

Implementasi Algoritma K-Means dengan Normalisasi Sigmoidal Untuk Klastering Data Ternak Sapi

Vingki Rapika Sari*¹, Efori Buulolo², Fadlina³

¹Teknik Informatika, Universitas Budi Darma, vingkirapikasari23@gmail.com

²Teknik Informatika, Universitas Budi Darma, eforibuulolo@gmail.com

³Teknik Informatika, Universitas Budi Darma, fadlina123@gmail.com

Keywords:

Algorithm;
K-Means;
Sigmoidal Normalization;
Data Clustering;
Cattle

ABSTRACT

Today's computer technology has been widely used in the field of statistics. This is due to the computing speed of computers which can reduce the time needed to produce the desired final value. One use of computer technology is in grouping cattle data to determine good quality cattle from several shipments that come in one year. The K-Means algorithm is a method used to group cattle data based on the closest and shortest distance or based on similar characteristics. Grouping cattle data determines the criteria for buying weight and selling weight using a calculation process using sigmoidal normalization. The desired result of this research is that the K-Means algorithm in grouping cattle data helps workers to determine good quality cattle. This grouping provides suggestions to help workers determine the best quality cattle.

Kata Kunci

Algoritma;
K-Means;
Normalisasi Sigmoidal;
Klastering Data;
Ternak Sapi

ABSTRAK

Teknologi komputer pada zaman sekarang telah banyak dimanfaatkan dalam bidang statistika. Hal ini disebabkan oleh kecepatan komputasi dari komputer yang dapat memangkas waktu yang dibutuhkan untuk menghasilkan nilai akhir yang diinginkan. Salah satu pemanfaatan teknologi komputer adalah dalam pengelompokan data ternak sapi untuk menentukan sapi berkualitas baik dari beberapa shipment yang masuk dalam satu tahun. Algoritma K-Means merupakan suatu metode yang digunakan untuk mengelompokkan data ternak sapi berdasarkan jarak terdekat dan terpendek atau berdasarkan kesamaan karakteristik. Pengelompokan data ternak sapi menentukan kriteria bobot beli, bobot jual dengan proses perhitungan menggunakan normalisasi sigmoidal. Hasil yang diinginkan dari penelitian ini ialah Algoritma K-Means pada pengelompokan data ternak sapi membantu para pekerja untuk menentukan sapi berkualitas baik, dengan adanya pengelompokan ini memberikan saran untuk membantu para pekerja dalam mengetahui kualitas sapi terbaik.

Korespondensi Penulis:

Vingki Rapika Sari,
Universita Budi Darma,
Jl. Sisingamangaraja No. 338 Kota Medan

Diajukan: 03-12-2023 | Diterima: 11-12-2023 | Diterbitkan: 30-12-2023

1. PENDAHULUAN

PT. Indofarm Sukses Makmur merupakan perusahaan yang melakukan kegiatan usaha Budidaya penggemukan sapi potong yang dapat memenuhi kebutuhan protein hewani manusia. Perusahaan ini didirikan pada tahun 2008. Kandang yang dibangun merupakan kandang tradisional. Pada tahun 2008 s/d 2023 perusahaan ini mulai mengalami perubahan terutama pada kandang, dari kandang tradisional menjadi kandang modern dengan kapasitas ± 3.000 ekor dan penyediaannya mengutamakan sapi *import*.

Perusahaan melakukan import sapi bakalan jenis Brahman Cross dari Australia sebab jenis sapi ini mempunyai kualitas baik dalam penggemukan sapi di PT. Indofarm Sukses Makmur. Hingga sekarang pengembangan usaha masih terus dilakukan, guna untuk meningkatkan populasi sapi diperusahaan.

Data ternak sapi tersebut disimpan di media penyimpanan dan belum dimanfaatkan untuk klastering sapi yang berkualitas dan sapi yang kurang berkualitas. PT. Indofarm Sukses Makmur mengalami masalah yang sering terjadi adalah klastering atau pengelompokan data ternak sapi untuk mengetahui keberhasilan produksi ternak sapi untuk menentukan sapi berkualitas baik dari beberapa shipment atau jenis Brahman Cross yang masuk dalam satu tahun pada PT. Indofarm Sukses Makmur. Hal ini menyebabkan perusahaan mengalami kerugian karena salah klastering atau pengelompokan data ternak dikarenakan kualitas sapi berkaitan dengan harga jual.

Oleh karena itu, dibutuhkan teknik atau algoritma dalam klastering atau pengelompokan data ternak sapi dengan menggunakan kriteria tertentu. Salah satu algoritma klastering dalam data mining, algoritma K-Means Merupakan algoritma yang mengelompokkan data berdasarkan jarak terdekat dan terpendek atau berdasarkan kesamaan karakteristik [1]. Pengelompokan data ternak sapi menggunakan kriteria Bobot Awal, Bobot Beli, Bobot Jual. Dalam proses perhitungan menggunakan normalisasi sigmoidal tujuannya adalah untuk menyederhanakan dalam proses perhitungan.

Berdasarkan penelitian yang dilakukan oleh Ulil Ma'rifat, dengan berjudul "Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng" menyimpulkan bahwa Pengelompokan data penyakit pasien kedalam kelompok penyakit "Akut" atau penyakit "Tidak Akut". Menggunakan metode k-means yang kemudian menghasilkan output kelompok penyakit berdasarkan hitungan dari metode k-means [2].

Penelitian juga pernah dilakukan oleh Fajar Nur Rahmat Fauzan, Budi Darma Setiawan, dan Issa Arwani dengan judul " Implementasi Algoritma K-Means untuk Klasterisasi Kinerja Akademik Mahasiswa", peneliti mencoba untuk mengekstrak pengetahuan yang bisa menggambarkan kinerja prestasi akademik mahasiswa pada akhir semester dan hasil dari penelitian tersebut menunjukkan bahwa dari semua jumlah cluster yang dimasukkan untuk cluster yang berjumlah 3 memiliki nilai silhouette coefficient yang paling mendekati nilai $S_i = 1$, yaitu dengan nilai 0,108690751 [3].

Penelitian juga pernah dilakukan oleh Tria Pratiwi Sutriyani, Amril Mutoi Sigegar dan Dwi Sulistya Kusumaningrum pada tahun 2018 tentang " Implementasi Algoritma K-Means Terhadap Pengelompokan Nilai Ujian Nasional Tingkat SMP di Provinsi Jawa Barat ", Menyimpulkan bahwa akan dilakukan penelitian untuk menambang data menjadi informasi, pada penelitian ini data yang digunakan adalah nilai rata-rata UN tingkat SMP di Provinsi Jawa Barat menggunakan teknik data mining [4].

Penelitian juga pernah dilakukan oleh Efori Buulolo dan Rian Syahputra, dengan judul "Implementasi Algoritma Clustering K-Means Untuk Mengelompokkan Mahasiswa Baru Yang Berpotensi (Studi Kasus : STMIK Budi Darma)" menyimpulkan bahwa agar pengelompokan calon mahasiswa baru berdasarkan potensi yang dimiliki maka digunakan algoritma clustering K-Means. Dasar pengelompokan dengan algoritma clustering K-Means berdasarkan nilai jarak terdekat dari kriteria yang berbasis numerik dengan output dalam bentuk cluster. Cluster yang telah dibentuk digunakan sebagai dasar pengelompokan calon mahasiswa baru ke dalam satu kelas [5].

Penelitian juga pernah dilakukan oleh Abu Salam, Diyan Adiatma dan Junita Zeniarja, dengan judul " Implementasi Algoritma K-Means dalam Pengklasteran Untuk Rekomendasi Penerima Beasiswa PPA di UDINUS " menyimpulkan bahwa Dari proses penelitian yang dilakukan didapatkan hasil dari 441 dataset pendaftar beasiswa PPA tahun 2016 seperti pada gambar 5 dan gambar 6, terdapat 154 data mengikuti cluster 0 yang artinya direkomendasikan diterima atau mendapatkan beasiswa sedangkan 287 data lainnya mengikuti cluster 1 yang artinya tidak direkomendasikan atau tidak mendapatkan beasiswa [6].

Penelitian terdahulu menyimpulkan bahwa proses pengelompokan algoritman K-means memiliki kekurangan dan kelebihan dalam sebuah pengelompokan data. Adapun beberapa kelebihan dan kekurangan sebagai berikut. Kelebihan Algoritma K-means memberikan informasi yang akurat dan tepat dalam mengenai pengelompokan data ternak sapi dan kekurangan Algoritma K-Means tampilan system yang di bangun sangat sederhana.

2. METODE PENELITIAN

Metode penelitian yang telah dapat dijabarkan dalam pembahasan masing-masing prosedur dalam penelitian sebagai berikut:

1. Identifikasi Masalah

Yakni menguraikan permasalahan secara konteks pada objek penelitian menggunakan algoritma K-Means, sehingga dapat menentukan batasan dari masalah yang akan diteliti.

2. Penelitian Lapangan

Teknik yang dilakukan untuk memperoleh data secara langsung agar data yang diperoleh lebih akurat. Dalam hal ini pengambilan data dan sampel dilakukan dengan cara :

a. Wawancara

Peneliti melakukan Tanya jawab kepada manager indofarm mengenai kriteria yang digunakan dalam pengelompokan data ternak sapi,

b. Observasi

Penelitian melakukan observasi langsung di lokasi penelitian untuk mendapatkan informasi mengenai data ternak sapi.

c. Sampling

Penelitian mengumpulkan sampel data dan menganalisisnya sesuai dengan prosedur dalam pengelompokan data ternak sapi

3. Pengumpulan Literatur

Dalam pengumpulan Literatur, penulis memperoleh dengan cara membaca, memahami dan mengutip hal-hal yang bersangkutan dengan penelitian yang diangkat seperti pada buku, jurnal online, artikel dan media cetak online lainnya yang membahas tentang normalisasi sigmoidal dengan algoritma K-means untuk pengelompokan data ternak sapi.

4. Analisa dan Pembahasan

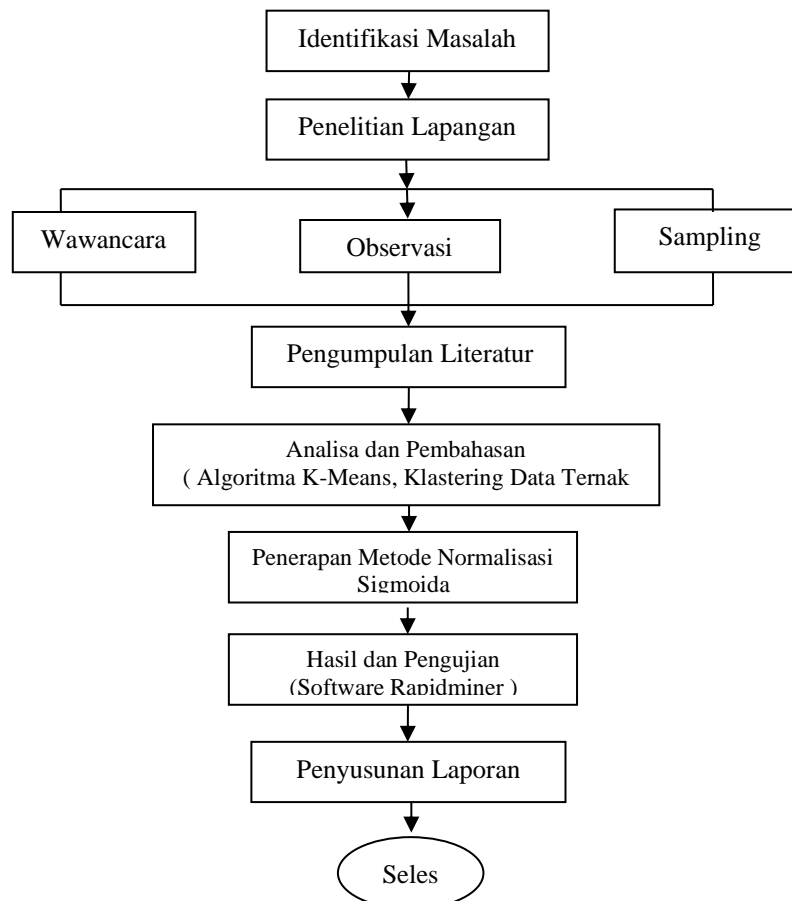
Menganalisa metode yang digunakan berdasarkan kasus penelitian yang akan diteliti pada data ternak sapi. Data yang diperoleh dan dikumpulkan dikelola dengan menggunakan metode yang dikombinasi dan menganalisis cara kerja metode tersebut.

5. Hasil dan Pengujian

Pada tahapan ini peneliti akan menjabarkan keluaran akhir yang dihasilkan system mulai dari perhitungan menggunakan algoritma K-Means, lalu menormalisasikan hasilnya dengan sigmoidal dalam pengelompokan data ternak sapi.

6. Penyusunan Laporan

Peneliti melakukan penyusunan laporan sesuai dengan kerangka kerja penelitian yang telah dilakukan pada PT. Indofarm Sukses Makmur.



Gambar 1. Kerangka Kerja Penelitian

2.1 Data Mining

Data mining merupakan kegiatan pengumpulan data yang bertujuan untuk menemukan pola tersembunyi pada data yang telah diproses. Dari informasi yang diolah dengan teknik data mining, terciptalah informasi baru yang berasal dari informasi lama yang hasilnya dapat digunakan untuk pengambilan keputusan di masa depan [7].

2.2 Algoritma K-Means

K-means adalah algoritma pengelompokan. Pengelompokan K-means adalah salah satu algoritma pembelajaran mesin tanpa pengawasan yang paling sederhana dan paling populer. K-means clustering adalah metode analisis data atau metode data mining yang melakukan pemodelan tanpa pengawasan dan merupakan metode pengelompokan data menggunakan sistem partisi.

K-means clustering adalah metode analisis non-hierarchical clustering yang bertujuan untuk membagi objek yang ada menjadi satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek dengan karakteristik yang sama dikelompokkan ke dalam cluster yang sama dan objek dengan karakteristik yang berbeda dikelompokkan ke dalam cluster yang sama. dikelompokkan ke dalam cluster lain [8].

Dalam algoritme pembelajaran ini, komputer mengelompokkan datanya ke dalam inputnya tanpa pengetahuan sebelumnya tentang kategori target. Pembelajaran ini termasuk dalam pembelajaran tanpa pengawasan. Input yang diterima adalah data atau objek dan k-cluster yang diinginkan. Algoritma ini mengelompokkan data atau objek ke dalam k grup. Setiap cluster memiliki pusat (centroud) yang mewakili cluster tersebut.

Dalam teknik clustering terdapat dua metode *clustering* yaitu clustering hirarkis dan clustering non-hierarkis. Pengelompokan hierarki adalah metode pengelompokan data yang bekerja dengan cara mengelompokkan dua atau

lebih objek yang sejenis atau serupa, kemudian proses dilanjutkan ke objek lain yang dekat dengan keduanya. Proses ini berlanjut hingga cluster membentuk semacam pohon. hierarki atau tingkat objek yang jelas dari yang paling mirip hingga yang paling tidak mirip. Namun secara logika, semua objek pada akhirnya hanya membentuk sebuah cluster [9].

Sedangkan *non-hierarchical* clustering pada teknik ini dimulai dengan menentukan jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, empat *cluster* atau lebih), kemudian setelah jumlah *cluster* yang diinginkan, proses clustering dimulai tanpa proses hirarki, metode ini juga merupakan sering disebut dengan metode K-Means Clustering [10].

Pengelompokan data menggunakan metode pengelompokan K-Means biasanya dilakukan dengan menggunakan algoritma dasar berikut :

1. Tentukan jumlah cluster
2. Secara acak menetapkan data ke cluster
3. Hitung centroid/rata-rata data untuk setiap cluster
4. Ratakan semua data ke centroid/mean terdekat
5. Kembali ke langkah 3 jika masih ada data dalam cluster, atau jika nilai centroid berubah, ada yang melebihi ambang batas yang ditentukan, atau jika fungsi tujuan yang digunakan berubah melebihi ambang batas yang ditentukan.

Beberapa masalah yang sering terjadi saat menggunakan metode tersebut K-Means untuk melakukan pengelompokan data adalah:

1. Temukan beberapa model pengelompokan yang berbeda
2. Pemilihan jumlah cluster tercepat
3. Kegagalan
4. *Outlier*
5. Bentuk cluster

Algoritma pengelompokan K-Means adalah sebagai berikut:

1. Pilih pusat K secara acak
2. Mengelompokkan data sehingga K cluster terbentuk sedemikian rupa sehingga titik centroid dari setiap cluster adalah titik centroid yang telah dipilih sebelumnya.
3. Perbarui nilai pusat gravitasi
4. Ulangi langkah 2 dan 3 hingga nilai titik pusat tidak lagi berubah

Data dapat dikelompokkan menjadi sebuah cluster dengan menghitung jarak terpendek dari data ke centroid.

Algoritma k-means memiliki beberapa keunggulan, yaitu:

1. Mudah diimplementasikan dan digunakan.
2. Dibutuhkan waktu yang relatif cepat untuk melakukan penelitian ini.
3. Mudah disesuaikan.
4. Sering digunakan.

Algoritma K Means memiliki beberapa kelebihan, tetapi juga memiliki beberapa kekurangan. Kelemahan dari algoritma ini adalah sebagai berikut:

1. Sebelum menjalankan algoritme, k titik diinisialisasi secara acak, sehingga pengelompokan data yang dihasilkan mungkin berbeda. Jika nilai acak inisialisasi kurang baik, pengelompokan yang dihasilkan akan kurang optimal.
2. Dapat mungkin mengalami masalah yang disebut curse of dimensionality. Hal ini bisa terjadi jika training memiliki dimensi yang sangat tinggi (misalnya jika data training terdiri dari 2 atribut maka dimensinya akan menjadi 2 dimensi. Tetapi jika ada 2 atribut maka dimensinya adalah 20). Salah satu cara algoritma ini bekerja adalah mencari jarak terpendek antara k titik dengan titik lainnya. Namun, jika Anda mencari jarak antar titik dalam dua dimensi, itu mudah dilakukan. Tapi bagaimana mencari jarak antar titik bila ada dimensi 20. Akan sulit.
3. Jika hanya ada beberapa titik sampel dalam data, cukup mudah untuk membaca dan menemukan titik terdekat dengan titik-k yang diinisialisasi secara acak. Namun, jika terdapat banyak titik data (misalnya, satu miliar titik data), akan memakan banyak waktu untuk menghitung dan menemukan titik terdekat. Proses ini dapat dipercepat, tetapi membutuhkan struktur data yang lebih kompleks seperti kD-Tree atau hashing.

2.3 Normalisasi Sigmoid

Sigmoidal merupakan metode normalization melakukan normalisasi data secara nonlinier ke dalam range 1-1 dengan menggunakan fungsi sigmoid.

Rumus (1)

$$\text{Newdata} = (1 - e^{-x}) / (1 + e^{-x})$$

$$X = (\text{data} - \text{mean}) / \text{std}$$

$$e = \text{nilai eksponensial (2,718281828)}$$

Metode ini sangat berguna pada saat data-data yang ada melibatkan data-data outlier. Data outlier data yang keluar jauh dari jangkauan data lainnya [11].

2.4 Hewan Sapi

Sapi adalah hewan ternak yang merupakan salah satu sumber protein hewani favorit masyarakat. Budidaya ternak sapi banyak dilakukan oleh masyarakat di karenakan ketersediaan pakannya yang melimpah dan pemeliharaan tidak sulit yaitu cukup menggunakan kandang terbuat dari kayu dan sapi dapat ditenakan.

Oleh karena itu, diperlukan sebuah pengelompokan data ternak sapi dengan metode algoritma K-means dengan menggunakan kriteria bobot beli, bobot awal, bobot jual untuk mengetahui keuntungan yang diperoleh dari usaha penggemukan sapi potong. Dalam proses perhitungan keuntungan menggunakan normalisasi sigmoidal tujuannya adalah untuk menyederhanakan dalam proses perhitungan [12].

2.5 Aplikasi Rapid Miner

RapidMiner dikenal dengan sebutan YALE (*Yet Another Learning Environment*), dikemukakan oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer merupakan Unit Kecerdasan Buatan Universitas Teknik Dortmund tahun 2001. Pada tahun 2006, perkembangannya didorong oleh Rapid-I, sebuah perusahaan yang didirikan oleh Ingo Mierswa dan Ralf Klinkenberg pada tahun yang sama. Pada tahun 2007, nama perangkat lunak itu berubah dari YALE menjadi RapidMiner. Pada tahun 2013, perusahaan melakukan rebranding dari Rapid-I menjadi RapidMiner. RapidMiner sebuah perangkat lunak ilmu yang dikembangkan oleh perusahaan yang bersama-sama menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam penambangan teks dan analisis prediktif. RapidMiner digunakan dalam bidang bisnis, selain itu juga dalam bidang pendidikan, penelitian, pelatihan, prototype yang mendukung proses pembelajaran mesin seperti visualisasi, model, persiapan data dan optimasi [13].

3. HASIL DAN ANALISIS

3.1 Analisis

Tahap analisa merupakan tahap yang penting dalam melanjutkan setiap tahapan selanjutnya pada penelitian. Dalam tahap ini akan dijelaskan langkah-langkah menormalisasikan data untuk pengelompokan data ternak sapi. **Tabel 1** dapat dilihat [Klik](#)

Selanjutnya dilakukan konversi huruf, pada Algoritma K-Means data yang terbentuk karakter tidak dapat diproses oleh karena itu perlu dilakukan perubahan data yang dalam bentuk karakter ke dalam bentuk angka.

| Sex | Angka |
|--------|-------|
| Heifer | 1 |
| Steer | 2 |
| Bull | 3 |

Berikut dapat dilihat **Tabel 3** [Klik](#)

Selanjutnya dilakukan normalisasi data dengan model sigmoida tujuannya adalah untuk menghindari ketimpangan antar data yang digunakan untuk klastering data, lihat data **Tabel 4** [Klik](#)

Berdasarkan tabel 4 berikut perhitungan nilai rata-rata masing-masing kolom

- Nilai rata-rata (Mean) untuk sex = $46 / 27 = 1.704$
- Nilai rata-rata (Mean) untuk BB Beli = $9003 / 27 = 333.444$
- Nilai rata-rata (Mean) untuk BB Jual = $13717 / 27 = 508.037$
- Nilai rata-rata (Mean) untuk Selisih BB = $4714 / 27 = 174.593$
- Nilai rata-rata (Mean) untuk DOF = $3719 / 27 = 137.741$
- Nilai rata-rata (Mean) untuk ADG = $34,54 / 27 = 1.279$

Selanjutnya mencari standart deviasi dengan rumus

Standart deviasi untuk sex

$$S = \sqrt{\frac{((1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (3 - 1,704)^2 + (3 - 1,704)^2 + (3 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (1 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2 + (2 - 1,704)^2)}{27}} = 0,656$$

Standart deviasi untuk kriteria lainnya dapat dilihat [Klik](#)

Tabel hasil dapat dilihat pada **Tabel 5** dilihat [Klik](#)

Selanjutnya adalah mengelompokkan data dengan algoritma K-Means. Untuk data ditentukan sebanyak 3 *cluster*, sehingga dapat ditentukan untuk nilai *k* adalah $k = 3$

Iterasi I :

Mentukan nilai centroid pada iterasi I ditentukan secara acak dari 27 simple data.

Tabel 6. Nilai Centroid Iterasi I

| | | | | | | |
|----|--------|--------|--------|--------|--------|--------|
| C1 | -0.490 | -0.436 | -0.729 | -0.765 | -0.301 | -0.729 |
| C2 | 0.222 | 0.077 | 0.250 | 0.313 | 0.096 | 0.236 |
| C3 | 0.756 | 0.553 | 0.332 | -0.028 | 0.006 | -0.048 |

a. Pencarian jarak untuk C1

$$dc1 = \sqrt{(-0,490 - (-0,490))^2 + (-0,436 - (-0,436))^2 + (-0,729 - (-0,729))^2 + (-0,765 - (-0,765))^2 + (-0,301 - (-0,301))^2 + (-0,729 - (-0,729))^2}$$

$$= 0.001$$

$$dc2 = \sqrt{(-0,490 - (-0,490))^2 + (0,159 - (-0,436))^2 + (-0,410 - (-0,729))^2 + (-0,673 - (-0,765))^2 + (-0,301 - (-0,301))^2 + (-0,613 - (-0,729))^2}$$

$$= 0.0691$$

$$dc3 = \sqrt{(-0,490 - (-0,490))^2 + (0,239 - (-0,436))^2 + (0,062 - (-0,729))^2 + (-0,123 - (-0,765))^2 + (-0,280 - (-0,301))^2 + (-0,036 - (-0,729))^2}$$

$$= 1.442$$

$$dc4 = \sqrt{(0,222 - (-0,490))^2 + (0,171 - (-0,436))^2 + (0,282 - (-0,729))^2 + (0,284 - (-0,765))^2 + (-0,259 - (-0,301))^2 + (0,441 - (-0,729))^2}$$

$$= 2.090$$

$$dc5 = \sqrt{(0,222 - (-0,490))^2 + (-0,335 - (-0,436))^2 + (-0,492 - (-0,729))^2 + (0,476 - (-0,765))^2 + (-0,259 - (-0,301))^2 + (-0,387 - (-0,729))^2}$$

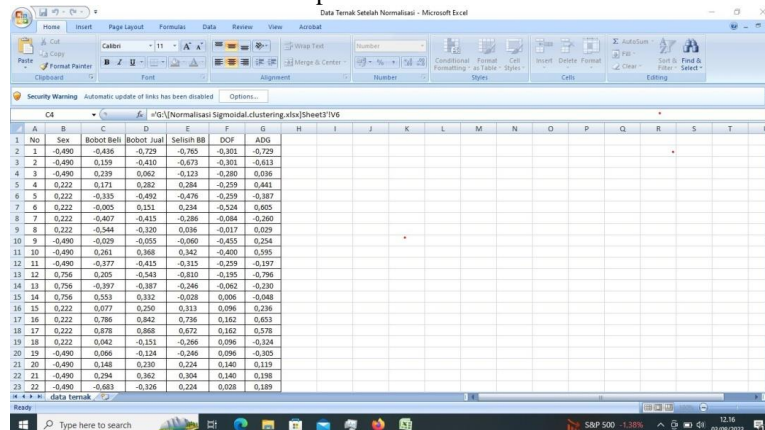
$$= 0.881$$

Proses selanjutnya dapat dilihat pada link berikut ini [Klik](#)
 Sehingga, Kelompok I / Cluster I data ternak sapi dengan data ternak rendah, Kelompok II / Cluster II data ternak sapi dengan data ternak menengah dan Kelompok III / Cluster III kelompok tertinggi.

3.2 Pengujian

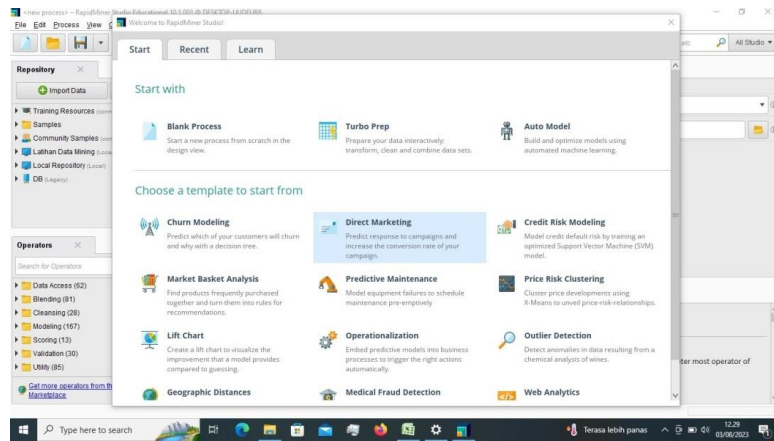
Pengujian data untuk pengelompokan data ternak sapi berkualitas baik pada penelitian menggunakan aplikasi Rapidminer. Adapun langkah pengujian yang dilakukan adalah sebagai berikut:

1. Siapkan data dalam bentuk. Tabel data ternak seperti berikut



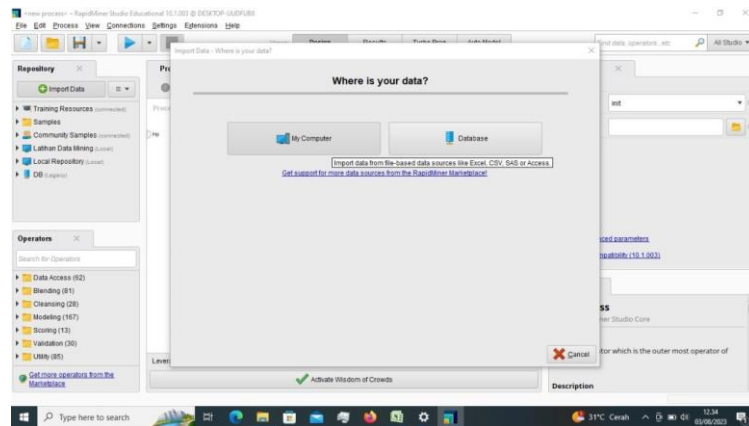
Gambar 2. Data dalam bentuk. Tabel Data Ternak

2. Buka aplikasi Rapidminer dan Klik Blank Proses



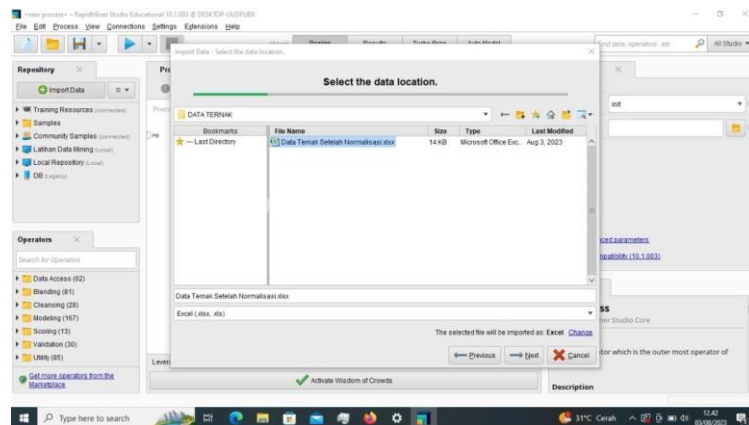
Gambar 3. Halaman Blank Proses

3. Klik File, pilih Import data, muncul di penyimpanan data, pilih dimana data kamu di simpan. My Computer atau Database.



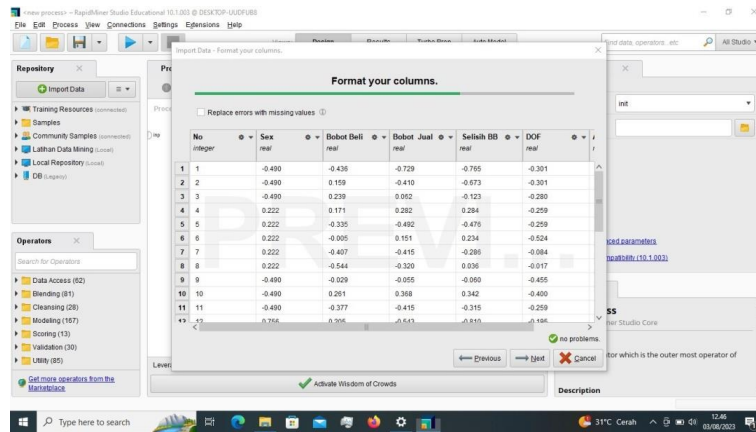
Gambar 4. Import Data

4. Selanjutnya cari file tersimpan di my computer



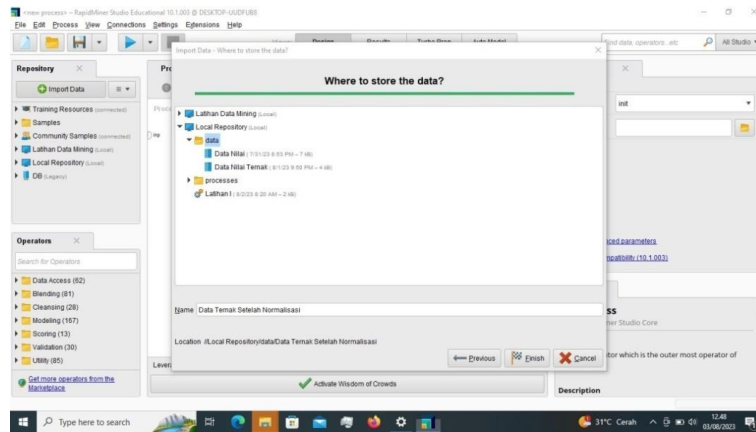
Gambar 5. Halaman pencarian file. Table Data Ternak

6. Selanjutnya pilih file tersimpan di my computer
7. Setelah ditemukan file yang tersimpan di my computer, klik next untuk proses selanjutnya.
8. Setelah proses pencarian file data, lalu muncul proses yang namanya format your columns atau disebut dengan tentukan jenis tipe data berdasarkan attribute.



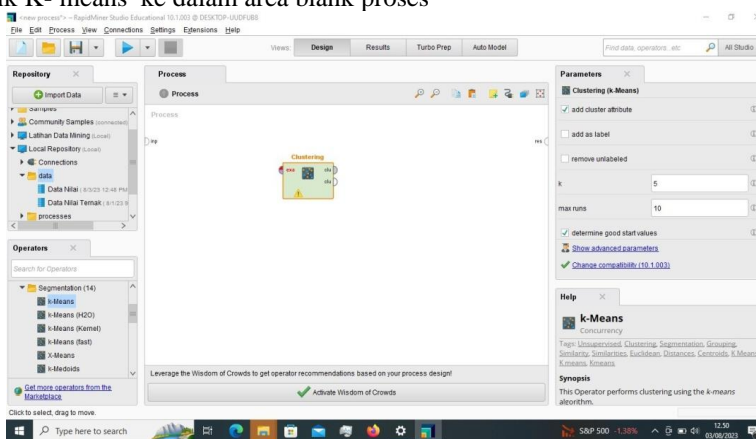
Gambar 6. Penentuan tipe data

9. Klik Next untuk proses selanjutnya.
10. Selanjut proses penyimpanan data sudah di normalisasikan.
11. Pilih Tempat Penyimpanan di local Repository dan pilih folder data, simpan Dengan nama folder Data Nilai Ternak.



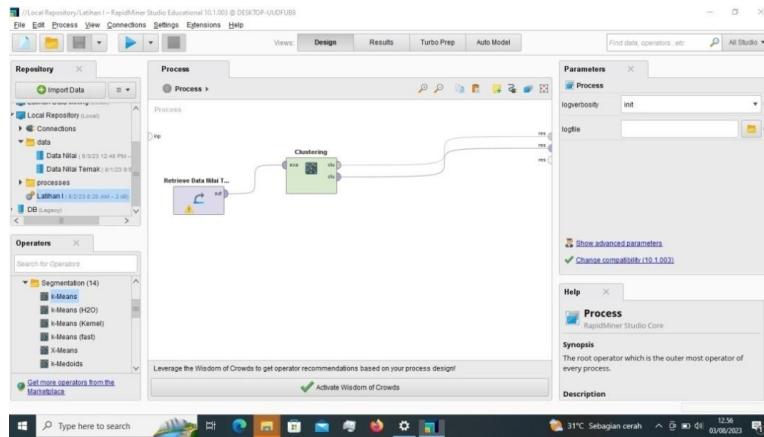
Gambar 7. Finish proses import data

12. Selanjutnya kita proses data dengan Algoritma K-Means
13. Klik File, Pilih new proses
14. Pilih modeling, clustering dan segmentation dan pilih K-means
15. Klik tahan dan tarik K-means ke dalam area blank proses



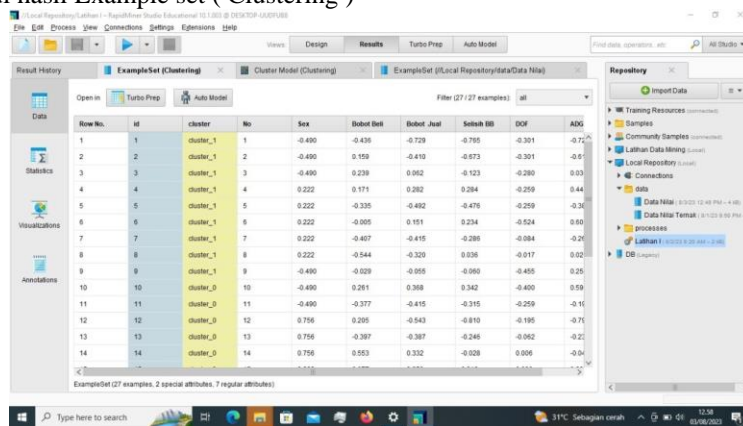
Gambar 8. Pengambilan komponen K-Means

16. Selanjutnya tahan dan tarik data nilai ternak yang digunakan ke dalam blank proses
17. Hubungkan data nilai ternak ke K-Means
18. Hubungkan clu ke res



Gambar 9. Pengambilan komponen data nilai ternak

19. Jalankan dengan mengklik play
20. Selanjutnya muncul hasil Example set (Clustering)



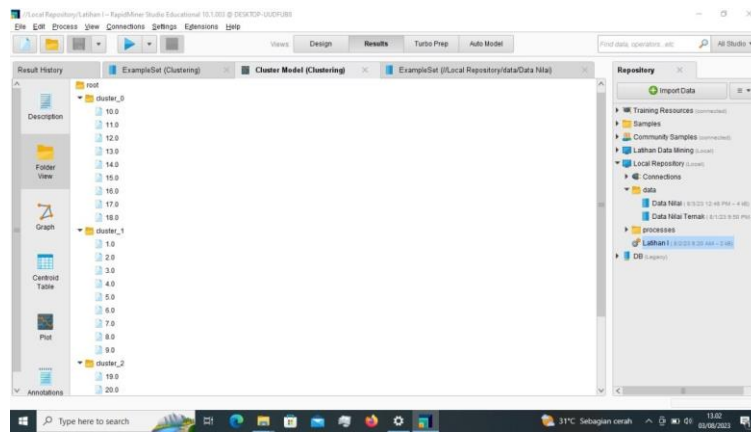
Gambar 10. Hasil Example set (Clustering)

21. Selanjutnya muncul hasil Cluster model



Gambar 11. Hasil Cluster model

22. Hasil pengelompokkan data sebagai berikut



Gambar 12. Hasil pengelompokan data

Berdasarkan hasil pengujian data dengan aplikasi rapidminer, dapat disimpulkan hasil pengelompokan data dengan algoritma K-Means dengan hasil pengujian aplikasi memiliki hasil yang hamper sama dengan tingkat kesamaan yaitu 90%.

4. KESIMPULAN

Kesimpulan dari penelitian ini yaitu Pertama, algoritma K-Means dapat digunakan untuk pengelompokan data ternak sapi untuk menentukan sapi berkualitas baik dengan menggunakan kriteria Notel, RFID, Shipment, Sex, Tanggal, Bobot, DOF dan ADG. Kedua, penggunaan normalisasi data pada pengelompokan data ternak sapi memberikan hasil yang pasti terhadap cluster data ternak karena kriteria data yang digunakan tidak saling mereduksi. Ketiga, hasil pengelompokan data ternak sapi dikelompok menjadi 3 (tiga) kelompok yaitu rendah, menengah dan tinggi.

REFERENSI

- [1] a. K. K. A. Solichin, "Klasterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta," vol. 5, p. 2, 2020.
- [2] U. Ma'rifatin, "Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng," p. 1, 2020.
- [3] F. N. R. F. J. D. S. A. Aziz, "Implementasi Algoritma K-Means Untuk Klasterisasi Kinerja Akademik Mahasiswa," J.Pengemb. Teknol. Inf. dan Ilmu Komput, vol. 2, p. 6, 2018.
- [4] A. D. Tria, "Implementasi Algoritma K-Means Terhadap Pengelompokan Nilai Ujian Nasional Tingkat SMP di Provinsi Jawa Barat," vol. 3, p. 1, 2018.
- [5] E. B. a. R. Syahputra, "Implementasi Algoritma Clustering K-Means Untuk Mengelompokkan Mahasiswa Baru Yang Berpotensi (Studi Kasus : Stmik Budi Darma)," vol. 2, pp. 17-24, 2019.
- [6] D. A. J. A. Salam, "Implementasi Algoritma K-Means dalam Pengklasteran Untuk Rekomendasi Penerima Beasiswa PPA di UDINUS," vol. 5, p. 1, 2020.
- [7] E. Buulolo, "Data Mining Untuk Perguruan Tinggi. Yogyakarta: deepublish," 2020.
- [8] Muktiani, Sukses Usaha Penggemukan Sapi Potong, Yogyakarta: Pustaka Baru Press, 2019.
- [9] N. A. Ali Ikhwan, "Implementasi Data Mining Untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means," vol. 4, p. 2, 2020.
- [10] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means Algorithm: A Comprehensive Survey and Performance Evaluation." *Electronics*, vol. 9, no. 8, p. 1295, 2020, doi: 10.3390/electronics9081295.
- [11] R. A. Indraputra and R. Fitriana, "K-Means clustering data COVID-19," *Jurnal Teknik Industri*, vol. 3, no. 10, pp. 275–282, 2020.
- [12] S. Sarbaini, W. Saputri, Nazaruddin, and F. Muttakin, "Cluster Analysis Menggunakan Algoritma Fuzzy K-Means Untuk Tingkat Pengangguran Di Provinsi Riau," *Jurnal Teknologi dan Manajemen Industri Terapan*, vol. 1, no. 2, pp. 78–84, Jun. 2022, doi: <https://doi.org/10.55826/tmit.v1i1I.30>.
- [13] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," *INFORMAL: Informatics Journal*, vol. 5, no. 1, p. 10, Apr. 2020, doi: <https://doi.org/10.19184/isj.v5i1.17071>.
- [14] G. Fadhlullah, Z. Abdurrahman Baizal, and N. Ikhsan, "JURNAL MEDIA INFORMATIKA BUDIDARMA Conversational Recommender Systems Based on Mobile Chatbot for Culinary," vol. 5, pp. 1233–1241, 2021, doi: <https://doi.org/10.30865/mib.v5i3.3113>.
- [15] E. Patel and D. S. Kushwaha, "Clustering Cloud Workloads: K-Means vs Gaussian Mixture Model," *Procedia Computer Science*, vol. 171, pp. 158–167, 2020, doi: <https://doi.org/10.1016/j.procs.2020.04.017>.