M. F. Rizkiana, and F. Retnaningtyas, “TTG Pembuatan Virgin Coconut Oil (VCO) untuk Peningkatan Ekonomi Masyarakat Desa Nogosari Kabupaten Jember,” vol. 7, no. 6, pp. 2–8, 2023.
- [16] K. Kanchymalay, N. Salim, A. Sukprasert, R. Krishnan, and U. R. A. Hashim, “Multivariate Time Series Forecasting of Crude Palm Oil Price Using Machine Learning Techniques,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 226, no. 1, 2017, doi: 10.1088/1757-899X/226/1/012117.
- [17] A. Radovanovic, X. Ye, J. V. Milanovic, N. Milosavljevic, and R. Storchi, “Application of the k-medoids partitioning algorithm for clustering of time series data,” *IEEE PES Innov. Smart Grid Technol. Conf. Eur.*, vol. 2020-October, pp. 645–649, 2020, doi: 10.1109/ISGT-Europe47291.2020.9248796.
- [18] L. Agustini and I. O. Kirana, “Pengelompokan Data Janjang Panen Kelapa Sawit Menggunakan Algoritma K-Medoids Pada PT SIR MANDAU,” *J. Mach. Learn. Data Anal.*, vol. 01, no. 01, pp. 36–44, 2022.
- [19] P. Kaur, S. S. Khurm, and G. S. Josan, “Analysis for classification of similar documents among various websites using rapid miner,” *Proc. 2014 Int. Conf. Issues Challenges Intell. Comput. Tech. ICICT 2014*, pp. 465–470, 2014, doi: 10.1109/ICICT.2014.6781327.
- [20] A. K. Singh, S. Mittal, P. Malhotra, and Y. V. Srivastava, “Clustering Evaluation by Davies-Bouldin Index(DBI) in Cereal data using K-Means,” *Proc. 4th Int. Conf. Comput. Methodol. Commun. ICCMC 2020*, no. Iccmc, pp. 306–310, 2020, doi: 10.1109/ICCMC48092.2020.ICCMC-00057.
- [21] N. Darmeli, “Optimasi Jumlah Cluster Metode K-Medoids Berdasarkan Nilai Dbi Pada Pengelompokan Data Luas Tanaman Dan,” vol. 09, no. 02, 2022.
- [22] R. Adhitama et al, “Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clusterin,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 3, no. 1, pp. 1–5, 2020, doi: 10.33387/jiko.v3i1.1635.
- [23] A. A. Az-zahra, A. F. Marsaoly, I. P. Lestyani, R. Salsabila, and W. O. Z. Madjida, “PENERAPAN ALGORITMA K-MODES CLUSTERING DENGAN VALIDASI DAVIES BOULDIN INDEX PADA PENGELOMPOKKAN TINGKAT MINAT BELANJA ONLINE DI PROVINSI DAERAH ISTIMEWA YOGYAKARTA,” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 9, no. 1, p. 24, 2021, doi: 10.24252/msa.v9i1.18555.