

Implementasi Metode Agglomerative Hierarchical Clustering Pada Segmentasi Naskah Aksara Jawa Dengan Complete Linkage

Paguh Esatrio

195314146

Prodi Informatika, Fakultas Sains & Teknologi Yogyakarta

e-mail: paguh2020@gmail.com

Abstrak

Naskah aksara Jawa atau Hanacaraka/Carakan merupakan salah satu Aksara tradisional yang ada di Indonesia. Aksara Jawa digunakan oleh masyarakat Jawa. Hanacaraka umumnya digunakan untuk menulis naskah seperti cerita (serat), catatan sejarah (babad), tembang kuno (kakawin), atau ramalan (primbon). Namun Seiring berjalannya waktu, maka tidak menutup kemungkinan, bahwa pengetahuan atau informasi yang ada pada Naskah aksara tersebut bisa saja hilang, seperti tulisan yang mulai memudar atau naskah tersebut menjadi lapuk.

Penelitian ini akan dilakukan pengelompokan aksara jawa dengan menggunakan metode pendekatan Agglomerative Hierarchical Clustering untuk menentukan pengelompok dari setiap citra aksara jawa. Terdapat beberapa Langkah yang dilakukan pada penelitian ini. Yang pertama adalah pengumpulan data yang berupa buku aksara jawa yang diambil di Perpustakaan Pustaka Artati Universitas Sanata Dharma Kampus II. Kemudian melakukan preprocessing untuk mendapatkan data yang bersih. Setelah itu akan dilakukan segmentasi untuk mendapatkan tiap aksaranya. Lalu akan di ambil ciri untuk tiap aksaranya dengan menggunakan Intensity of Character menggunakan ekstrakciri_3unit dan ekstrakciri_4unit. Kemudian Data ciri yang telah didapatkan dikelompokkan menggunakan Agglomerative Hierarchical Clustering dengan Complete Linkage yang mana akan menghasilkan cluster-cluster untuk tiap data ciri aksara yang kemudian dibuat dendogram dan dibandingkan antara dendogram ekstrakciri_3unit dan dendogram ekstrakciri_4unit akan dibandingkan.

Dari pengujian yang telah yaitu membandingkan data ciri dari metode Agglomerative Hierarchical Clustering menggunakan Intensity of Character mendapatkan hasil pengelompokan yang bagus dari 2 metode tersebut untuk menentukan kelompok dari setiap citra aksara jawa untuk ekstrakciri_3unit didapat hasil akhir 162.65 dan ekstrakciri_4unit didapat hasil akhir 247.168 yang artinya ekstrakciri_3unit lebih baik

Kata kunci: Aksara Jawa, Clustering, Agglomerative Hierarchical Clustering, Intensity of Character, Complete Linkage

1. PENDAHULUAN

Naskah aksara Jawa atau Hanacaraka/Carakan merupakan salah satu Aksara tradisional yang ada di Indonesia. Aksara Jawa digunakan oleh masyarakat Jawa. Hanacaraka umumnya digunakan untuk menulis naskah seperti cerita (serat), catatan sejarah (babad), tembang kuno (kakawin), atau ramalan (primbon). Namun Seiring berjalannya waktu, maka tidak menutup kemungkinan, bahwa pengetahuan atau informasi yang ada pada Naskah aksara tersebut bisa saja hilang, seperti tulisan yang mulai memudar atau naskah tersebut menjadi lapuk.

Dengan semakin majunya perkembangan teknologi khususnya dalam penerapan pengolahan citra dokumenterpikirlah sebuah ide untu membuat program yang dapat membaca aksara yang ada pada naskah dengan aksara Jawa tersebut. Ada beberapa tahap yang harus

dilakukan agar dapat membentuk sistem tersebut, dan salah satunya ialah membentuk *database* berdasarkan aksara yang ada pada naskah tersebut

Namun sebelum itu harus dilakukan segmentasi citra yang berfungsi mengubah representasi dari suatu citra menjadi sesuatu yang lebih berarti dan mudah untuk dianalisis yang dimana Terdapat 825 citra aksara Jawa yang mana aksara tersebut susah untuk dikenali agar dijadikan satu dengan kelompok aksaranya sendiri, yang berefek pada pemberian label pada masing-masing aksara tersebut.

Untuk mengatasi masalah pengelompokan tersebut, maka dapat diselesaikan dengan menggunakan Segmentasi citra dengan menggunakan metode clustering dapat digunakan dengan berbagai macam algoritma, salah satunya adalah Agglomerative Hierarchical Clustering (AHC) metode *clustering*, sehingga citra aksara-aksara tersebut dapat berkumpul dengan aksara yang sejenisnya, yang mana hal ini akan membuat pemberian label menjadi lebih mudah pada tahap pembuatan *database*. Metode AHC mempunyai kemampuan menggabungkan data dengan membuat hirarki, dimana yang memiliki kemiripan akan ditempatkan di hirarki yang berdekatan dan yang tidak memiliki kemiripan ditempatkan pada hirarki berjauhan (Donny, 2013). Pendekatan yang dilakukan oleh metode AHC berupa *bottom-up* yang berarti proses pengelompokan data akan dimulai dari yang terkecil hingga ke pengelompokan yang besar.

Pada salah satu penelitian mengenai implementasi metode ini yaitu pengelompokan citra aksara bali menggunakan metode single linkage hierarchical clustering oleh David Thanlian Kurniawan pada tahun 2020 Dalam penelitian ini menjelaskan dari Penggunaan ciri mempengaruhi penggunaan rumus jarak agar terbentuknya cluster yang dikatakan layak (nilai silhouette coefficient sama dengan 0,51). Seperti pada IoT ukuran 7x7, pengelompokan cluster yang dikatakan layak akan lebih efektif terbentuk dengan menggunakan jarak City Block, sedangkan IoT ukuran 4x4, 5x5, 6x6 dan 8x8, pengelompokan cluster yang dikatakan layak akan lebih efektif terbentuk dengan menggunakan jarak Euclidean.

Terdapat penelitian lain mengenai implementasi metode ini yaitu Implementasi Algoritma Agglomerative Hierarchical Clustering Untuk Mengelompokkan Capaian Belajar Siswa SD oleh Aloysius Ari Kurniawan pada tahun 2017. Dalam penelitian ini menjelaskan dari 3 metode yang memberikan hasil paling seimbang yaitu metode *average linkage*. Hasil perhitungan *silhouette index* mendapatkan hasil sebanyak seluruh data sebanyak 14 berada tepat di *cluster*, kemudian *cluster* 2 ada sebanyak 10 data yang tepat dari 22 data dan *cluster* 3 sebanyak 4 dari 31 data. Disimpulkan bahwa metode Agglomerative Hierarchical Clustering dengan baik mengelompokkan data dari nilai siswa dan menghasilkan *cluster* yang cukup seimbang untuk digunakan sebagai acuan pengelompokan kelas siswa.

Penelitian lain tentang penerapan Agglomerative Hierarchical Clustering untuk segmentasi pelanggan oleh Widyawati yang dipublikasi pada januari 2020. Dijelaskan bahwa dalam penelitian ini metode yang digunakan yaitu *agglomerative hierarchical clustering* berhasil dalam melakukan pengelompokan pelanggan dan dapat dikaitkan pada strategi pemasaran.

Penelitian lain tentang implementasi Metode Agglomerative Hierarchical Clustering Pada Segmentasi Pelanggan Barbershop (Studi Kasus : RichDjoe Barbershop Malang) oleh Rhayhana Putri Justitia, Nurul Hidayat, dan Edy Santoso menyimpulkan bahwa Penerapan metode Agglomerative Hierarchical Clustering pada penelitian ini menghasilkan 2 parameter jarak terbaik yaitu *single linkage* dan *average linkage*. Kedua parameter jarak ini dalam proses pengelompokan data memberikan hasil paling optimal pada titik potong 391 dengan jumlah 8 *cluster*.

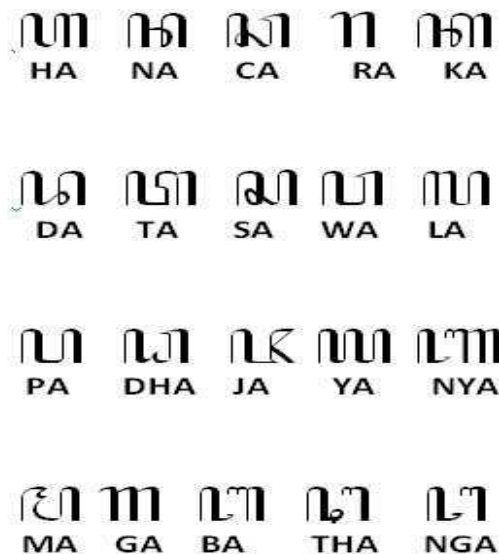
Penelitian lain tentang implementasi Metode Clustering Pada Data Sentimen Bpjs Kesehatan Menggunakan Algoritma Agglomerative Hierarchical Clustering Complete Linkage oleh Tinton Aji Sadewo, Purba Daru Kusuma, dan Casi Setianingsih Pada penerapan metode agglomerative hierarchical clustering Complete Linkage mendapatkan hasil pengujian pada penelitian ini dengan data negatif yang memiliki nilai jarak terbaik dan paling optimal. Pada data negatif memberikan hasil yang optimal pada jumlah cluster 8 serta titik potong 683.

2. METODE PERCOBAAN

2.1 LANDASAN TEORI

2.1.1 Aksara Jawa

Aksara Jawa, atau dikenal dengan nama Hanacaraka adalah aksara jenis abugida turunan aksara Brahmi yang digunakan atau pernah digunakan untuk penulisan naskah-naskah berbahasa Jawa, bahasa Makasar, bahasa Sunda, dan bahasa Sasak Bentuk aksara Jawa yang sekarang digunakan (modern) sudah tetap sejak masa Kesultanan Mataram (abad ke-17) tetapi bentuk cetaknya baru muncul pada masa zaman ke-19. Aksara ini adalah modifikasi dari aksara Kawi atau dikenal dengan Aksara Jawa Kuno yang juga merupakan abugida yang digunakan sekitar masa zaman ke-8 – masa zaman ke-16. Menurut Hadiprijono (2013) aksara Jawa terdiri dari 20 aksara, yaitu dari aksara ha sampai nga, seperti pada Gambar 1



Gambar 2.1.1 Gambar Pokok Aksara Jawa

2.1.2 Pengertian Citra

Citra atau gambar adalah suatu bentuk fungsi kontinu dari intensitas dalam bidang dua dimensi (2-D), pada citra dapat dinyatakan secara matematis :

$$0 < f(x, y) < \infty \quad (1)$$

Yang mana $f(x,y)$ merupakan intensitas cahaya pada citra yang terletak dilokasi (x,y) . Citra digital dapat diartikan sebagai citra $f(x,y)$ yang nilai nya berbentuk diskrit, baik untuk nilai koordinat (x,y) maupun nilai intensitas cahaya nya. Citra digital digambarkan sebagai matrik yang terdiri atas baris dan kolom, yang mana setiap baris dan kolom akan berisikan elemen atau nilai intensitas kecerahan tertentu untuk setiap $f(x,y)$ atau *pixel* (titik) pada bidang citra digital (Widiarti & Himamunanto, 2013).

2.1.3 Preprocessing

Tahapan *preprocessing* berguna untuk menyiapkan data seperti mengubah ukuran citra,

membuang *noise*, atau memisahkan latar belakang citra dengan objek nya, sebelum citra tersebut masuk ke tahap ekstraksi ciri. Pada tahapan ini, akan banyak mengimplementasikan metode-metode pemrosesan citra dalam mengolah citra aksara.

2.1.4 Binerisasi

Binerisasi adalah tahap pertama dalam pengolahan data, dimana tahapan ini memisahkan objek yang ada di citra dengan latar belakang (*background*) yang tidak dibutuhkan. Tujuan utama dari proses ini adalah secara otomatis menentukan nilai ambang *threshold* yang akan membagi citra dalam dua kelompok, yaitu kelompok obyek dan kelompok latar belakang (*background*) (Widiarti & Himamunanto, 2013).

2.1.5 Filtering

Filtering merupakan tahapan untuk dilakukan nya reduksi derau atau *noise reduction* (Widiarti & Himamunanto, 2013). Adapun jenis-jenis *noise* yang mungkin dapat terjadi antara lain :

- a) *Gaussian noise* : *noise* berupa titik-titik yang berwarna yang jumlahnya sama dengan presentase *noise*.
- b) *Speckle noise* : *noise* berupa warna hitam pada titik yang terkena *noise*.
- c) *Salt and pepper noise* : *noise* berupa warna putih pada titik yang terkena *noise*.

2.1.6 Segmentasi

Segmentasi merupakan proses pemecahan citra untuk memperoleh objek-objek yang terkandung dalam citra tersebut. Segmentasi terbagi menjadi dua tahap. Tahap pertama pemisahan objek dengan objek lain yang berada pada baris yang sama, lalu pada tahap kedua, yaitu mendapatkan huruf tersebut secara individu dengan mendapatkan kolom-kolom karakter citra aksara dari baris yang telah dipotong pada tahap pertama. Adapun salah satu metode umum yang dapat digunakan untuk proses segmentasi ini, yaitu Profil Proyeksi.

Profil proyeksi merupakan tahapan segmentasi citra, yang mana hasil dari tahapan ini ialah mendapatkan objek yang terpisah dari latar belakang nya dengan cara memotong secara vertical dan horizontal. Profil proyeksi vertical berguna untuk mendapatkan baris awal pixel objek hingga baris akhir pixel objek, sedangkan profil proyeksi horizontal berguna untuk mendapatkan kolom awal pixel objek hingga kolom akhir pixel objek (Widiarti & Himamunanto, 2013).

2.1.7 Ekstraksi Ciri

Ekstraksi ciri merupakan tahapan untuk memperoleh sifat pola dari suatu objek (dalam hal ini yaitu citra aksara Jawa) yang ada pada citra, yang mana menjadikan objek tersebut berbeda dengan objek lainnya (Widiarti & Himamunanto, 2013).

2.1.8 Agglomerative Hierarchical Clustering

Hierarchical Clustering merupakan metode *clustering* yang digunakan untuk mengorganisir data dalam bentuk *tree* (Han & Kamber, 2006). Jenis *Hierarchical* yang digunakan pada penelitian ini ialah Bottom-Up (Agglomerative). Pada *Agglomerative hierarchical clustering*, setiap data ke-*i* diperlakukan sebagai suatu *cluster* tersendiri, lalu dilakukannya iterasi penggabungan antar data hingga membentuk suatu *cluster* tunggal. Adapun langkah pengerjaan nya :

1. Menghitung Matrik Jarak antar data.
2. Menggabungkan dua kelompok terdekat berdasarkan parameter kedekatan

yang ditentukan.

3. Memperbaharui Matrik Jarak antar data untuk merepresentasikan kedekatan diantara kelompok baru dan kelompok yang masih tersisa.
4. Mengulangi langkah 2 dan 3 hingga hanya satu kelompok yang tersisa.

2.1.9 Single linkage

Single linkage merupakan sebuah parameter jarak berdasarkan jarak paling kecil tiap data. Algoritma parameter jarak *single linkage* dimulai dengan memilih jarak terkecil dalam matriks data.

$$d_{(UV)W} = \min(d_{UW}, d_{VW}) \quad (2)$$

2.1.10 Complete Linkage

Complete linkage merupakan sebuah parameter jarak berdasarkan jarak paling besar tiap data. Algoritma parameter jarak *complete linkage* dimulai dengan memilih jarak terbesar dalam matriks data.

$$d_{(UV)W} = \max(d_{UW}, d_{VW}) \quad (3)$$

2.1.11 Average Linkage

Average Linkage merupakan sebuah parameter jarak berdasarkan jarak rata-rata tiap data. Algoritma parameter jarak *Average Linkage* dimulai dengan memilih jarak rata-rata dalam matriks data.

$$d_{(UV)W} = \frac{(d_{UW} + d_{VW})}{n(UV) + n(W)} \quad (4)$$

2.1.12. Euclidean Distance

Euclidean distance merupakan jarak antara dua titik pada garis lurus. Metode perhitungan ini menggunakan rumus Pythagoras, yaitu perhitungan jarak antar data yang sering digunakan dalam machine learning. [4] Rumus dari Euclidean Distance sebagai berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (5)$$

Dimana:

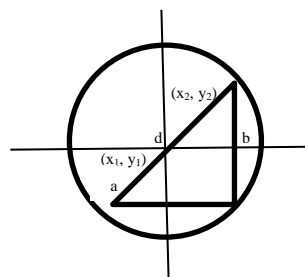
d_{ij} = perhitungan jarak untuk kemiripan

n = jumlah vector

x_{ik} = vector citra masukan

x_{jk} = vector citra pembanding

Dari persamaan rumus diatas, bentuk dari perhitungan jarak Euclidean Distance adalah lingkaran yang ditunjukkan pada gambar berikut:



Gambar 2.1.12 Rumus Pythagoras

Keterangan:

$$a = x_2 - x_1$$

$$b = y_2 - y_1$$

Rumus Pythagoras:

$$1. \quad a^2 + b^2 = d^2 \quad (6)$$

$$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 \quad (7)$$

2.2 Algoritma

2.2.1 Preprocessing

1. Binarisasi Citra yaitu pertama-tama baca data citra yang akan diolah menggunakan function imread dan akan disimpan pada sebuah variabel bernama img, lalu data yang tersimpan dalam variabel img akan diubah menjadi citra hitam-putih dengan function im2bw. Lalu hasil binerisasi akan disimpan pada variabel img.
2. Filtering citra yaitu pertama-tama data citra biner yang tersimpan dalam variabel img akan dilakukan proses filtering dengan function bwareaopen. Lalu hasil filtering akan disimpan pada variabel img.

2.2.2 Segmentasi

1. Setelah selesai pada tahap preprocessing, lalu dilakukan pengambilan karakter untuk tiap baris menggunakan function bwlable(). Hasilnya akan disimpan ke dalam variabel bernama T.
2. Kemudian lakukan proses pengukuran untuk setiap karakter yang didapat menggunakan function regionprops lalu disimpan didalam variabel property, lalu lakukan proses boundingbox untuk setiap karakter yang didapat setelah diukur pada proses sebelumnya menggunakan function rectangle yang disimpan didalam variabel img.
3. Kemudian memotong setiap karakter yang telah diboundingbox, lalu hasil perpotongan akan disimpan kedalam folder pc milik kita menggunakan function imwrite yang disimpan kedalam variabel Bernama ~x1.
4. Lakukan tahap resize citra pada tahap ini hasil semua segmentasi pada tahap 3 akan diubah ukurannya. Tahap ini dilakukan agar seluruh data memiliki ukuran yang sama sehingga mempermudah proses ekstraksi ciri. Pada tahapan ini digunakan function autoresize yang merupakan function yang saya buat sendiri.

2.2.3 Ekstrasi Ciri

Lakukan Intensity of Character Pada tahap ini saya menggunakan function Intensity_Of_Character yang saya telah buat, dimana data aksara yang telah melalui tahap preprocessing dan segmentasi akan dilakukan proses ekstraksi ciri dengan IoC. Perhitungan IoC dilakukan dengan mengubah matrik hasil segmentasi menjadi matriks baru dengan ukuran NxN dengan cara menambah jumlah piksel hitam sebanyak ukuran gambar dibagi dengan ukuran matriks IoC. Banyak data citra yang dipakai adalah 733. Ukuran IoC yang digunalkan adalah 3x3 yang menghasilkan 9 koIom dan 733 baris, 4x4 yang menghasilkan 16 kolom.

1. Baca data-dat yang telah segmentasi yang disimpan didalam folder
2. Bagi aksara menjadi 9/16 bagian menggunakan function ekstrakciri_3unit / ekstrakciri_4unit
3. Cari jumlahan piksel yang bernilai 0 disetiap bagian karakter
4. Simpan ciri tersebut menjadi ciri dari karakter bersangkutan ke dalam file DataEkstraksiCiri_3unit / DataEkstraksiCiri_4unit

2.2.4 Menghitung Euclidean Distance

1. Membaca data ciri dari karakter bersangkutan yang disimpan di file DataEkstraksiCiri_3unit / DataEkstraksiCiri_4unit.
2. Menghitung Euclidean distance data ciri dari karakter bersangkutan menggunakan function squareform yang disimpan didalam variable distance_matrix.

2.2.5 Klustering Agglomerative Hierarchical Clustering complete Linkage

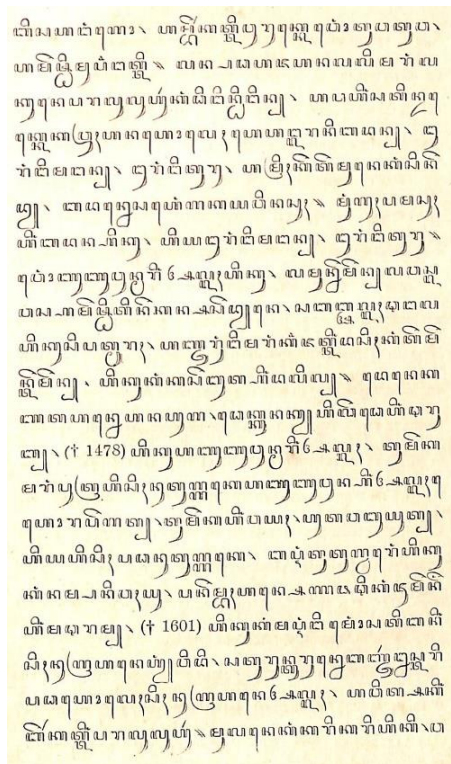
1. Setelah menghitung Euclidean Distance data ciri dari karakter bersangkutan maka selanjutnya akan masuk ketahap Klustering Agglomerative Hierarchical Clustering Complete Linkage menggunakan function linkage ('Complete').
2. Kemudian tahap selanjutnya membuat dendrogram dari hasil Klustering Agglomerative Hierarchical Clustering Complete Linkage untuk data ciri dari karakter bersangkutan menggunakan function dendrogram.

2.2.6 Evaluasi Clustering untuk Intensity of Character (matriks 3x3) dan Intensity of Character (matriks 4x4)

1. Evaluasi Klustering untuk Intensity of Character (matriks 3x3) pada tahap ini data hasil clustering berupa dendrogram menggunakan Agglomerative Hierarchical Clustering Complete Linkage untuk tiap cluster.
2. Evaluasi Klustering untuk Intensity of Character (matriks 4x4) pada tahap ini data hasil clustering berupa dendrogram menggunakan Agglomerative Hierarchical Clustering Complete Linkage untuk tiap cluster.
3. Bandingkan hasil evaluasi pada tahap ini hasil clustering berupa dendrogram dari Intensity of Character (matriks 3x3) akan dibandingkan hasil clustering berupa dendrogram dari Intensity of Character (matriks 4x4), untuk melihat clustering yang ideal menggunakan yang mana.

2.3 Data Penelitian

Data yang dipakai merupakan buku aksara jawa yang diambil di Perpustakaan Pustaka Artati Universitas Sanata Dharma Kampus II yang digunakan dalam penelitian ini:



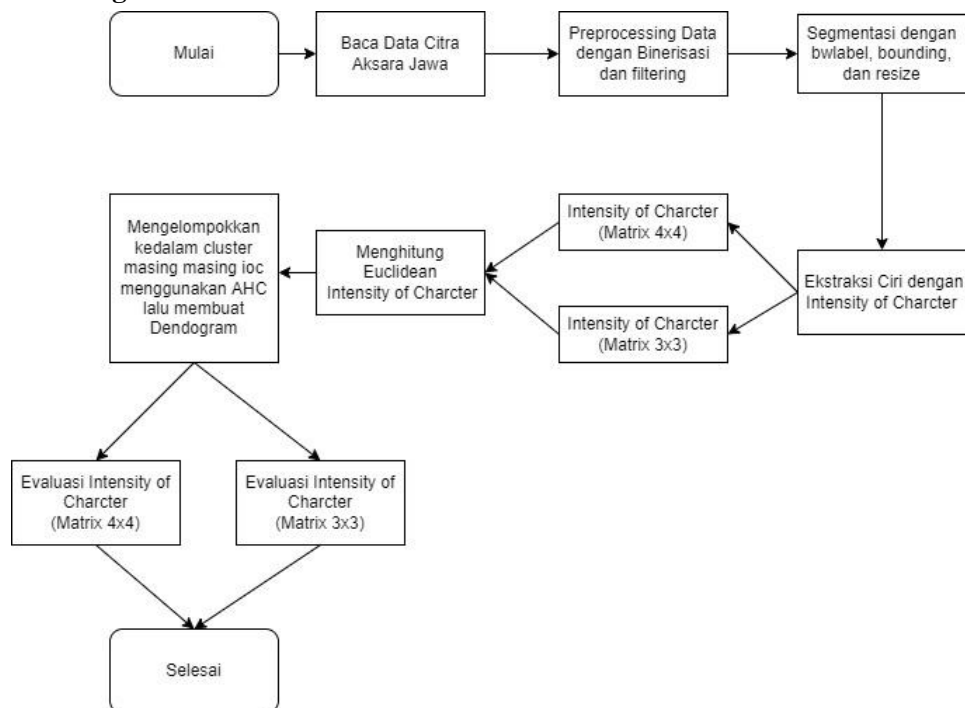
Gambar 2.3 Citra Aksara Jawa (1)

Data yang digunakan studi kasus penelitian ini adalah salah satu halaman dari sebuah buku dengan tulisan aksara jawa yang dimana jumlah aksara setelah segmentasi dan di potong dengan menggunakan bounding box yang berjumlah 733 aksara



Gambar 2.3 Citra Perpotongan Aksara Jawa (2)

2.4 Desain Program



Gambar 4 Flowchart

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

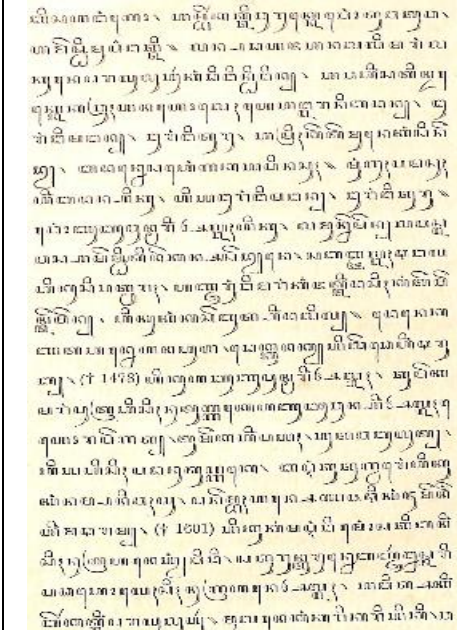
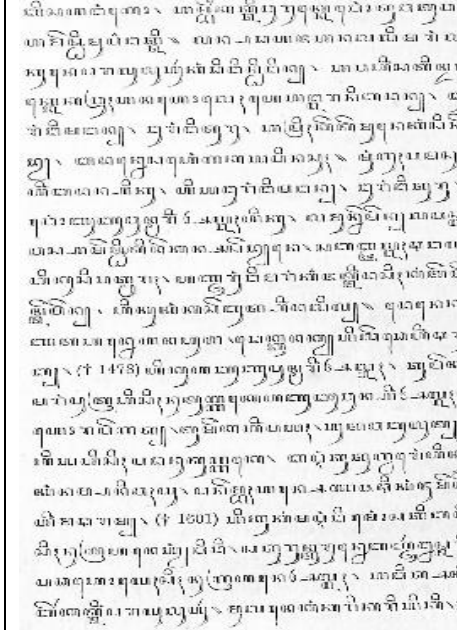
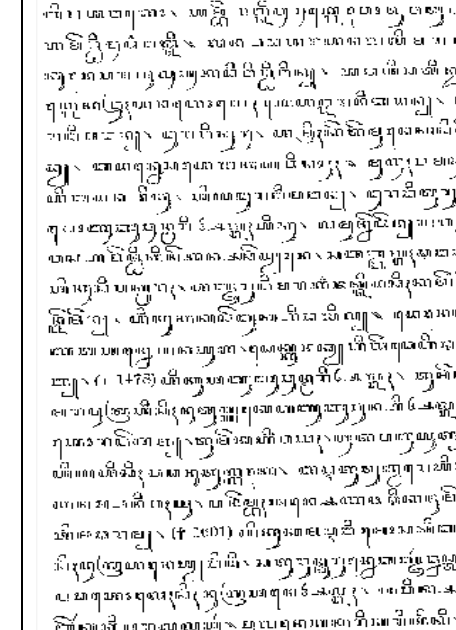
a. Implementasi hasil penelitian ini menggunakan dua ciri Intersity of Character (IoC) yaitu dengan ukuran 3x3, dan 4x4, untuk semua data yang berjumlah 733. Percobaan Klustering Agglomerative Hierarchical Clustering Complete Linkage pertama akan menggunakan data IoC 3x3, lalu percobaan Klustering Agglomerative Hierarchical Clustering Complete Linkage kedua akan menggunakan data IoC 4x4.

b. Dari ekstraksi ciri tersebut kemudian akan dikelompokkan dengan metode Klustering Agglomerative Hierarchical Clustering complete linkage dari 1 sampai dengan 733. Kemudian

hasil clustering akan dianalisis dengan melihat hasil dari dendrogram tiap percobaan dengan memfokuskan pengelompokan cluster untuk beberapa jumlah aksara Jawa yang ada di dendrogram.

3.2 Hasil Preprocessing

3.2.1 Hasil Binerisasi

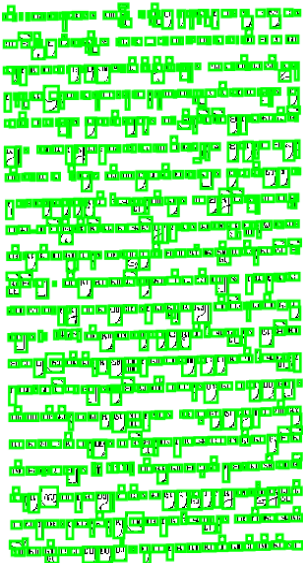
| Citra Awal | Citra Hasil Binerisasi | Citra Hasil Filtering |
|--|--|---|
| <p>  </p> | <p>  </p> | <p>  </p> |

Tabel 1 Perbandingan Citra Awal, Binerisasi, dan Filtering

Pada Tabel 1 hasil perbandingan, dapat dilihat dimana citra Binerisasi yang terdapat bayangan keabuan sudah menghilang pada citra Filtering

3.2.2 Hasil Segmentasi

a. Hasil Pengambilan Karakter dan Pengukuran untuk Setiap Karakter





Gambar 5 Hasil Pengambilan Karakter dan Pengukuran untuk Setiap Karakter

b. Hasil Menyimpan Setiap Karakter



Gambar 6 Hasil Menyimpan Setiap Karakter

c. Hasil Resize

| Hasil Segmentasi Ukuran 33x33 | Hasil Segmentasi Ukuran 44x44 |
|---|---|
|  |  |

Tabel 2 Perbandingan Hasil Resize

Terlihat perubahan ukuran aksara secara drastic di table 2 dimana awalnya ukuran gambar aksara sama

3.3 Implementasi Ekstraksi Ciri

| Data | Ciri 1 | Ciri 2 | Ciri 3 | Ciri 4 | Ciri 5 | Ciri 6 | Ciri 7 | Ciri 8 | Ciri 9 |
|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Aksara 1 | 28 | 14 | 0 | 31 | 39 | 40 | 1 | 0 | 13 |
| Aksara 2 | 27 | 15 | 17 | 29 | 0 | 17 | 13 | 22 | 0 |
| Aksara 3 | 8 | 0 | 8 | 19 | 17 | 9 | 25 | 32 | 33 |
| Aksara 4 | 17 | 22 | 30 | 24 | 25 | 28 | 29 | 23 | 28 |
| Aksara 5 | 20 | 13 | 5 | 24 | 20 | 10 | 0 | 0 | 9 |

Tabel 3 Hasil 5 Aksara Awal Matrikx IoC 3x3

Pada Tabel 3 disini data aksara yang diambil adalah 5 aksara dari 733 aksara yang dimana aksara tersebut memiliki 9 kolom ciri dimana karena awalnya Ioc berbentuk 3 baris x 3 kolom akan diubah menjadi 1 baris dengan 9 kolom untuk mempermudah proses clustering.

| Data | Ciri 1 | Ciri 2 | Ciri 3 | Ciri 4 | Ciri 5 | Ciri 6 | Ciri 7 | Ciri 8 | Ciri 9 | Ciri 10 | Ciri 11 | Ciri 12 | Ciri 13 | Ciri 14 | Ciri 15 | Ciri 16 |
|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|---------|---------|---------|---------|---------|---------|
| Aksara 1 | 26 | 36 | 4 | 0 | 19 | 10 | 22 | 13 | 38 | 54 | 20 | 56 | 1 | 0 | 0 | 16 |
| Aksara 2 | 26 | 14 | 19 | 4 | 34 | 0 | 0 | 23 | 37 | 2 | 0 | 5 | 2 | 36 | 6 | 0 |
| Aksara 3 | 10 | 0 | 0 | 6 | 24 | 19 | 5 | 26 | 22 | 9 | 4 | 0 | 20 | 36 | 37 | 35 |
| Aksara 4 | 36 | 30 | 29 | 31 | 13 | 21 | 24 | 32 | 31 | 28 | 28 | 21 | 33 | 21 | 26 | 32 |
| Aksara 5 | 21 | 15 | 14 | 1 | 19 | 7 | 5 | 1 | 17 | 18 | 15 | 5 | 0 | 0 | 0 | 17 |

Tabel 4 Hasil 5 Aksara Awal Matrikx IoC 4x4

Pada Tabel 4 disini data aksara yang diambil adalah 5 aksara dari 733 aksara yang dimana aksara tersebut memiliki 16 kolom ciri dimana karena awalnya Ioc berbentuk 4 baris x 4 kolom akan diubah menjadi 1 baris dengan 9 kolom untuk mempermudah proses clustering.

3.4 Hasil Euclidean Distance

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|----|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1 | 0 | 56.0535 | 65.1844 | 54.8817 | 37.6431 | 54.1849 | 68.2642 | 67.1789 | 69.5126 | 57.3411 | 53.9815 | 70.5408 | 39.7618 |
| 2 | 56.0535 | 0 | 49.5278 | 46.1519 | 37.4833 | 49.8999 | 38.2884 | 33.8969 | 38.0789 | 15.6205 | 41.4970 | 32 | 46.2277 |
| 3 | 65.1844 | 49.5278 | 0 | 40.2616 | 50.8134 | 47.7598 | 37.3497 | 25.0998 | 42.9767 | 50.5272 | 57.0351 | 37.7492 | 49.5177 |
| 4 | 54.8817 | 46.1519 | 40.2616 | 0 | 52.8678 | 57.2713 | 56.5332 | 40.3609 | 28.6007 | 54.6260 | 42.5676 | 31.7175 | 47.7179 |
| 5 | 37.6431 | 37.4833 | 50.8134 | 52.8678 | 0 | 29.7153 | 44.1701 | 48.9081 | 60.1249 | 29.8161 | 44.3734 | 53.8052 | 19.1833 |
| 6 | 54.1849 | 49.8999 | 47.7598 | 57.2713 | 29.7153 | 0 | 45.2548 | 51.6043 | 63.7652 | 44.3396 | 52.7257 | 57.3411 | 32.2955 |
| 7 | 68.2642 | 38.2884 | 37.3497 | 56.5332 | 44.1701 | 45.2548 | 0 | 28.0891 | 54.2586 | 35.4401 | 48.7647 | 45.6289 | 50.3488 |
| 8 | 67.1789 | 33.8969 | 25.0998 | 40.3609 | 48.9081 | 51.6043 | 28.0891 | 0 | 32.1714 | 37.4566 | 48.0937 | 24.8797 | 52.5167 |
| 9 | 69.5126 | 38.0789 | 42.9767 | 28.6007 | 60.1249 | 63.7652 | 54.2586 | 32.1714 | 0 | 48.7032 | 44.8553 | 15.6205 | 58.3009 |
| 10 | 57.3411 | 15.6205 | 50.5272 | 54.6260 | 29.8161 | 44.3396 | 35.4401 | 37.4566 | 48.7032 | 0 | 45.9565 | 39.8246 | 41.1947 |
| 11 | 53.9815 | 41.4970 | 57.0351 | 42.5676 | 44.3734 | 52.7257 | 48.7647 | 48.0937 | 44.8553 | 45.9565 | 0 | 46.4758 | 38.9230 |
| 12 | 70.5408 | 32 | 37.7492 | 31.7175 | 53.8052 | 57.3411 | 45.6289 | 24.8797 | 15.6205 | 39.8246 | 46.4758 | 0 | 53.9351 |
| 13 | 39.7618 | 46.2277 | 49.5177 | 47.7179 | 19.1833 | 32.2955 | 50.3488 | 52.5167 | 58.3009 | 41.1947 | 38.9230 | 53.9351 | 0 |
| 14 | 68.4763 | 29.5466 | 40.2741 | 41.6533 | 43.1509 | 44.4185 | 37.5366 | 31.8748 | 34.2491 | 30.4795 | 41.4126 | 23.0434 | 42.5441 |
| 15 | 73.6885 | 54.2033 | 38.9487 | 36.5787 | 63.8984 | 57.7927 | 54.4243 | 45.2659 | 36.4417 | 60.5145 | 49.8598 | 37.8153 | 55.9196 |
| 16 | 52.3737 | 36.9188 | 40.4969 | 17.8045 | 42.8486 | 52.5833 | 49.4267 | 36.5513 | 31 | 43.0929 | 37.7492 | 27.6586 | 39.5221 |
| 17 | 79.2654 | 55.7943 | 32.2490 | 46.5940 | 63.9687 | 51.1175 | 46.4866 | 32.5883 | 40.5339 | 59.7244 | 57.3672 | 39.1024 | 62.5140 |

Gambar 7 Hasil Matriks IoC 3x3 Dari Euclidean Distance

Terlihat bentuk matrix 733x733 yang diambil dari data ciri Matriks IoC 3x3, kemudian matrix 733x733 tersebut akan dihitung euclidean distance nya untuk mempermudah clustering.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|----|----------|---------|----------|---------|---------|---------|----------|----------|----------|---------|---------|----------|----------|
| 1 | 0 | 94.4934 | 105.9623 | 81.5782 | 73.3076 | 89.3364 | 102.7132 | 110.5351 | 110.4355 | 89.8833 | 85.4225 | 104.8904 | 69.0290 |
| 2 | 94.4934 | 0 | 64.3195 | 84.1784 | 57.8705 | 72.2772 | 58.2838 | 60.7865 | 68.8113 | 26.5707 | 83.1023 | 57.9224 | 74.5386 |
| 3 | 105.9623 | 64.3195 | 0 | 74.4782 | 70.7248 | 58.6600 | 54.8270 | 38.5487 | 59.0762 | 61.7495 | 74.8665 | 49.4975 | 82.7103 |
| 4 | 81.5782 | 84.1784 | 74.4782 | 0 | 79.2528 | 80.2371 | 91.6024 | 74.9200 | 67.0895 | 87.0976 | 79.0190 | 66.5808 | 68.0441 |
| 5 | 73.3076 | 57.8705 | 70.7248 | 79.2528 | 0 | 41.8210 | 59.3970 | 75.5513 | 89.5768 | 39.8873 | 77.5306 | 78.8289 | 35.4260 |
| 6 | 89.3364 | 72.2772 | 58.6600 | 80.2371 | 41.8210 | 0 | 57.1752 | 74.3572 | 86.6891 | 55.1725 | 85.0294 | 76.7919 | 52.4404 |
| 7 | 102.7132 | 58.2838 | 54.8270 | 91.6024 | 59.3970 | 57.1752 | 0 | 50.1199 | 80.5978 | 45.7275 | 79.8937 | 68.0735 | 79.7935 |
| 8 | 110.5351 | 60.7865 | 38.5487 | 74.9200 | 75.5513 | 74.3572 | 50.1199 | 0 | 53.7401 | 61.0492 | 78.1985 | 40.0250 | 90.7138 |
| 9 | 110.4355 | 68.8113 | 59.0762 | 67.0895 | 89.5768 | 86.6891 | 80.5978 | 53.7401 | 0 | 75.1199 | 93.3113 | 24.3721 | 91.2853 |
| 10 | 89.8833 | 26.5707 | 61.7495 | 87.0976 | 39.8873 | 55.1725 | 45.7275 | 61.0492 | 75.1199 | 0 | 81.2158 | 63.2218 | 64.5136 |
| 11 | 85.4225 | 83.1023 | 74.8665 | 79.0190 | 77.5306 | 85.0294 | 79.8937 | 78.1985 | 93.3113 | 81.2158 | 0 | 87.4128 | 81.2034 |
| 12 | 104.8904 | 57.9224 | 49.4975 | 66.5808 | 78.8289 | 76.7919 | 68.0735 | 40.0250 | 24.3721 | 63.2218 | 87.4128 | 0 | 85.3639 |
| 13 | 69.0290 | 74.5386 | 82.7103 | 68.0441 | 35.4260 | 52.4404 | 79.7935 | 90.7138 | 91.2853 | 64.5136 | 81.2034 | 85.3639 | 0 |
| 14 | 94.3133 | 58.2237 | 65.0308 | 67.4389 | 56.7010 | 60.2495 | 63.3798 | 57.5239 | 58.7282 | 55.6417 | 92.4878 | 45.4863 | 60.7289 |
| 15 | 97.5654 | 80.8826 | 68.1542 | 49.3761 | 74.0473 | 67.6461 | 72.0486 | 67.3869 | 58.7452 | 79.3473 | 88.5664 | 57.0351 | 67.8380 |
| 16 | 77.8010 | 64.1872 | 69.8785 | 38.8330 | 63.8201 | 70.2424 | 70.7602 | 65.7951 | 58.8473 | 63.2930 | 74.3909 | 54.7996 | 60.1831 |
| 17 | 123.9798 | 91.9456 | 54.5619 | 81.3757 | 93.3435 | 72.4983 | 67.5500 | 54.2863 | 64.1015 | 87.3499 | 93.7017 | 61.2944 | 99.5691 |
| 18 | 110.2021 | 76.7462 | 44.2500 | 72.2076 | 80.0207 | 84.3724 | 77.0000 | 24.1211 | 40.5601 | 70.6610 | 86.7516 | 42.0465 | 100.1400 |

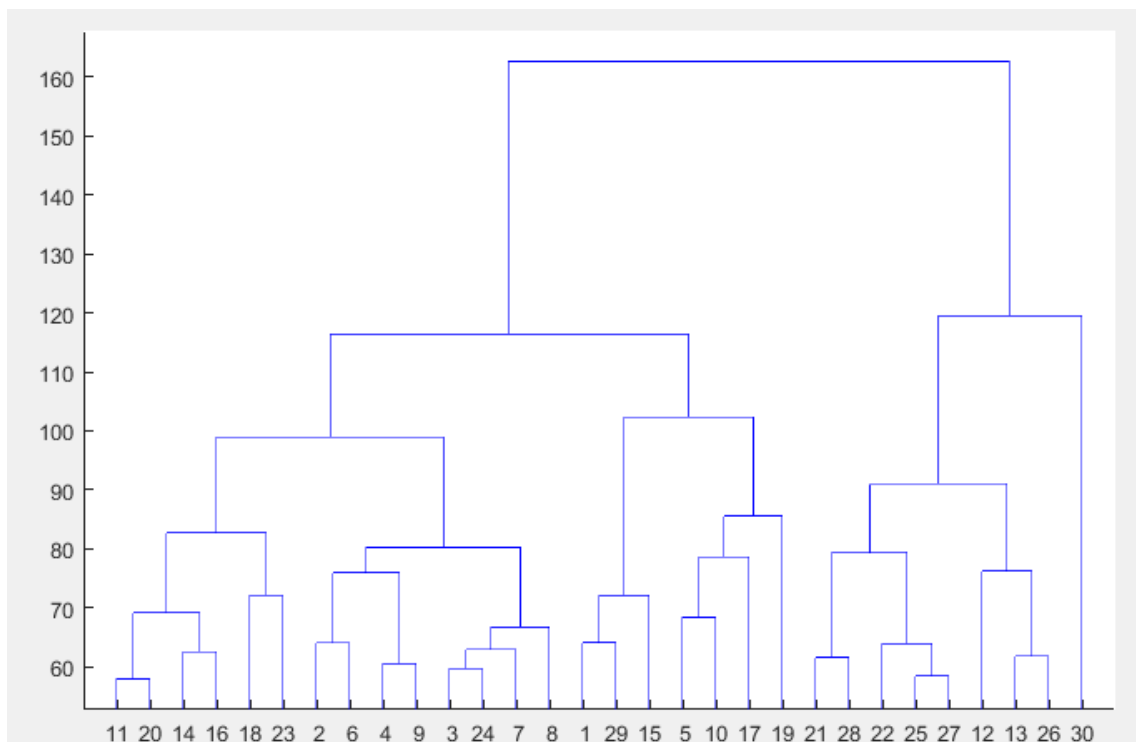
Gambar 8 Hasil Matriks IoC 3x3 Dari Euclidean Distance

Terlihat bentuk matrix 733x733 yang diambil dari data ciri Matriks IoC 4x4, kemudian matrix 733x733 tersebut akan dihitung euclidean distance nya untuk mempermudah clustering.

3.5 Hasil dari Agglomerative Hierarchical Clustering Complete Linkage



Terlihat diatas merupakan hasil *clustering* 733 citra aksara Jawa menggunakan *Agglomerative Hierarchical Clustering Complete Linkage*. Pada dendrogram terlihat bahwa data yang ditunjukkan hanya 1 sampai 30 saja, yang dikarenakan *function cluster* yang ada di MATLAB hanya menampilkan data tidak lebih dari 30 agar menghindari terjadinya penumpukan object pada dendrogram.

3.5.1 Hasil AHC Complete Linkage 3X3



Gambar 8 Hasil Dendrogram untuk Data Ciri Matriks IoC 3x3

Dendrogram diatas merepresentasikan data-data hasil ekstraksi ciri, yang mana mereka akan dijadikan sepasang jika jarak antar data dengan data yang lain memiliki nilai minimum. Sebagai contohnya, pada Dendrogram terlihat data nomor 11 dipasangkan dengan data nomor 20. Hal ini dapat terlihat pada tabel diatas bahwa data yang bernomor 11 dianggap mirip dengan data yang bernomor 20.

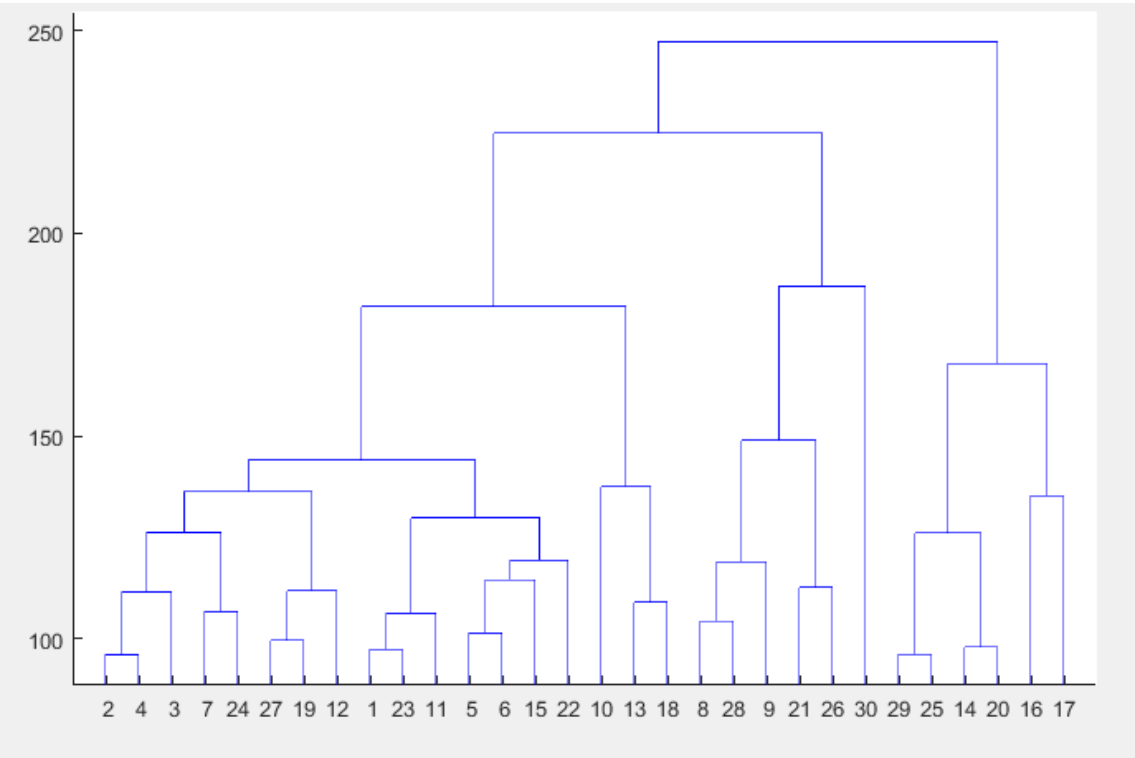
| Gambar Aksara 11 | Gambar Aksara 20 |
|---|---|
|  |  |

Pada gambar diatas terlihat sangat berbeda dengan hasil representasi dendrogram yang ada pada Pada bagian ini, Dendrogram memasang dua aksara yang memiliki bentuk berbeda, hal ini bisa terjadi dikarenakan saat proses *hierarchival*, maka citra nomor 11 dan nomor 20 dianggap memiliki jarak terpendek begitu pun dengan aksara yang di pasangkan lainnya



Sehingga hasil dari dendrogram dari 3x3 tersebut didapat jaraknya yaitu 162.65

3.5.2 Hasil AHC Complete Linkage 4X4

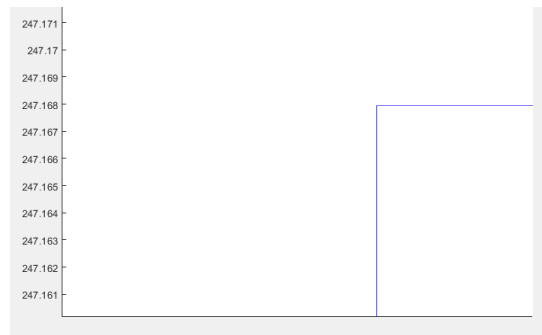


Gambar 8 Hasil Dendrogram untuk Data Ciri Matriks IoC 4x4

Dendrogram diatas merepresentasikan data-data hasil ekstraksi ciri, yang mana mereka akan dijadikan sepasang jika jarak antar data dengan data yang lain memiliki nilai minimum. Sebagai contohnya, pada Dendrogram terlihat data nomor 5 dipasangkan dengan data nomor 6. Hal ini dapat terlihat pada tabel diatas bahwa data yang bernomor 5 dianggap mirip dengan data yang bernomor 6.

| Gambar Aksara 5 | Gambar Aksara 6 |
|-----------------|-----------------|
| | |

Pada gambar diatas terlihat sangat berbeda dengan hasil representasi dendrogram yang ada pada. Pada bagian ini, Dendrogram memasangkan dua aksara yang memiliki bentuk berbeda, hal ini bisa terjadi dikarenakan saat proses *hierarchical*, maka citra nomor 11 dan nomor 20 dianggap memiliki jarak terpendek begitu pun dengan aksara yang di pasangkan lainnya



Sehingga hasil dari dendrogram dari 4x4 tersebut didapat jaraknya yaitu 247.168

4. KESIMPULAN

Berdasarkan penelitian tentang Pengelompokan Citra Aksara Jawa Menggunakan *Hierarchical Clustering* yang telah dilakukan, maka didapatkan kesimpulan sebagai berikut :

1. Dari hasil percobaan terdapat 733 pengelompokan citra aksar Jawa
2. Pada dendrogram Data Ciri Matriks IoC 3x3 menggunakan clustering Agglomerative Hierarchical Clustering jenis complete Linkage untuk dimana dendrogram tersebut menampilkan clustering 30 pada dendrogram dengan jarak 162.65
3. Ada juga kelebihan dalam metode ini adalah pemakain matriks IoC 3x3 yang lebih mudah ditemukan ciri karakter aksara nya sehingga mempercepat dalam pengelompokan cluster dengan Agglomerative Hierarchical Clustering complete Linkage. Sedangkan kekurangan yaitu dengan pemakain matriks IoC 3x3 membuat hasil dari pengelompokan cluster dengan Agglomerative Hierarchical Clustering complete Linkage lebih akurat karena semakin kecil nilai jarak eulidean maka semakin mirip.
4. Pada dendrogram Data Ciri Matriks IoC 4x4 menggunakan clustering Agglomerative Hierarchical Clustering jenis complete Linkage untuk dimana dendrogram tersebut menampilkan clustering 30 pada dendrogram dengan jarak 247.168
5. Ada juga serta kelebihan yaitu dengan pemakain matriks IoC 4x4 membuat hasil dari pengelompokan cluster dengan Agglomerative Hierarchical Clustering complete Linkage agak kurang akurat karena semakin kecil nilai jarak eulidean maka semakin mirip. Sedangkan kekurangan dalam metode ini adalah pemakain matriks IoC 4x4 yang lebih

susah ditemukan ciri karakter aksara nya sehingga memperlambat dalam pengelompokan cluster dengan Agglomerative Hierarchical Clustering Complete Linkage

5. SARAN

- a. Perlu adanya GUI dalam pembuatan dendogram agar cluster terlihat lebih baik
- b. Perlu dilakukan percobaan dengan ukuran ciri IoC yang lainnya untuk memperoleh hasil yang lebih baik.
- c. Perlu dilakukan proses preprocessing yang lebih baik agar hasil pengelompokan dengan dengan klastering Agglomerative Hierarchical Clustering jenis Complete Linkage menjadi lebih baik.

TESTIMONI / KESAN

Kesan / pengalaman yang menarik bagi saya selama mengikuti perkuliahan PCD di semester genap 2021/2022 ini yaitu saya banyak mempelajari banyak hal dalam pemanfaatan citra dokument untuk arsip terutama dalam citra aksara jawa seperti preprocessing citra, segmentasi citra, melakukan clustering citra aksara dari dokumen, dan melakukan classification citra aksar dari dokumen.

Dan juga saat melakukan kegiatan pembelajar diluar kampus seperti ke Diaroma Arsip Joja untuk melihat dokumen lama dan cara pengarsipanya serta berkeliling ke museum diorama di sana. Kemudian Perpustakaan Pustaka Artati kampus merican untuk melihat aksara-aksara yang tersimpan disana. Dan yang paling berkesan adalah Kelas Pengenalan Citra Dokumen itu sendiri yang dimana jumlah mahasiswanya hanya 13 orang saja yang membuat pertemanan di kelas ini menjadi lebih dekat sehingga bisa mengenal satu sama lain lebih mendalam

DAFTAR PUSTAKA

- [1] Hadiprijono (2013). Trampil Maca lan Nulis Aksara Jawa. Kanisius Yogyakarta. Holle, K. (1876). *Hamong Tani*
 - [2] Gonzalez, R.C., & Woods, R.E. (2002). Digital Image Processing. 2nd Edition, Prentice Hall, Upper Saddle River.
 - [3] Widiarti, A. R., & Himamunanto, A. R. (2012). Teori Dan Aplikasi Pengolahan Citra Digital Transliterasi Otomatis Citra Dokumen Teks Aksara Jawa. *Lintang Pustaka Utama*.
 - [4] F. H. Saad, O. I. E. Mohamed, and R. E. Al-qutaish, "C Omparison of H Ierarchical a Gglomerative a Lgorithms F or C Lustering M Edical," vol. 3, no. 3, pp. 1–15, 2012.
 - [5] S. Viriyavisuthisakul, P. Sanguansat, P. Charnkeitkong, and C. Haruechaiyasak, "A comparison of similarity measures for online social media Thai text classification," *ECTICON 2015 - 2015 12th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol.*, pp. 0–5, 2015, doi: 10.1109/ECTICon.2015.7207106.
-

LAMPIRAN

1. Segmentasi Program

```
% Melakukan Segmentasi gambar aksara
clear all;clc;close all;

%% Membaca Gambar
[fname path] = uigetfile('*.','Masukan Gambar');
fname = strcat(path,fname);
img = imread(fname);

%% Menampilkan gambar asli
figure(1),imshow(img),title('Gambar Asli');

%% Mengubah gambar keskala abu-abu
if size(img,3) == 3
    img = rgb2gray(img);
end

%% Mengkonversikan gambar kecitra biner
thresholding = graythresh(img);
img = ~im2bw(img,thresholding);

%% Menghapus semua objek yang mengandung kurang dari 50 piksel pada Gambar
img = bwareaopen(img,50);
pause(1)

%% Menampilkan Gambar Setelah BoundingBox Tanpa Noise
%% Melabeli objek yang terhubung
[T Na] = bwlabel(img);

%% Mengukur objek pada gambar
properti = regionprops(T,'BoundingBox');
hold on

%% Membuat BoundingBox objek gambar
for x = 1:size(properti,1)
    rectangle('Position',properti(x).BoundingBox,'EdgeColor','g','LineWidth',2)
end
hold off
pause(1)

%% Menampilkan dan Menyimpan setiap perpotongan objek pada gambar
figure
for x = 1:Na
    [m,n] = find(T == x);
    x1 = img(min(m):max(m),min(n):max(n));
    imshow(~x1)
    pause(0)
    namefile = char(strcat('D:\citra uas\3x3\', 'Aksara-',num2str(x),'.jpeg'));
    imwrite(~x1,namefile);
end
```


2. Ekstraksi Ciri

```
% Melakukan ekstraksi ciri gambar aksara
clear all;clc;close all;

%% Membaca Semua Baris Citra
imagefiles = dir('*.jpeg');
nfiles = length(imagefiles);% Jumlah file yang ditemukan
for ii=1:nfiles
    currentfilename = imagefiles(ii).name;
    currentimage = imread(currentfilename);
    images{ii} = currentimage;
end

%% Resize Citra
resize = autoresize(images)

%% Intensity of Character
Intensity_Of_Character = autoIntensity_Of_Character(resize)
Intensity_Of_Character = Intensity_Of_Character';
Intensity_Of_Character = cell2mat(Intensity_Of_Character);

%% Menyimpan data dari ekstraksi ciri
save ('DataEkstraksiCiri.mat','Intensity_Of_Character');
```

- Ekstraksi ciri 3 unit

```
function fitur = ekstrakciri_3unit(kar)
data=kar;
unit=3;
ukuran=size(data,1);
uk_unit=ukuran/unit;
hasil=zeros(unit,unit);
%ciri baris atas
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=1:uk_unit
        for z=1:uk_unit
            if (kar(y,indek+z)==0)
                sum=sum+1;
            end
        end
    end
    hasil(1,x)=sum;
end
%ciri baris 2
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=uk_unit+1:2*uk_unit
        for z=1:uk_unit
```

```

        for z=1:uk_unit
            if (kar(y,indek+z)==0)
                sum=sum+1;
            end
        end
        end
        end
        hasil(2,x)=sum;
    end
    %ciri baris 3
    for x=1:unit
        sum=0;
        indek=(x-1)*uk_unit;
        for y=2*uk_unit+1:3*uk_unit
            for z=1:uk_unit
                if (kar(y,indek+z)==0)
                    sum=sum+1;
                end
            end
        end
        hasil(3,x)=sum;
    end
    fitur=hasil;
end

```

- Ekstrasi ciri 4 unit

```

function fitur = ekstrakciri_4unit(kar)
data=kar;
unit=4;
ukuran=size(data,1);
uk_unit=ukuran/unit;
hasil=zeros(unit,unit);
%ciri baris atas
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=1:uk_unit
        for z=1:uk_unit
            if (kar(y,indek+z)==0)
                sum=sum+1;
            end
        end
    end
    hasil(1,x)=sum;
end
%ciri baris 2
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=uk_unit+1:2*uk_unit
        for z=1:uk_unit

```

```

    for z=1:uk_unit
        if (kar(y,indek+z)==0)
            sum=sum+1;
        end
    end
    end
    hasil(2,x)=sum;
end
%ciri baris 3
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=2*uk_unit+1:3*uk_unit
        for z=1:uk_unit
            if (kar(y,indek+z)==0)
                sum=sum+1;
            end
        end
    end
    hasil(3,x)=sum;
end
%baris paling bawah
for x=1:unit
    sum=0;
    indek=(x-1)*uk_unit;
    for y=3*uk_unit+1:4*uk_unit
        for z=1:uk_unit
            if (kar(y,indek+z)==0)
                sum=sum+1;
            end
        end
    end
    hasil(4,x)=sum;
end
fitur=hasil;
end

```

- AutoResize

```

function images = autoresize(images)
for k=1:length(images)
    cellContents = images{k}; % Ekstrak array dari sel ini.

    %cellContents = imresize(cellContents, [44,44]); % Ubah ukuran menjadi 44x44.
    cellContents = imresize(cellContents, [33,33]); % Ubah ukuran menjadi 33x33.
    %Pilih salah satu CELLCONTENTS
    filename = char(strcat('D:\citra uas\3x3\', 'aksara-', num2str(k), '.jpeg'));
    imwrite(cellContents,filename);
    images{k} = cellContents; % Stuff back in.
end
end

```

- autoIntensity_Of_Character

```
function Intensity_Of_Character = autoIntensity_Of_Character(Intensity_Of_Character)
for k=1:length(Intensity_Of_Character)
    cellContents = Intensity_Of_Character{k}; % Extract array from this cell.

    %cellContents = ekstrakciri_5unit(cellContents); %Intensity_Of_Character
    %cellContents = ekstrakciri_4unit(cellContents); %Intensity_Of_Character
    cellContents = ekstrakciri_3unit(cellContents); %Intensity_Of_Character

    cellContents=cellContents(:);
    cellContents=cellContents';
    Intensity_Of_Character{k} = cellContents; % Stuff back in.
end
```

3. AHC Aksara

```
% Melakukan AHC gambar aksara
clear all;clc;close all;

%% Meload data ciri yang telah sebelumnya diekstraksi
load DataEkstraksiCiri.mat

%% Membuat setiap elemen berisi vektor jarak antara sepasang objek.
distance = pdist(Intensity_Of_Character);

%% Memformat ulang vektor jarak menjadi matriks
distance_matrix = squareform(distance)

%% Mengelompokkan kedalam cluster menggunakan average
Z = linkage(Intensity_Of_Character,'complete')

%% Membuat dendrogram dengan 10 iterasi
dendrogram(Z)
```