

BAB II TINJAUAN PUSTAKA

2.1. Studi Pustaka

Dalam penyusunan penelitian ini, penulis terinspirasi dan mereferensi dari penelitian-penelitian sebelumnya yang berkaitan dengan latar belakang masalah pada penelitian ini. Adapun penelitian yang berhubungan dengan penelitian ini dapat dilihat dari Tabel 2.1:

Tabel 2.1. Penelitian Terkait

Peneliti	Judul Penelitian	Masalah	Metode
Irfan Helmy (2021)	Strategi pembangunan Kawasan pedesaan peternakan sapi di Kabupaten Kebumen	Bagaimana upaya menggerakkan ekonomi pedesaan yang lebih terarah dan strategik	<i>Focus Group Discussion</i>
Achmad zaki Yamani, Tri Ginanjar laksana, Novanda Alim Setya Nugraha dan Ridho Ananda (2021)	Pengembangan Sistem <i>Big Data</i> peternakan sapi berbasis dual <i>layer APPS</i> dan <i>machine learning</i> di Kabupaten Kebumen	Bagaimana cara mengetahui gambaran otentik tentang ketersediaan sapi yang di potong secara <i>update</i> , bagaimana upaya pengoptimalan pencatatan sapi yang telah dipotong secara <i>real time</i> di Kabupaten Kebumen	Pendekatan model <i>waterfall</i> berbasis android
Moch. Sugiarto, Syarifuddin Nur dan Oetoeng Edy Djamiko (2017)	Optimal curahan waktu kerja dan penambahan skala optimal usaha perbibitan sapi PO kebumen	Mengidentifikasi curahan waktu kerja peternak dalam usaha perbibitan sapi PO kebumen	<i>Survey</i>
Mikaela Aditya Wahyu Krisnamurti (2017)	Penerapan metode <i>K-Means</i> clustering untuk mengelompokkan potensi produksi buah buahan di Provinsi Daerah Istimewa Yogyakarta	Hasil panen buah yang bervariasi di Provinsi Daerah Istimewa Yogyakarta sehingga diperlukan pengelompokkan	<i>K-means</i>

Peneliti	Judul Penelitian	Masalah	Metode
		daerah potensial penghasil buah	
Andrew Donda Muthw (2019)	Penerapan <i>clustering time series</i> untuk menggerombolkan provinsi di Indonesia berdasarkan nilai produksi padi	Memenuhi kebutuhan pangan nasional, tanaman padi merupakan komoditas utama yang diupayakan setiap tahun agar produksinya terus meningkat	<i>Clustering time series</i>
Luh Putu Widya, Pardomuan Robinson Sihombing (2021)	Analisis <i>clustering time series</i> dalam pengelompokan Provinsi di Indonesia berdasarkan nilai PDRB	Perbandingan jarak <i>eulicid</i> dan DTW dalam mengelompokn provinsi-provinsi yang ada di Indonesia	<i>Clustering time series</i>
M. Alfian Alfian Riyadi, Kartika Fithriasari	<i>Data mining</i> peramalan konsumsi listrik dengan pendekatan <i>cluster time series</i> sebagai <i>preprocessing</i>	Permasalahan proses identifikasi model yang sesuai untuk tiap series	Metode ARIMA dan ANN
Ardiansyah Muhammad (2014)	Penggunaan jarak <i>Dynamic Time Wrapping (DTW)</i> pada analisis <i>cluster</i> data deret waktu	Bagaimana Pengelompokan Dana Pihak Ketiga (DPK) dengan menggunakan jarak DTW dan metode pengelompokan secara hirarki	<i>Clustering time series</i> secara Hirarki

Berdasarkan Tabel 2.1. penelitian yang terkait yaitu penelitian Helmy (2021) membahas mengenai tantangan membuat strategi pengembangan Kawasan pedesaan peternakan sapi di Kabupaten Kebumen dalam kegiatan pemanfaatan atau budidaya sumber daya Kabupaten Kebumen. Salah satunya yaitu dengan menyusun rencana strategi pembangunan kawasan pedesaan sapi PO di Kabupaten Kebumen, yang diharapkan dapat menjadi panduan dalam menganalisis isu strategis kawasan pedesaan, strategi dan arah kebijakan Kawasan pedesaan dan dasar mengalokasikan kebutuhan pendanaan dalam rangka

penyelenggara pembangunan Kawasan pedesaan sapi di Kabupaten Kebumen. Penelitian ini menghasilkan program yang menjadi strategi pengembangan Kawasan pedesaan yang mandiri, unggul dan berkelanjutan. Metode yang digunakan dalam penelitian ini yaitu, koordinasi dengan pihak dispermades P3A, *Focus Group Discussion* (FGD) dan penerapan rencana program.

Penelitian Kusuma dkk (2020) membahas mengenai permasalahan untuk mengetahui bobot sapi Peranakan Ongole (PO) di Kabupaten Kebumen. Penelitian ini menggunakan 150 ekor sapi PO yang terdiri 25 ekor sapi umur 7 bulan, 36 ekor sapi umur 5-6 bulan dan 20 ekor sapi umur 4 bulan. Berdasarkan hasil penelitian rata-rata bobot lahir jantan umur sapi 7 bulan yaitu 183,87 Kg, sapi umur 5-6 bulan rata-rata memiliki bobot 135,18 Kg, dan sapi umur 4 bulan memiliki berat rata-rata 178,76 Kg, sapi umur 5-6 bulan memiliki berat rata-rata 147,87 Kg. Metode yang digunakan dalam penelitian ini yaitu metode kuantitatif.

Penelitian Sugiarto dkk (2017) membahas mengenai permasalahan dalam meningkatkan pendapatan melalui penguatan efisiensi usaha khususnya peningkatan skala usaha sapi PO Kabupaten Kebumen. Bertujuan untuk mengidentifikasi curahan waktu kerja perternak dalam pembibitan dan dapat merumuskan skala usaha pembibitan sapi PO Kebumen sebagai usaha sampingan. Penelitian ini menggunakan *survey* melalui wawancara, saran dari peneliti ini yaitu peternak sapi PO Kabupaten Kebumen disarankan menggunakan teknologi pengolahan pakan dan pembukaan lahan pakan apabila menambah usaha sapi potong yang dimilikinya.

Penelitian Alim dkk (2021) membahas permasalahan mengenai pencatatan pemotongan sapi dan *monitoring* sapi di Kabupaten Kebumen, membuat pemerintah daerah sulit dalam menentukan kebijakan terkait kebutuhan konsumsi daging sapi. Berdasarkan kondisi tersebut, maka penelitian ini mengusulkan untuk optimalisasi pencatatan dan monitoring sapi di Kabupaten Kebumen menggunakan pendekatan model *waterfall* berbasis *android*. Hasil penelitian ini berbentuk sistem informasi daging sapi berbasis *dua layer apps* dan *machine learning* yang bernama SIDAK (Sistem Informasi Daging Sapi Kebumen), aplikasi SIDAK dapat menjadi

masukannya terhadap perumusan arah kebijakan pengendalian serta pengelolaan sapi di Kabupaten Kebumen.

Penelitian oleh Murti (2017) membahas mengenai pengelompokan daerah potensial penghasil buah untuk dapat mengetahui daerah yang menghasilkan buah dengan jumlah banyak ataupun sedikit. Metode yang digunakan pada penelitian ini yaitu metode *k-means clustering*, dengan menggunakan pendekatan *k-means clustering* pembagian daerah dapat dilakukan berdasarkan luas panen, produksi dan tahun panen. Tujuan dari penelitian ini yaitu memudahkan pengelompokan suatu daerah dengan hasil produksi buah banyak, sedang dan rendah. Hasilnya adalah gambaran yang menunjukkan pengelompokan daerah berdasarkan hasil pertanian buah

Penelitian Putu Widya Adnyani & Robinson Sihombing (2021) membahas mengenai upaya pengembangan pembangunan di Indonesia salah satunya sektor ekonomi dimana salah satu indikatornya yaitu dengan melihat Produk Domestik regional Bruto (PDRB). Untuk itu peneliti dilakukan pengelompokan berdasarkan nilai PDRB menggunakan metode *clustering time series*. Analisis *clustering time series* dilakukan menggunakan jarak *Dynamic Time Wrapping* (DTW) dan jarak *Euclid* dengan membandingkan hasil penggerombolan masing-masing jarak tersebut mengenai nilai Pendapatan Domestik Regional Bruto (PDRB).

Pada penelitian Muhammada (2014) membahas mengenai pengelompokan Dana Pihak Ketiga (DPK). DPK merupakan sumber dana terbesar yang dimiliki oleh bank, oleh karena itu bank umum memiliki peran yang besar dalam menggerakkan perekonomian di Indonesia. Metode yang digunakan yaitu pengelompokan menggunakan jarak DTW dan validasi menggunakan koefisien *silhouette*. Berdasarkan hasil penelitian diperoleh dua kelompok dengan nilai koefisien *silhouette* 0.9775

Berdasarkan penelitian-penelitian sebelumnya belum terdapat penelitian yang membahas mengenai pengelompokan kecamatan penghasil sapi lokal di Kabupaten Kebumen. Penelitian mengenai Pengelompokan kecamatan penghasil sapi lokal di Kabupaten Kebumen dilakukan untuk memudahkan pemerintah

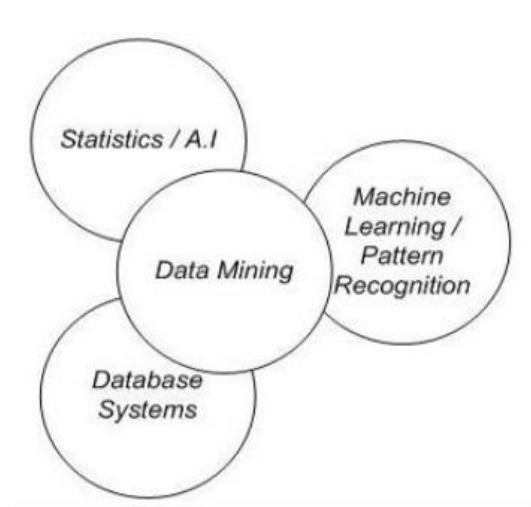
daerah dalam mengambil kebijakan upaya pengembangan daerah kawan penghasil sapi lokal di Kabupaten Kebumen. Metode yang digunakan dalam penelitian pengelompokan kecamatan penghasil daging sapi lokal di Kabupaten Kebumen menggunakan data *mining* pengelompokan, dengan metode *clustering time series* menggunakan perhitungan jarak *Dynamic Time Wrapping* (DTW) dan perhitungan *multidimensional scalling*.

Penelitian ini menggunakan *clustering time series* karena *clustering time series* merupakan metode yang tepat untuk mengelompokan data produksi daging sapi setiap kecamatan yang berbentuk data runtun waktu dan menggunakan perhitungan jarak *Dynamic Time Wrapping* (DTW) karena berdasarkan penelitian Widya & Sihombing (2021) yang membanding penggunaan jarak *Dynamic Time Wrapping* (DTW) dan jarak *Euclide* menghasilkan nilai *coefficient silhoutte* jarak *Dynamic Time Wrapping* (DTW) lebih besar dibandingkan nilai *coefficient silhoutte* jarak *euclide*.

2.2.Dasar Teori

2.2.1. Data Mining

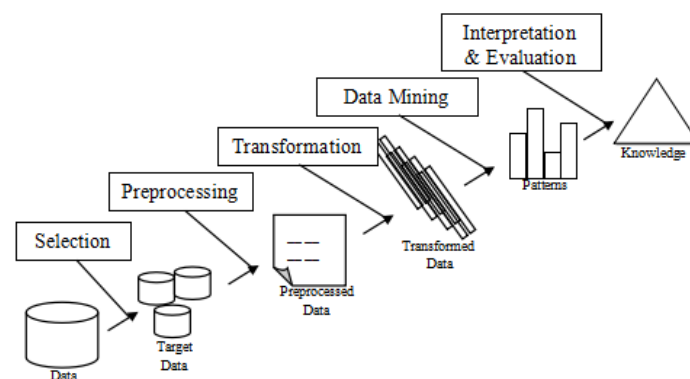
Data Mining merupakan proses pengambilan informasi baru diambil dari sebuah data besar yang membantu dalam pengambilan keputusan atau sebagai proses pengumpulan informasi penting dari suatu data yang besar untuk diubah menjadi sebuah pengetahuan (Sembiring, 2019). Data mining merupakan area yang mengintegrasikan berbagai macam metode, meliputi statistika, basis data, kecerdasan buatan, *machine learning*, pengenalan pola, serta pemodelan mengenai ketidakpastian. Gambar 2.1 menunjukkan hal tersebut (Wardani, 2020).



Gambar 2.1 *Data Mining* merupakan irisan beberapa bidang

Sumber : (Nirwani, 2020)

Tujuan dari penerapan *data mining* yaitu untuk menemukan, menggali atau menambah pengetahuan dari data atau informasi yang kita miliki (Tahyudin Imam, 2014). Proses dalam mendapatkan nilai tambah dari sekumpulan data sering dikenal juga sebagai penemuan pengetahuan dari pangkalan data (*Knowledge Discovery in Database*=KDD). *Knowledge Discovery in Database* (KDD) yaitu tahapan yang dilakukan untuk mendapatkan pengetahuan dari sekumpulan data (Puspitasari & Juliono, 2019).



Gambar 2.2. Proses *Knowledge Discovery in Database* (KDD)

Sumber : (Puspitasari & Juliono, 2019)

Berdasarkan Gambar 2.2, dapat diuraikan sebagai berikut :

- a. Menggali pengetahuan awal serta sasaran pengguna dan memahami dominan aplikasi untuk mengetahui
- b. Membuat target data-set meliputi fokus *subset* data dan pemilihan data.
- c. Transformasi data dan pembersihan meliputi eliminasi *derau*, *outlier*, *missing value* serta pemilihan fitur dan reduksi dimensi.
- d. Penggunaan algoritma dari *data mining* yang terdiri asosiasi, sekuensial, klasifikasi dan lain-lain.
- e. Interpretasi, evaluasi dan visualisasi pola dalam melihat apakah ada sesuatu yang baru dan menarik serta dilakukan iterasi jika diperlukan (Setiawan, 2016).

Data mining dikenal dengan istilah *pattern recognition*. Menyebutkan bahwa KDD, merupakan proses terstruktur, yaitu sebagai berikut :

1. *Data selection* pemilihan (seleksi) data dari serangkaian data operasional harus dilakukan sebelum tahap penggalian atau ekstrasi informasi dalam KDD dimulai. Data yang terpilih akan digunakan untuk proses *data mining* disimpan dalam suatu *file*, terpisah dari basis data operasional
2. *Data cleaning* adalah proses membersihkan data dari data *noise* dan tidak konsisten
3. *Data integration* adalah proses untuk menggabungkan dari beberapa sumber berbeda.
4. *Pattern evaluation* adalah proses mengidentifikasi pola.
5. *Data transformation* adalah proses mengubah bentuk data menjadi data yang sesuai untuk proses *mining*.
6. *Knowledge presentation* mempresentasikan informasi yang dibutuhkan, proses di mana informasi yang telah didapat kemudian digunakan oleh pemilik data (I. M. Sari dkk., 2017).

Menurut Larosse, pembagian *data mining* menjadi beberapa kelompok berdasarkan tugas yang dilakukan, yaitu

a. Deskripsi

Penelitian atau analisis melakukan percobaan untuk mendapatkan cara mendeskripsikan pola serta kecenderungan pada data di mana

deskripsi tersebut seringkali memberi kemungkinan penjelasan terhadap satu pola.

b. Prediksi

Prediksi merupakan memperkirakan nilai masa yang akan datang. Teknik ini hampir sama dengan klasifikasi kecuali dalam prediksi nilai *item* dari hasil akan ada di masa mendatang.

c. Estimasi

Estimasi hampir mirip dengan klasifikasi, hanya saja yang membedakan yaitu variabel target estimasi cenderung mengarah kepada numerik dibanding mengarah ke kategori.

d. Klasifikasi

Suatu Teknik yang mencakup perilaku dan atribut dari kelompok yang sudah diidentifikasi, Teknik ini memungkinkan memberikan klasifikasi pada data baru dengan memanipulasi data yang ada dan telah diklasifikasi menggunakan hasilnya untuk memberikan sejumlah aturan,

e. Asosiasi

Analisis asosiasi yaitu teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. Analisis asosiasi juga sering disebut *market basket analysis*.

f. Pengelompokan

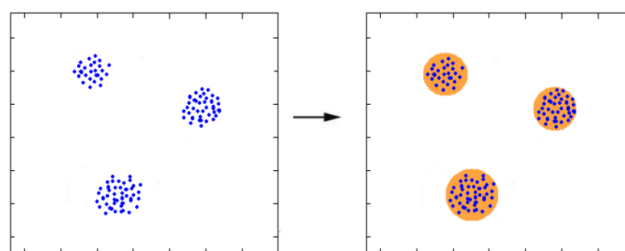
Pengelompokan yaitu mengelompokkan *record*, mengamati dan membangun kelas dari objek yang memiliki kesamaan. Kluster merupakan sekumpulan data yang saling mempunyai kesamaan dan mempunyai ketidaksamaan terhadap kluster yang berbeda (Sulastri & Gufroni, 2017).

2.2.2. Clustering

Clustering merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*) yang memiliki arti bahwa karakteristik tiap kluster

tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut - atribut dari suatu kelompok. Data *mining* ada dua jenis metode *clustering* yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering*. Metode *clustering* berusaha untuk menempatkan objek yang mirip dalam satu *cluster* dan membuat jarak antar *cluster* sejauh mungkin. Artinya objek dalam satu *cluster* sangat mirip satu dengan lain dan berbeda dengan objek dalam *cluster-cluster* yang lain. Metode *non - hierarchical (Partitioning)* dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan. Setelah jumlah *cluster* diketahui baru proses *cluster* dilakukan metode ini biasa disebut dengan *K-means*.

Tujuan utama dari metode *clustering* yaitu mengelompokkan banyak informasi atau objek-objek ke dalam *cluster* (grup), sehingga untuk setiap *cluster* ada berisi informasi yang hampir mirip atau semirip mungkin, yang berarti data atau objek dalam satu *cluster* sangat mirip satu sama lain dan berbeda dengan objek dalam *cluster-cluster* lain (K. Sari, 2020). Ada dua metode *clustering* yang dikenal saat ini yaitu *hierarchical clustering* dan *partitioning clustering*. Pada metode *hierarchical clustering* terdiri dari *complete linkage clustering*, *single linkage clustering*, *average linkage clustering* serta *centroid linkage clustering* sedangkan metode *partitioning* terdiri dari *K-means* serta *fuzzy K-Means* (Paramartha dkk., 2017).



Gambar 2.3 contoh proses *clustering*

Sumber (K. Sari, 2020)

Menurut Hair Jr. dkk (2017), terdapat lima macam algoritma untuk membentuk kelompok dengan metode hirarki, yaitu:

a. *Single Linkage* (pautan tunggal)

Metode pautan tunggal didasarkan atas jarak minimum. Dua objek yang memiliki jarak terdekat dikelompokkan. Apabila terdapat objek ketiga yang memiliki jarak terdekat dengan salah satu obyek dalam kelompok yang telah terbentuk, maka obyek tersebut dapat digabung ke dalam kelompok tersebut.

b. *Complete Linkage* (pautan Lengkap)

Metode ini pada dasarnya sama dengan metode pautan tunggal. Hanya saja jarak yang digunakan adalah jarak maksimal. Alasan dipilih jarak maksimum adalah agar objek-objek yang mempunyai sedikit kesamaan dapat dihubungkan.

c. *Avarage- Linkage* (pautan rata-rata)

Metode pautan rata-rata juga memiliki kesamaan dengan dua metode pautan tunggal. Hanya jarak yang digunakan adalah rata-rata jarak dari semua objek dalam satu kelompok dengan objek lain di luar kelompok tersebut. Metode ini mengelompokkan dua objek berdasarkan jarak rata-rata yang didapat dengan melakukan rata-rata semua jarak antar objek terlebih dahulu.

d. *Ward's method*

Dalam metode *Ward*, perhitungan jarak berdasarkan jumlah kuadrat anatara dua kelompok untuk semua peubahnya. Metode ini dapat digunakan apabila jumlah pengamatan tidak terlalu besar. Pada umumnya jarak yang digunakan adalah jarak kuadrat *Dynamic Time Wrapping* (DTW) dan jarak kuadrat *Euclidean*. Metode *Ward's* bertujuan untuk memperoleh *cluster* yang memiliki varians internal *cluster* yang sekecil mungkin.

e. *Centroid method* (metode titik pusat)

Jarak yang digunakan dalam metode ini adalah jarak (kuadrat *Euclidean*) antara titik pusat dua kelompok. Titik pusat kelompok ini adalah nilai tengah objek setiap peubah dalam satu kelompok. Dalam metode ini setiap kali terbentuk kelompok baru, maka titik pusatnya

berubah. Keuntungan dari metode ini adalah kecilnya pengaruh penciran dalam pembentukan kelompok.

2.2.3. Analisis Runtun Waktu (*Time Series*)

Analisis *time series* pertama kali di perkenalkan dan dikembangkan yaitu pada tahun 1970, digunakan untuk memprediksi struktur atau mekanisme yang mendasari terciptanya data runtun waktu. Data *time series* yang dimaksud merupakan hasil pengamatan atas variabel yang terjadi pada kurun waktu tertentu. Misalkan notasi z_t merupakan pengamatan dalam sebuah data *time series* saat t maka runtut waktu saat n pengamatan dinyatakan dengan z_1, z_2, \dots, z_n . Dalam analisis *time series* diberikan perkiraan bahwa setiap pengamatan korelasi atau hubungan. Agar dapat melihat adanya korelasi dapat dilakukan dengan uji korelasi antar pengamatan yang sering dikenal dengan *Autocorrelation function* (ACF) dan *parsial autocorrelation function* (PACF) (Alim dkk., 2021).

Menurut Makridakis dan Wheelwright (1999), pola pada data *time series* dibedakan menjadi empat jenis yaitu:

1. Pola Musiman (S)

Pola musiman terjadi jika data membentuk pola yang teratur yang bersifat musiman

2. Pola Horizontal (H)

Pola horizontal terjadi apabila data mengalami fluktuasi disekitar rataaan.

3. Pola Tren (T)

Pola tren terjadi jika data mengalami kenaikan atau penurunan.

4. Pola Siklik (C)

Pola siklis terjadi apabila data mengalami fluktuatif jangka Panjang.

2.2.4. Analisis *Clustering Time Series*

Cluster time series merupakan metode pengelompokan data *time series* yang dapat digunakan dalam mengelompokan daerah-daerah berdasarkan variabel yang ada. Analisis *cluster* dapat menggunakan jarak *Dynamic Time*

Warping (DTW) dan jarak *Euclid*. Penggunaan kedua jarak untuk dibandingkan untuk mengetahui hasil pengelompokan terbaik (Putu Widya Adnyani & Robinson Sihombing, 2021). Tujuan dari pengelompokan data *time series* sebagai berikut :

1. Data *time series* berisi informasi berharga yang dapat diolah melalui penemuan pola, pengelompokan merupakan solusi umum untuk mengungkap pola-pola pada data *time series*.
2. Data *time series* yang sangat besar tidak mudah untuk digali informasinya oleh karena itu lebih memilih untuk menggunakan data yang sudah dikelompokkan
3. Pengelompokan *time series* merupakan pendekatan yang lebih banyak digunakan sebagai teknik eksplorasi dan juga sebagai sub-rutin di lebih algoritma data yang kompleks
4. Dapat mempermudah pengguna memahami struktur data (Aghabozorgi dkk., 2015).

Langkah-langkah analisis data yang dilakukan dalam penelitian ini adalah sebagai berikut:

- a. Melakukan Eksplorasi data
- b. Melakukan *clustering* dengan menggunakan ukuran jarak *Dynamic Time Wrapping* (DTW)

Jarak DTW adalah jarak minimum antar dua pasang titik yang memperhitungkan pergeseran atau perpindahan. Misalkan mempunyai dua deret waktu, yaitu $Q = q_1, q_2, \dots, q_i, \dots, q_n$ dan $r = r_1, r_2, \dots, r_j, \dots, r_m$ penerapan ukuran jarak DTW menggunakan perhitungan melalui matriks $n \times m$ dimana elemen (i, j) dari matriks tersebut terdapat jarak (q_i, r_j) . Jarak *Euclidean* digunakan untuk menghitung lintasan *warping* dengan $W = w_1, w_2, \dots, w_k$ dimana $\max(m, n) \leq k \leq m + n - 1$ merupakan himpunan elemen matriks yang memenuhi tiga kendala yaitu kondisi batas, kontinuitas dan monotonisitas. Kendala kondisi atas membutuhkan lintasan *warping* untuk memulai dan menyelesaikan

perhitungan secara diagonal berlawanan sudut sel dari matriks, dengan $W_l = (1,1)$ dan $W_k = (m,n)$. jalur *wrapping* rumusnya dapat dilihat di persamaan (i)

$0 < i \leq n$ dan $0 < j \leq m$, kita dapat mendefinisikan sebagai :

$$d_{DTW} = \text{Min} \frac{\sum_{k=1}^K w_k}{K} \quad (\text{i})$$

dengan:

d_{DTW} = jarak *dynamic time warping*

w_k = Lintasan *Wrapping ke- K*

c. Melakukan perhitungan *multidimensional scaling*

Multidimensional scaling (MDS) merupakan analisis multivariat yang menunjukkan hubungan antar sejumlah objek dalam ruang multidimensional didasarkan pada kemiripan / kedekatan objek. Kedekatan antar objek diperoleh menggunakan jarak *Dynamic Time Wrapping* (DTW). MDS digunakan untuk mengetahui hubungan interdependensi atau saling ketergantungan antar variabel atau data.

d. Ukuran Ketepatan Banyak *Cluster*

Pengujian sebuah model dilakukan bertujuan untuk mendapatkan informasi sedekat apa hubungan antara satu objek dengan objek lain pada sebuah *cluster* dan seberapa jauh antara sebuah *cluster* dengan *cluster* yang lainnya. Banyaknya cluster yang tepat perlu diketahui agar pemodelan yang dilakukan memiliki tingkat keakuratan yang tinggi atau *error* yang rendah. Banyaknya kluster pada *time series* adalah koefisien *silhouette* dengan rumus:

$$S(i) = 1 + \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (\text{ii})$$

keterangan $a(i)$ adalah rata-rata jarak antar anggota dalam kluster $b(i)$ adalah jarak terkecil antar anggota kluster yang berdekatan. Banyaknya kluster yang tepat ditandai dengan nilai $s(i)$ yang semakin mendekati 1. Menurut Kaufman dan Rousseeuw (1990) interpretasi koefisien *silhouette* dapat dinyatakan seperti pada tabel 2.2.

Tabel 2.2. Kategori Pengelompokan berdasarkan koefisien *silhouette*

No	Koefisien <i>Silhouette</i>	Kategori pengelompokan
----	-----------------------------	------------------------

1	0.71-1.00	<i>Strong Classification</i>
2	0.51-0.70	<i>Good Classification</i>
3	0.26-0.50	<i>Weak Classification</i>
4	0.00-0.25	<i>Bad Classification</i>

Adapun *silhouette coefficient* terdapat pada angka antara nilai -1 sampai 1 dengan nilai *silhouette coefficient* semakin mendekati nilai 1, maka semakin bagus pengelompokan objek-objek kedalam sebuah *cluster* dan sebaliknya jika *silhouette coefficient* sudah mendekati angka -1, maka akan makin buruk metode pengelompokan datanya pada *cluster* tersebut, metode pengukuran ini menggabungkan metode *cohesion* dengan *separation* (K. Sari, 2020).