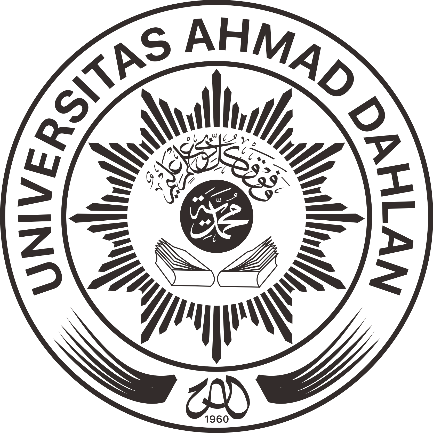
Pembimbing : Lisna Zahrotum, S.T.,M.Cs

**PENGELOMPOKAN DATA BANK SAMPAH KOTA YOGYAKARTA MENGGUNAKAN METODE *AGGLOMERATIVE HIRARCHICAL CLUSTERING (AHC)***

**SKRIPSI**

**Untuk Memenuhi Sebagian Persyaratan   
Mencapai Derajat Sarjana**

****

**Disusun Oleh:**

**Mayang Notri Silawati**

**1500018102**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INDUSTRI**

**UNIVERSITAS AHMAD DAHLAN**

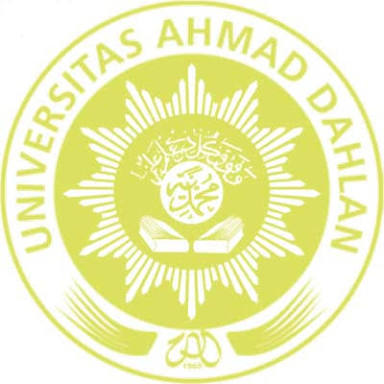
**YOGYAKARTA**

**2020**

# HALAMAN PERSETUJUAN

**SKRIPSI**

**PENGELOMPOKAN DATA BANK SAMPAH KOTA YOGYAKARTA MENGGUNAKAN METODE *AGGLOMERATIVE HIRARCHICAL CLUSTERING (AHC)***

 yang dipersiapkan untuk dipertahankan oleh :

**Mayang Notri Silawati**

**1500018102**

Kepada :

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS AHMAD DAHLAN

YOGYAKARTA

Telah disetujui oleh :

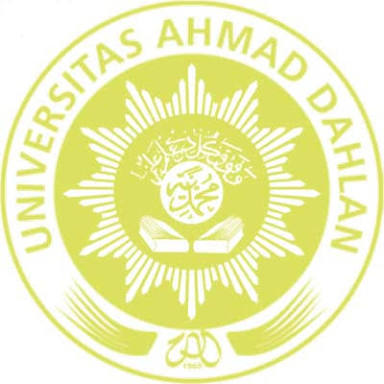
Pembimbing

**(Lisna Zahrotun, S.T., M.Cs)  
NIY. 60150773**

# HALAMAN PENGESAHAN

**SKRIPSI**

**PENGELOMPOKAN DATA BANK SAMPAH KOTA YOGYAKARTA MENGGUNAKAN METODE *AGGLOMERATIVE HIRARCHICAL CLUSTERING (AHC)***

yang dipersiapkan dan disusun oleh :

**Mayang Notri Silawati**

**1500018102**

Telah dipertahanlkan di depan Dewan Penguji

Pada tanggal .........................

dan dinyatakan telah memenuhi syarat

Susunan Dewan Penguji :

PEMBIMBING : Lisna Zahrotun, S.T., M.Cs.

PENGUJI I : Sri Handayaningsih, S.T., M.T.

PENGUJI II :

Yogyakarta, .................................. 2020

Dekan

Fakultas Teknologi Industri

Universitas ahmad Dalan Yogyakarta

**Sunardi, S.T., M.T., Ph.D.  
NIY. 60010312**

**MOTTO**

*“Jika kamu ingin hidup bahagia, terikatlah pada tujuan, bukan orang atau benda.” – Albert Einstein*

"*Sesungguhnya perbuatan baik itu dapat menghapus perbuatan buruk.*" *– Surat Hud ayat 114*

*“Ilmu pengetahuan adalah keindahan bagi masing-masing ahlinya (bidang), di dunia maupun di akhirat.” – HR. Ar-Rabii’*

# HALAMAN PERSEMBAHAN

*Al-hamdu lillahi rabbil 'alamin*

Puji Syukur saya ucapkan kepada Allah SWT. Karena berkat dan karunia-Nya sehinga penulis bisa menyelesaikan karya tulis ini yang sipersembahkan untuk :

1. Kedua orang tua saya, Ayahanda yang tercinta dan terkasih Azhari Marhaena wijaya dan Ibuda tercinta Susmiwati, terimakasih banyak atas doa, semangat, nasehat, kasih sayang, motivasi, dan pelajaran hidup yang sangat berkesan ini.
2. Untuk kakak laki-laki tercinta Bayu Sagita Wijaya dan kakak perempuan terkasih Hemi Puspitawati, terimakasih atas doa dan semangat yang telah diberikan selama ini sehingga penulis dapat menyelesaikan skripsi ini.
3. Untuk orang terkasih Luthfi Ryanto, terimakasih atas doa, semangat, motivasi selama ini dan terimakasih telah menemani saya dalam suka dan duka, terimakasih atas dukungannnya selama proses sampai pembuatan laporan skripsi ini selesai.
4. Ibu Lisa Zahrotun, S.T., M.Cs. sebagai dosen pembimbing skripsi, terimakasih atas bimbingannya ibu kepada saya selama ini.
5. Untuk sahabat tercinta terimakasih kalian selalu menjadi penyemangatku. Sahabat seperjuangan dari awal perkuliah hingga saat ini yang selalu ada suka maupun duka terimakasih karna kalian selalu melengkapi hari-hariku.
6. Dan semua pihak yang tidak bisa disebutkan satu persatu, terimakasih atas semangat yang selama ini telah diberikan kepada saya.

**PENGELOMPOKAN DATA BANK SAMPAH KOTA YOGYAKARTA MENGGUNAKAN METODE *AGGLOMERATIVE HIRARCHICAL CLUSTERING (AHC)***

ABSTRAK

**Abstrak**

Sampah masih menjadi suatu permasalahan yang kerap didengar, serta banyak menimbulkan masalah serius untuk wilayah Indonesia terutama di Kota Yogyakarta, bahkan setiap harinya timbunan sampah yang dihasilkan Kota Yogyakarta bisa mencapai ratusan ton perhatinya. Pemerintah mulai kesulitan dalam menanggulangi sampah yang semakin meingkat setiap harinya, salah satu cara untuk meningkatkan kepedulian masyarakat terhadap sampah yaitu memiliki solusi yang sudah berjalan. Salah satunya dengan dibentuknya bank sampah. Bank sampah adalah salah satu tempat yang digunakan untuk mengumpulkan sampah kering yang sudah dipilah serta memiliki menejemen seperti perbankan, akan tetapi yang ditabung bukan uang melainkan sampah. Data bank sampah masih didata dengan manual menggunakan *Microsoft Excel.* Data bank sampah juga belum dilakukan pengolahan yang dapat mempermudah peneliti bank sampah untuk mengetahui bank sampah yang baik dan bank sampah yang tidak baik berdasarkan konstruksi bangunan, menejemen bak sampah, mekanisme bank sampah, dan pelaksanaan bank sampah.

Berdasarkan permasalahan tersebut, maka diperlukan pengelompokan untuk mengetahui bank sampah yang baik berdasarkan kontruksi bangunan, menejemen bank sampah,mekanisme bank sampah, dan pelaksanaan bank sampah menggunakan metode *Agglomerative Hierarchical Clustering (AHC).* Tahapan data *mining* untuk penelitian ini adalah *selection data, transformasi, cleaning data*, *euclidean distance*, metode *Agglomerative Hierarchical Clustering* dan pengujiannya menggunakan *Shilhoutte Coeficent.*

Dalam penelitian ini data yang digunakan adalah data Bank Sampah Kota Yogyakarta. Dari 126 data yang didapatkan, terbentuk 4 *cluster* yang menghasilkan akurasi sebesar 0,343 yang artinya nilai *Shilhoutte Coeficent* tergolong rendah.

*Keyword:* *Agglomerative Hierarchical Clustering,* Data *Mining, Euclidean Distance, Shilhoutte Coeficent,* *Pengelompokan.*

# KATA PENGANTAR

****

***Assalamu’alaikum Wr.Wb.***

Segala puji bagi Alloh S.W.T yang telah memberikan rahmat dan hidayahnya kepada kita sehingga Laporan Tugas Proyek **“PENGELOMPOKAN DATA BANK SAMPAH KOTA YOGYAKARTA MENGGUNAKAN METODE *AGGLOMERATIVE HIRARCHICAL CLUSTERING (AHC)* ”** dapat disusun dan diselesaikan sesuai dengan waktunya, dan semoga rahmat tersebut dapat menjadikan kita lebih beriman dan bertaqwa kepada Allah S.W.T

Sholawat serta salam tetap tercurahkan kepada junjungan besar kita Nabi Muhammad S.A.W, keluarga, dan para sahabatnya, dan semoga syafaatnya sampai kepada kita sebagai umatnya diakherat kelak.

Dalam pembuatan Laporan skripsi ini, Penyusun tidak dapat menyelesaikan Laporan ini tanpa bantuan dan dorongan dari berbagai pihak yang bersangkutan, oleh karena itu tim penyusun mengucapkan terimakasih kepada :

1. Lisna Zahrotun, S.T.,M.Cs. Selaku Dosen Pembimbing Skripsi Prodi Teknik Informatika Universitas Ahmad Dahlan Yogyakarta.
2. Fiftin Noviyanto, S.T.,M.Cs Selaku Dosen Pembimbing Akademik Prodi Teknik Informatika Universitas Ahmad Dahlan Yogyakarta.
3. Nur Rocmah D.P.A, S.T.,M.Kom Selaku Dosen Pembimbing Akademik (Sambung) Prodi Teknik Informatika Universitas Ahmad Dahlan Yogyakarta.
4. Untuk bebeb yang selalu menemani saya dalam suka dan duka, terimakasih atas dukungannnya selama proses sampai pembuatan laporan skripsi ini selesai.
5. Untuk Teman-teman tercinta saya yang sudah memberikan motivasi dan mendorong saya untuk menyelesaikan skripsi ini.
6. Untuk kakak Laki-Laki saya Bayu Sagita Wijaya dan kakak Perempuan saya Hemi Puspita Wati terimagaji telah selalu menanyakan skripsi saya, agar saya tidak malas menyelesaikan skripsi ini.

Penyusun menyadari Laporan skripsi ini tidak sempurna, karena kesempurnaan hanyalah milik tuhan semata, tetapi penulis sangat berharap untuk kebergunaan Laporan Skripsi ini bagi generasi atau para mahasiswa kreatif lainnya untuk lebih mengembangkan dan memajukan teknologi bersama untuk bisa dipersaingkan di era globalisasi ini.

Dan Penyusun Mengharapkan kritik dan saran yang bersifat membangun untuk kreatifitas bersama terutama untuk memajukan Universitas Ahmad Dahlan Yogyakarta.

07 Oktober 2018

Penyusun

Mayang Notri Silawati

# DAFTAR ISI

[HALAMAN AWAL **Error! Bookmark not defined.**](#_Toc33120054)

[HALAMAN PERSETUJUAN ii](#_Toc33120055)

[HALAMAN PENGESAHAN iii](#_Toc33120056)

[HALAMAN PERSEMBAHAN iv](#_Toc33120057)

[ABSTRAK v](#_Toc33120058)

[KATA PENGANTAR vi](#_Toc33120059)

[DAFTAR ISI viii](#_Toc33120060)

[DAFTAR TABEL x](#_Toc33120061)

[DAFTAR GAMBAR xi](#_Toc33120062)

[DAFTAR LISTING xii](#_Toc33120063)

[BAB I PENDAHULUAN 13](#_Toc33120064)

[A. Latar Belakang 13](#_Toc33120065)

[B. Identifikasi Masalah 16](#_Toc33120066)

[C. Batasan Masalah 16](#_Toc33120067)

[D. Rumusan Masalah 17](#_Toc33120068)

[E. Tujuan Penelitian 17](#_Toc33120069)

[F. Manfaat Penelitian 17](#_Toc33120070)

[BAB II KAJIAN PUSTAKA 18](#_Toc33120071)

[A. Penelitian Yang Berhubungan 18](#_Toc33120072)

[B. Landasan Teori 26](#_Toc33120073)

[BAB III METODELOGI PENELITIAN 31](#_Toc33120074)

[A. Objek Penelitian 31](#_Toc33120075)

[B. Metode Pengumpulan data 31](#_Toc33120076)

[C. Alat Penelitian 32](#_Toc33120077)

[BAB IV HASIL DAN PEMBAHASAN 38](#_Toc33120078)

[A. Hasil Pengumpulan Data 38](#_Toc33120079)

[B. Tahapan Data Mining 39](#_Toc33120080)

[C. Implementasi 48](#_Toc33120081)

[BAB V KESIMPULAN DAN SARAN 66](#_Toc33120082)

[A. Kesimpulan 66](#_Toc33120083)

[B. Saran 66](#_Toc33120084)

[DAFTAR PUSTAKA 68](#_Toc33120085)

# DAFTAR TABEL

[Tabel 2. 1 Penelitian Terdahulu 21](#_Toc32783415)

[Tabel 4. 1 Load Data 39](#_Toc33119888)

[Tabel 4. 2 Selection Data 41](#_Toc33119889)

[Tabel 4. 3 Cleaning Data 41](#_Toc33119890)

[Tabel 4. 4 Transformasi Data 42](#_Toc33119891)

[Tabel 4. 5 Euclidean Distance 43](#_Toc33119892)

[Tabel 4. 6 Tahap 1 *Complete Linkage* 44](#_Toc33119893)

[Tabel 4. 7 Tahap 2 *Complete Linkage* 45](#_Toc33119894)

[Tabel 4. 8 Tahap 3 *Complete Linkage* 45](#_Toc33119895)

[Tabel 4. 9 Tahap 4 *Complete Linkage* 46](#_Toc33119896)

[Tabel 4. 10 Tahap 5 *Complete Linkage* 46](#_Toc33119897)

[Tabel 4. 11 Tahap 6 *Complete Linkage* 46](#_Toc33119898)

[Tabel 4. 12 Tahap 7 *Complete Linkage* 47](#_Toc33119899)

[Tabel 4. 13 Tahap 8 *Complete Linkage* 47](#_Toc33119900)

[Tabel 4. 14 Hasil Silhoutte Coefficient 56](#_Toc33119901)

# DAFTAR GAMBAR

[Gambar 3. 1 Tahapan Penelitian Data Mining 33](#_Toc32783546)

[Gambar 4. 1 Interface Dashboard 49](#_Toc33119902)

[Gambar 4. 2 Dataset 51](#_Toc33119903)

[Gambar 4. 3 *Interface* hasil *Clustering* 53](#_Toc33119904)

[Gambar 4. 4 *Interface* hasil *Clustering* (Pengenalan ) 54](#_Toc33119905)

[Gambar 4. 5 *Interface* hasil *Clustering* (Tumbuh 1 ) 54](#_Toc33119906)

[Gambar 4. 6 *Interface* hasil *Clustering* (Tumbuh 2) 55](#_Toc33119907)

[Gambar 4. 7 *Interface* hasil *Clustering* (Berkembang) 55](#_Toc33119908)

[Gambar 4. 8 Data *Cluster* 1 59](#_Toc33119909)

[Gambar 4. 9 Data *Cluster* 2 61](#_Toc33119910)

[Gambar 4. 10 Data *Cluster* 3 63](#_Toc33119911)

[Gambar 4. 11 Data *Cluster* 4 65](#_Toc33119912)

# DAFTAR LISTING

[Listing 4. 1 Listing Dashboard 50](#_Toc33119914)

[Listing 4. 2 Dataset 52](#_Toc33119915)

[Listing 4. 3 Hasil *Clustering* 53](#_Toc33119916)

[Listing 4. 4 Hasil Clustering Pengenalan,Tumbuh 1,Tumbuh 2,Berkembang. 56](#_Toc33119917)

# BAB I PENDAHULUAN

## Latar Belakang

Sampah masih menjadi suatu permasalahan yang kerap didengar, serta banyak menimbulkan masalah serius untuk wilayah Indonesia terutama di Kota Yogyakarta, bahkan setiap harinya timbunan sampah yang dihasilkan Kota Yogyakarta bisa mencapai ratusan ton perhatinya. Tanpa adanya kepedulian masyarakarat terhadap sampah dan dianggap hal yang tidak penting serta tidak dihiraukan. Pemerintah mulai kesulitan dalam menanggulangi sampah yang semakin meingkat setiap harinya, salah satu cara meningkatkan kepedulian masyarakat terhadap sampah yaitu memiliki solusi yang sudah berjalan. Salah satunya dengan dibentuknya bank sampah. Bank sampah adalah salah satu tempat yang digunakan untuk mengumpulkan sampah kering yang sudah dipilah serta memiliki menejemen seperti perbankan, akan tetapi yang ditabung bukan uang melainkan sampah. Hasil dari pengumpulan sampah yang sudah dipilah akan disetorkan ke tempat pembuatan kerajinan dari sampah atau ke tempat pengepul sampah.

Berdasarkan Undang-Undang Nomor 18 Tahun 2008 Pasal 1 berbunyi Sampah adalah sisa kegiatan sehari-hari manusia dan proses alam yang berbentuk padat. Pengelolaan sampah yang dimaksud adalah kegiatan yang sistematis, menyeluruh, dan berkeseimbangan yang meliputi pengurangan dan penanganan sampah. Berdasaarkan sifat dan manfaat sampah dikelola menjadi beberapa golongan : 1) sampah organik adalah sampah yang terdiri dari bahan-bahan yang bisa terurai secara alami/biologis, seperti sisa makanan dan guguraan daun sehingga sampah sejenis ini bisa disebut sampah basah; 2) sampah anorganim adalah sampah yang terdiri dari bahan-bahan yang sulit terurai secara biologis. Proses penghancuran membutuhkan penanganan lebih lanjut ditempat khusus, misalnya plastik, kaleng san styrofoam. Sampah jenis ini bisa disebut sampah kering; 3) sampah bahan berbahaya dan beracun (B3) adalah limbah dari bahan-bahan berbahaya dan beracun seperti limbah rumah sakit, limbah pabrik dan lain-lain.(Indonesia, 2012)

Berdasarkan hasil wawancara, dengan Ibu Utaminingsih Liniarti S.T., M.T. selaku Sekertaris Program Studi Teknik Industri Unversitas Ahmad Dahlan dalam penelitiannya bahwa pemerintah kurang memperhatikan kegiatan bank sampah yang sudah diterapkan sehingga data bank sampah masih dikelola secara manual dan masih dibagi menjadi kelompok-kelompok tertentu. Adapun program pemerintah yang akan dicapai pada tahun 2019 adalah pemerintah kota Yogyakarta menginginkan setiap RW memiliki bank sampah. Jumlah data bank sampah di Kota Yogyakarta berdasarkan masing-masing kelurahan, selanjutnya beliau mengambil data lebih rinci menggunakan kuisioner sehingga banyak data yang dihasilkan. Dari data tersebut perlu dilakukan pengenalan terhadap masing-masing bank sampah yang ada, kemudian dikelompokkan berdasarkan pengelolanya.

Pengelolaan sampah yang diselenggarakan oleh dinas terkait hanya berfokus pada pengumpulan dan pengangkutan ke Tempat Pemrosesan Akhir (TPA). Pengolahan data masih dilakukan secara manual, tanpa melalui pengolahan tertentu. Kebanyakan TPA bermasalah terhadap lingkungan hidup, misalnya TPA tidak dilapisi oleh lapisan anti air seperti *geotextile*, tidak ada pengolahan air limbah, dan masih diizinkannya praktik  *open dumping* dan *open burning*. Sehingga pada akhirnya menyebabkan banyak permasalahan seperti pencemaran air limbah ke air tanah, bau busuk dan pencemaran udara. Untuk menanggulangi semua itu didirikan Bank Sampah yang memiliki tujuan memanfaatkan sampah yang tidak dikelola oleh masyarakat, sehinggal dalam hal ini sampah akan dipilih berdasarkan jenisnya dan disetorkan ke Bank Sampah setiap minggunya guna mendapatkan sejumlah nominal uang yang akan disimpan ke setiap rekning nasabah yang terdaftar di Bank Sampah tersebut. Selain itu pengelola Bank Sampah juga memiliki tujuan untuk menyediakan wadah informasi kepada masyarakt sekitar agar masyarakat tahu bagaimana cara memanfaatkan sampah sehingga sampah tersebut tidak dibuang ataupun dibakar untuk memusnahkannya.

Selain itu salah satu informasi yang dapat digali dari data tersebut adalah pengelompokan data bank sampah. Data bank sampah ini terdiri dari beberapa atribut dengan jumlah *record* yang banyak sehingga diperlukan suatu proses *data mining* yang dapat mengelompokkan data tersebut, yaitu *clustering*. Dengan menggunakan *clustering* diharapkan dapat memberikan prediksi pengelompokan tersebut. Salah satu metode *clustering* yang dapat digunakan untuk mengelompokkan data adalah *Agglomerative Hierarchical Clustering (AHC).*

Secara umum *Agglomerative Hierarchical Clustering (AHC)* dikenal beberapa metode untuk membentuk *cluster*, diataranya yaitu *single linkage,complete linkage*, *average linkage* dan *ward*. *Single linkage* mengelompokkan objek pengamatan berdasarkan jarak minimum, *complete linkage* mengelompokkan objek pengamatan berdasarkan jarak maksimum, *average linkage* mengelompokkan objek pengamatan berdasarkan jarak rata-rata. Sedangkan metode *ward* dalam perhitungan jaraknya didasarkan pada jumlah kuadrat galat antara dua *cluster* untuk semua variablenya.

*Agglomerative Hierarchical Clustering (AHC)* merupakan suatu pengelompokan hirarki yang bersifat *bottom up* dimana keberadaan setiap titik data dalam *cluster* ditentukan oleh *proximity* antar titik tersebut. Metode *Agglomerative Hierarchical Clustering (AHC)* yang akan digunakan dalam penelitian ini ialah *ward method* .Dengan latar belakang masalah tersebut maka pada penelitian ini diharapkan dapat memberikan solusi dari permasalahan yang terjadi pada Bank Sampah yakni membangun sistem pengelompokan data agar memudahkan pengelolaan data Bank Sampah.

1. **Identifikasi Masalah**

Adapun identifikasi masalah yang telah dijabarkan diatas sebagai berikut :

1. Belum ada identifikasian data bank sampah untuk memanfaatkan sampah yang bisa dikelola ataupun yang tidak bisa dikelola oleh bank sampah.
2. Pemerintah kurang memperhatikan kegiatan bank sampah yang sudah diterapkan.
3. Jumlah data bank sampah yang banyak tetapi belum dilakukan pengelompokan.

## Batasan Masalah

Agar penelitian tidak meluas, perlu adanya batasan masalah. Adapun batasan masalah sebagai berikut :

1. Penelitian ini terfokus pada pengelompokan data Bank Sampah di Kota Yogyakarta berdarkan kontruksi bangunan, menejemen bank sampah,mekanisme bank sampah, dan pelaksanaan bank sampah.
2. Data bank sampah dalam bentuk *file excel* dengan jumlah data 126 data bank sampah.
3. Penelitian ini menggunakan metode *Agglomerative Hierarchical Clustering (AHC)*

## Rumusan Masalah

Adapun rumusan masalah dari penelitian ini ialah bagaimana menerapkan metode *Agglomerative Hierarchical Clustering (AHC)* dalam melakukan pengelompokan data bank sampah untuk menetukan data yang dapat digunakan untuk pengelompokkan data bank sampah itu sendiri.

1. **Tujuan Penelitian**

Tujuan penelitian adalah melakukan pengelompokkan data bank sampah yang ada di Kota Yogyakarta berdasarkan kontruksi bangunan, menejemen bank sampah,mekanisme bank sampah, dan pelaksanaan bank sampah.

## Manfaat Penelitian

Manfaat yang ingin diperoleh dari penelitian ini adalah :

1. Manfaat bagi peneliti bank sampah mendapatkan rekomendasi untuk meningkatkan kualitas bank sampah dari tiap kelompok bank sampah.
2. Bagi pengelola bank sampah untuk mengetahui kelompok dari bank sampah yang dikelola.

# BAB II KAJIAN PUSTAKA

## Penelitian Yang Berhubungan

Pada bagian ini akan dijabarkan penelitian yang berhubngan dengan bahan acuan dasar dalam penelitian. Selain kajian penelitian yang berhubungan, akan dijelaskan tentang kajian teori yang mendukung penelitian ini.

Penelitian yang dilakukan oleh (Zahrotun, 2015) membahas tentang *Clustering K-Means dan Agglomerative Hierarchical Clustering (AHC)*  untuk analisis pengelompokan jumlah penumpang bus trans jogja. Dalam penelitian tersebut menghasilkan pengelompokkan menggunakan metode *Clustering K-Means dan Clustering AHC* maka dalam penelitian ini terlihat hasil dari pengelompokkan *Clustering K-Means* memiliki kemiripan yang lebih seimbang untuk setiap kelompokknya dibandingkan dengan hasil dari pengelompokkan *Clustering AHC*. Sehingga dapat disimpulkan meskipun pada proses *Clustering K-Means* harus menentukan *centroid* atau titik terlebih dahulu namun dalam kasus ini dapat menghasilkan data pengelompokkan yang lebih baik. Sedangkan untuk hasil pengelompokkan menggunakan metode *AHC* meskipun lebih efisien karena tidak menggunakan titik pusat pada awal pengelompokkan dan hanya menghitung tingkat kemiripan memiliki hasil yang kurang baik.

Penelitian yang dilakukan oleh (Pitaloka, Maulana and Kurniati, 2009). Dalam penelitian ini bertujuan untuk mengimplementasikan metode *Agglomerative Hierarchical Clustering (AHC)*  untuk pengelompokan data dalam sebuah oerangkat lunak, menerapkan metode data mining . Metode yang digunakan dalam penelitian ini mampu menghasilkan data pelanggan yang akan direpresentasikan ke dalam

bentuk hirarki klaster yang selanjutnya akan dikelomokkan ke dalam kelompok yang berbeda. Selain itu, akan duhitung juga cophrnrtic correlatin ceofficient untuk mengukur seberapa baik sebuah hirarchical clustering memenuhi kesesuaian data.

Penelitian yang dilakukan oleh (Ilmiah *et al.*, no date). Dalam penelitian tersebut terdapat masalah yakni Berdasarkan observasi dan wawancara yang sudah dilakukan dengan ibu Denti selaku Wakil Kepala Sekolah Kurikulum di SMPN 19 Bandung, seluruh siswa kelas 9 pada semester genap diwajibkan mengikuti kelas tambahan yang disebut Pemantapan dan ujian try out guna mempersiapkan diri untuk menghadapi Ujian Nasional. Namun tetap saja ada siswa yang tertinggal dalam mata pelajaran tertentu karena setiap siswa memiliki potensi yang berbeda satu sama lain. Untuk dapat dinyatakan lulus, nilai siswa harus bisa melewati batas KKM (Kriteria Ketuntasan Minimal) dalam semua mata pelajaran. Selain pemantapan, seharusnya dilakukan juga pembentukan kelompok belajar. Untuk membentuk kelompok belajar biasanya sekolah hanya membagi kelompok dengan menggunakan urutan absen. Jika pengelompokan dilakukan berdasarkan absen, besar kemungkinan siswa yang memiliki nilai tinggi di mata pelajaran akan berada dalam satu kelompok yang sama dengan siswa yang tertinggal di mata pelajaran tersebut.

Hal ini akan menyebabkan siswa yang memiliki nilai tinggi akan merasa bosan karena materi yang sudah dipahaminya diulang-ulang untuk membantu siswa yang tertinggal memahami materi tersebut. Agar siswa dapat mengejar mata pelajaran yang tertinggal, kelompok belajar harus dibentuk sesuai dengan kemiripan nilai siswa. Sehingga akan mempermudah guru untuk mengidentifikasi dimanakah kesulitan yang dialami dari setiap kelompok belajar yang terbentuk. Sehingga metode yang digunakan dalam penelitian ini metode *data mining, clustering, agglomerative hierarchical clustering, single linkage,* matriks jarak. Metode yang digunakan dalam penelitian ini mampu menghasilkan setelah dilakukan analisis,desain dan implementasi dari perancangan perangkat lunak yang dibangun dan telah dikembangkan serta sarana-sarana yang akan memberikan catatan penting dan kemungkinan perbaikan yang dilakukan untuk membangun perangkat lunak sebelumnya.

Tabel 2. 1 Penelitian Terdahulu

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Aspek | Pustaka 1 | Pustaka 2 | Pustaka 3 | Pustaka 4 | Pustaka 5 | Topik yang akan dikembangkan |
| Nama penulis | M. Agphin Ramadhan | Dhita Prima Kusuma1, Yuli Astuti2 | Lisna Zahrotun | Novialita Pitaloka¹, Kiki Maulana², Angelina Prima Kurniati³ | Andre Catur Prasetyo | Mayang Notri Silawati |
| Nama Jurnal | PERBANDINGAN EFEKTIVITAS BANK SAMPAH  DI KOTA BANDUNG DAN KOTA YOGYAKARTA | SISTEM PENGOLAHAN DATA BANK SAMPAH (STUDY KASUS :  BANK SAMPAH BANGKIT PONDOK I  NGEMPLAK SLEMAN) | ANALISIS PENGELOMPOKAN JUMLAH PENUMPANG BUS TRANS JOGJA MENGGUNAKAN METODE *CLUSTERING K-MEANS* DAN *AGGLOMERATIVE HIERARCHICAL CLUSTERING* (AHC) | PENGELOMPOKAN DATA MENGGUNAKAN HIERARCHICAL CLUSTERING  (AHC) | PENERAPAN DATA MINING UNTUK MEMBENTUK KELOMPOK  BELAJAR MENGGUNAKAN METODE CLUSTERING DI SMPN 19  BANDUNG | Pengelompokan bank sampah menggunakan metode *Agglomerative Hirarchical Clustering (AHC)* |
| Volume, Nomor | Vol. XII No.1 | Volume 21 No 1 | Vol. 9, No. 1 | - | Edisi. .. Volume. .., | - |
| Bulan, Tahun Terbit | Mei 2016 | Juni 2017 | Januari 2015 | 2009 | Bulan 20.. | - |
| Penerbit | INERSIA | ISSN 2088-3943 | JURNAL INFORMATIKA | Telkom University | ISSN : 2089-9033 | - |
| Masalah Penelitian | BSWM lebih luas menjangkau RW-RW lain di sekitarnya. Sedangkan Bank Sampah ASRI masih  terfokus pada warga di RW 08; 8) Kemudian, dalam hal kegiatan bank sampah. BSWM memiliki beberapa kegiatan selain pengumpulan, penjualan, dan pemanfaatan sampah, yaitu: penyuluhan ke warga sekitar dan sekolah mengenai pemanfaatan sampah, serta aktif mengikuti pameran lingkungan hidup. Sedangkan Bank Sampah ASRI fokus pengumpulan, penjualan, dan pemanfaatan sampah. | Adapun permasalahan yang ada pada system pengolahan data Bank Sampah yakni Pelayanan terhadap nasabah saat melakukan proses transaksi masih semi manual dibantu dengan kalkulator untuk penghitungan hasil setoran dan akhirnya bisa ditulis dibuku tabungan nasabah. Pembuatan laporan dilakukan secara manual, sehingga petugas harus satu persatu mencari data kemudian menyimpulkan serta menuliskan hasil dari informasi data tersebut. Tidak ada validasi dalam penginputan data. Sehingga saat pembuatan laporan akan menyulitkan penulisan akhir laporan. | Namun algoritma ini juga memiliki kelemahan bergantung pada pemilihan teknik *intercluster similarity* yang lebih dikenal dengan istilah *linkage*. Beberapa kelemahan dari linkage tersebut adalah sensitif terhadap adanya *outlier*, kesulitan menangani variasi bentuk dan ukuran, dan memisahkan cluster yang besar. Sedangkan Berkhin menyebutkan beberapa kelemahan algoritma K‐*means* adalah: (1) sangat bergantung pada pemilihan nilai awal centroid, (2) tidak jelas berapa banyak cluster k yang terbaik, (3) hanya bekerja pada atribut numerik. | Data yang akan digunakan sebagai studi kasus adalah data tagihan pelanggan layanan PT.Telkom. Data yang akan digunakan dalam format MS.Excel dengan tipe \*.csv. | Untuk membentuk kelompok belajar biasanya sekolah hanya membagi kelompok dengan menggunakan urutan absen. Jika pengelompokan dilakukan berdasarkan absen, besar kemungkinan siswa yang memiliki nilai tinggi di mata pelajaran akan berada dalam satu kelompok yang sama dengan siswa yang tertinggal di mata pelajaran tersebut. Hal ini akan menyebabkan siswa yang memiliki nilai tinggi akan merasa bosan karena materi yang sudah dipahaminya diulang-ulang untuk membantu siswa yang tertinggal memahami materi tersebut. | Adapun identifikasi masalah yang telah dijabarkan diatas sebagai berikut :  Kesulitan dalam pendataan data dari bank sampah yang dilakukan secara tertulis.  Belum ada identifikasian data bank sampah untuk memanfaatkan sampah yang bisa dikelola ataupun yang tidak bisa dikelola oleh bank sampah. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tujuan Penelitian | Meningkatnya jumlah bank sampah di masing masing kota tersebut diharapkan dapat pula meningkatkan partisipasi masyarakat dalam pengelolaannya sehingga efektif mengurangi volume sampah di perkotaan. | Untuk menyediakan wadah informasi untuk masyarakat sekitar secara menarik tentang bagaimana cara memanfaatan sampah agar sampah tidak dibuang ataupun dibakar untuk memusnahkannya. | Bertujuan untuk mengelompokkan data ke dalam kluster sehingga setiap kluster berisi data yang semirip mungkin | Mengimplementasikan metode  *(AHC)* untuk pengelompokan data dalam sebuah perangkat lunak. Menerapkan metode *data mining, Agglomerative Hierarchical Clustering* *(AHC)* untuk membentuk hirarki dari data . Memberikan hasil pengelompokan data menggunakan metode *(AHC)* dengan pendekatan *single linkage, complete linkage* dan *average linkage* serta analisis hasilhirarkinya dengan *cophenetic distance*. | Tujuan yang akan dicapai dalam  penelitian ini adalah membantu pihak kurikulum  untuk membentuk kelompok belajar berdasarkan  kemiripan nilai siswa. | Adapun rumusan masalah yang telah diuraikan diatas, maka penelitian ini memiliki tujuan sebagai berikut : Membantu bank sampah dalam mengelompokkan data-data yang dikelola secara manual.  Membedakan bank sampah yang telah sukses dan belum sukses dalam sisi dana,sosialisasi dan penyuluhan. Menerapkan metode *Agglomerative Hierarchical Clustering (AHC)* untuk mengelompokkan data bank sampah. |
| Metode / teori yang digunakan | observasi dan wawancara, klasifikasi | Analisis PIECES | *Clustering K-Means* dan *Agglomerative Hierarchical Clustering* (AHC)” | AHC, Single Linkage, Complete Linkage, Average Linkage | *data mining, clustering, agglomerative*  *hierarchical clustering, single linkage,* matriks jarak | *Agglomerative*  *Hierarchical Clustering* |
| Hasil Penelitian | Sumber sampah baik pada BSWM maupun  Bank Sampah ASRI berasal dari sampah rumah tangga, seperti: plastik pembungkus makanan, botol plastik, botol kaca, kaleng, dan kertas, baikkertaskoran, HVS, maupunkarton.Di Indonesia,penggolongan sampah yang sering digunakan adalah sebagai (a) sampah  organik, atau sampah basah, yang terdiri atas  daun-daunan, kayu, kertas, karton, tulang, sisa-sisa makanan ternak, sayur, buah, dan lain-lain, dan sebagai (b) sampah anorganik, atau sampah kering yang terdiri atas kaleng, plastik, besi dan logam-logam lainnya, gelas dan mika. Sampah yang diterima pada kedua bank sampah ini masuk ke klasifikasi sampah anorganik. Karena jenis sampah ini masih dapat dijual ke pengepul atau dimanfaatkan menjadi kerajinan tangan. Sedangkan sampah-sampah organik belum dapat dimanfaatkan karena terbatasnya sarana, prasarana, dan SDM pengelola bank sampah. | Penelitian ini berhasil membangun system pengolahan data bank sampah pada bank sampah bangkit pondok 1. Sistem sudah sesuai dengan proses bisnis di pengolahan bank sampah yang ada di bank sampah bangkit pondok 1. Dengan adanya website yang bisa diakses masyarakat luas pihak pengelola dapat dengan mudah menyebarkan informasi tentang bank sampah. Pihak pengelola bisa menambah informasi baru, mengubah informasi, dan menghapus informasi di dalam website. | Dari hasilpengelompokkan menggunakan metode clustering K-Means dan clustering AHC maka dalam penelitian ini terlihat hasil dari pengelompokkan clustering K-Means memiliki kemiripan yang lebih seimbang untuk setiap kelompokknya dibandingkan dengan hasil dari pengelompokkan clustering AHC. Sehingga dapat disimpulkan meskipun pada proses clustering K-Means harus menentukan centroid atau titik terlebih dahulu namun dalam kasus ini dapat menghasilkan data pengelompokkan yang labih baik. Sedangkan untuk hasil pengelompokkan menggunakan metode AHC meskipun lebih efisien karena tidak menggunakan titik pusat pada awal pengelompokkan dan hanya menghitung tingkat kemiripan memiliki hasil yang kurang baik. | Metode Agglomerative Hierarchical Clustering (AHC) dengan pendekatan jarak single linkage, complete linkage dan average linkage  dapat digunakan untuk membangun hirarki dari data dan mengelompokkannya.  2) Performansi metode (AHC) dengan pendekatan average link pada dataset Iris lebih baik bila dibandingkan dengan (AHC) dengan pendekatan single link dan complete link yaitu 90,66% berbanding 68% untuk single linkage dan 84% untuk complete link.  3) Metode pendekatan jarak (proximity) sangat berpengaruh dalam  membangun hirarki klaster karena perbedaan metode ini menyebabkan hirarki yang dibangunnya pun berbeda.  4) Berdasarkan nilai CPCC yang diperoleh, hasil hirarki metode  (AHC) dengan pendekatan average linkage lebih baik dibandingkan dengan Agglomerative  Hierarchical Clustering (AHC) dengan pendekatan single linkage dan complete linkage. | Hasil-hasil yang diperoleh  setelah dilakukan analisis, desain, dan implementasi  dari perancangan perangkat lunak yang dibangun  dan telah dikembangkan serta saran-saran yang akan  memberikan catatan penting dan kemungkinan  perbaikan yang perlu dilakukan untuk  pengembangan perangkat lunak sebelumnya.  Bahwa aplikasi yang dibangun  membantu pihak bagian kurikulum untuk  membentuk kelompok belajar yang sesuai dengan  kemiripan nilai siswa. | Dapat memberikan solusi dari permasalahan yang terjadi pada Bank Sampah yakni membangun sistem pengelompokan data agar memudahkan pengelolaan data Bank Sampah. |

## Landasan Teori

1. **Data Mining**

Data mining merupakan proses menemukan pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, *data warehouse*, atau tempat penyimpanan data lainnya. Data mining juga didefinisikan sebagai bagian dari proses penggalian pengetahuan dalam *database* yang sering disebut dengan istilah *Knowledge Discovery in Database* (KDD). KDD merupakan suatu area yang mengintegrasikan berbagai metode. Adapun beberapa tahap data mining yang merupakan suatu rangkaian proses:

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*).
2. Integrasi data (penggabungan data dari beberapa sumber).
3. Transformasi data (data diubah menjadi bentuj yang sesuai untuk di-mining).
4. Aplikasi teknik data mining.
5. Evaluasi pola yang ditemukan(untuk menemukan yang menarik/bernilai).
6. Presentasi pengetahuan (dengan teknik visualisasi).
7. **Clustering**

Analisis cluster adalah proses pengelompokan objek ke dalam subset yang memiliki makna dalam konteks masalah tertentu. Benda-benda tersebut dengan demikian disusun menjadi representasi efisien yang mencirikan populasi yang diambil sampelnya. Tidak seperti klasifikasi, pengelompokan tidak bergantung pada kelas yang telah ditentukan sebelumnya. *Clustering* disebut sebagai metode pembelajaran yang tidak diawasi karena tidak ada informasi yang diberikan tentang “jawaban yang benar” untuk objek apa pun. Ini dapat mengungkap hubungan yang sebelumnya tidak terdeteksi dalam kumpulan data yang rumit. Banyak aplikasi untuk analisis cluster ada. Misalnya, dalam aplikasi bisnis, analisis *cluster* dapat digunakan untuk menemukan dan mengkarakterisasi grup pelanggan untuk tujuan pemasaran.

Dua jenis algoritma pengelompokan adalah *non-hierarchical* dan *hierarchical*. Dalam pengelompokan *non-hierarchical*, seperti algoritma *k-means*, hubungan antara *cluster* tidak ditentukan. *Hierarchical clustering* berulang kali menghubungkan pasangan *cluster* sampai setiap objek data termasuk dalam hirarki. Dengan kedua pendekatan ini, masalah yang penting adalah bagaimana menentukan kesamaan antara dua objek, sehingga kelompok dapat dibentuk dari objek dengan kesamaan tinggi satu sama lain. Umumnya, fungsi jarak, seperti fungsi jarak *Manhattan* dan *Euclidian*, digunakan untuk menentukan kesamaan. Fungsi jarak menghasilkan nilai yang lebih tinggi untuk pasangan objek yang kurang mirip satu sama lain. Kadang-kadang fungsi kesamaan digunakan sebagai gantinya, yang menghasilkan nilai lebih tinggi untuk pasangan yang lebih mirip.

1. **Metode *Agglomerative Hierarchical Clustering***

**Pengelompokan *hirarchical clustering* umumnya ada dua jenis yaitu *Agglomerative (Botton-up)* dan *Devise (Top-Down).* Namun untuk penelitian ini menggunakan  *Agglomerative Hierarchical Clustering* dikarenakan, hasil dari pengelompokan dapat dilihat dengan dendogram, tidak diperlukan penetuan jumlah *cluster* pada awal pengelompokan, dan *Agglomerative Hierarchical Clustering* dengan mengguakan pendekatan bawah-atas (*button-up)* dimana pengelompokan data dimulai dari kecil ke pengelompokan yang besar.**

***Agglomerative Hierarchical Clustering* (AHC) dengan menggunakan *button-up,* dimulai dari masing-masing data sebuah *cluster*, kemudian secara rekursif mencari kelompok terdekat sebagai pasangan yang kemudian akan digabungkan menjadi kelompok lebih besar. Proses tersebut diulang terus sehingga tampak bergerak keatas membentuk hirarki.(Kurniawan, 2017)**

Terdapat tiiga teknik kedekatan dalam ***Hierarchical Clustering* (AHC), yaitu; *single linkage* (jarak terdekat) atau tautan tunggal, *average linkage* (jarak rata-rata) atau tautan rata-rata, *complete linkage* (jarak terjauh) atau tautan lengkap.**

1. ***Single Linkage***

***Single linkage* menemukan kedektan diantara dua kelompk terdekat (terkecil) antara dua data dari *cluster* yang berbeda.**

**Rumus :**

Keterangan :

= jarak antara data U dan V terdekat dari masing-masing *cluster* U dan V

1. *Average Linkage*

*Average linkage* menetukan kedektan diantara daua kelomok dari jarak rata-rata antara dua data dari cluster yang berbeda.

Rumus :

Keterangan :

|U| dan |V| = jumlah data yang ada dalam *cluster* U dan V

1. *Complete Linkage*

*Complete Linkage* menentukan kedekatan diantara dua kelompok dari jarak terjauh antara dua data dari *cluster* yang berbeda.

Rumus :

Keterangan :

= jarak antara data U dan V dari masing-masing *cluster* U dan V

1. **Pengujian**

Pengujian yang dilakukan pada penelitian ini adalah *Shihoutte Coeficient. Shihoutte Coeficient* untuk mengetahui seberapa dekat relasi antara objek dalam sebuah *cluster* dan seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. *Silhouette Coefficient* merupakan gabungan dari dua metode yaitu metode *cohesion* yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah *cluster*, dan metode *separation* yang berfungsi untuk mengukur seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. Hasil nilai *Shihoutte Coeficient* mendekati 1 maka semakin baik kualitas *cluste*r, jika semakin mendekati -1 maka kualitas *cluster* kurang baik, semakin mendekati 0 menunjukkan semakin sedikit dokumen yang benar pengelompokkannya.

# BAB III METODELOGI PENELITIAN

## **Objek Penelitian**

Penelitian “Pengelompokan Data Bank Sampah Menggunakan Metode *Agglomerative Hierarchical Clustering (AHC).* Mengambil objek penelitian dari Bank Sampah yang dikelola oleh ibu Utaminingsih Liniarti S.T., M.T. dengan atribut penelitian utama, kontribusi bangunan,menejemen bank sampah, mekanisme kerja bank sampah, pelaksanaan bank sampah dalam bentuk *file excel.*

## Metode Pengumpulan data

1. **Wawancara**

Metode wawancara ini dilakukan dengan cara tanya jawab langsung kepada Utaminingsih Liniarti S.T., M.T. selaku Sekretaris Program Studi Teknik Industri Universitas Ahmad Dahla dalam bentuk *file Excel*.

1. **Studi Literatur**

Pengumpulan data dilakukan dengan studi literatur dari berbagai macam artikel,publikasi ilmiah untuk mempelajari :

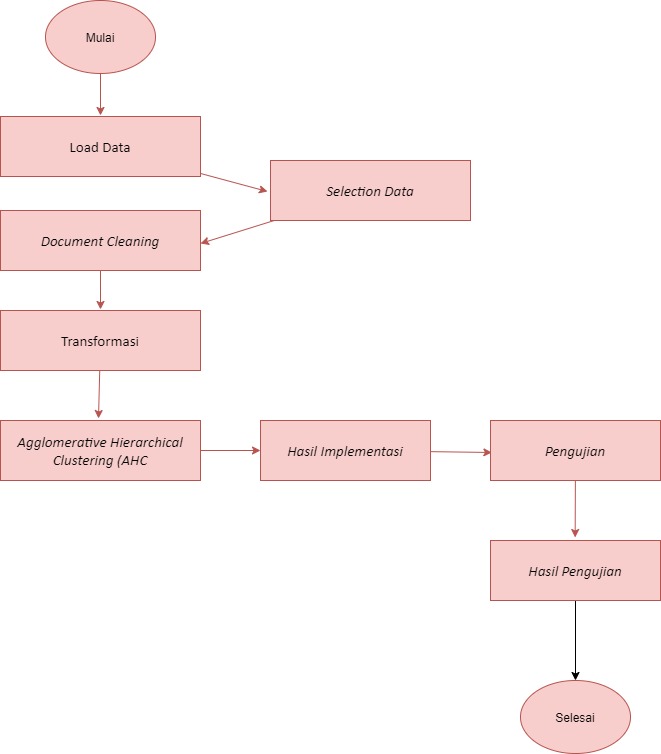
1. Mekanisme data *mining* dalam pengelompokan data.
2. Mekanisme *Similarity* dengan Menggunakan *Euclidean Distance*.
3. Pengelompokan data menggunakan metode *Agglomerative Hierarchical Clustering (AHC).*

## Alat Penelitian

Alat penelitian dibagi menjadi dua yaitu perangkat keras dan perangkat lunak. Alat yang dibutuhkan dalam melakukan penelitian ini sebagai berikut :

1. **Kebutuhan Perangkat Keras** 
   * 1. Processor minimal AMD A6 ~2.2GHz
     2. Memory 4 GB
     3. Kapasitas Hardisk minimal 500 GB
     4. Monitor 14” LCD
2. **Kebutuhan Perangkat Lunak**
3. Bahasa pemrograman *Python* 3 *, Micro framework flask* dan *Sublime Text.*
4. Web Browser seperti *Google Chrome*, Microsoft Edge, *UC Browser* dan lain sebagainya.
5. Modul bahasa pemrograman *Python* seperti *Anaconda Prompt.*
6. **Tahapan Data Mining**

Adapun tahapan yang akan dilakukan untuk menyelsaikan ini seperti pada gambar 3.1. tahapan penelitian data mining :



Gambar 3. 1 Tahapan Penelitian Data Mining

Gambar 3.1 menampilkan tahapan penelitian data mining dengan keterangan sebagai berikut:

1. Load Data

Load data merupakan awal dari tahapan pengembangan sistem. Data kelulusan yang dikumpulkan dalam bentuk *excel* kemudian di *load* kedalam program agar data kelulusan dapat di olah pada tahapan selanjutnya.

1. Selection Data

*Selection* data merupakan awal dari tahapan penelitian. Data bank sampah yang telah terkumpulkan dalam bentuk *file excel* kemudian diseleksi dengan menggunakan atribut bank sampah yang terkumpul yaitu : Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah, sehingga dapat diolah pada tahap selanjutnya

1. *Cleaning Data*

Setelah data bank sampah berhasil di *selection* maka selanjutnya adalah *cleaning* data. *Cleaning* data merupakan proses yang digunakan untuk membuat data yang tidak konsisten dan bersifat *noise*. Proses ini menggunakan fitur *dropna* yang ada pada Bahasa pemrograman *Python*.

1. Transformasi

Transformasi data pada penelitian ini merupakan proses pengkategorian atau perubahan data ke format yang sesuai untuk proses data mining. Transformasi dilakukan dengan cara mengubah format dan pengkategorian data sehingga lebih mudah untuk diolah.

1. *Agglomerative Hierarchical Clustering (AHC)*

Sebelum masuk ke tahapan pengelompokan data yang telah melalui tahapan clustering, kemujian dicari jarak kemiripan antara bank sampah yang satu degan bank sampah yang lainnya dengan menggunakan rumus *Euclidean Distance*  yang digunakan untuk mencari jarak kemiripan dari nilai masing-masing variable setiap data.

Dari perhitungan jarak antar bank sampah menghasilkan data yang dapat digunakan untuk pengelompokkan data yaitu menggunakan *Agglomerative Hierarchical Clustering (AHC)*. Data dikelompokkan berdasarkan kemiripan jarak yang dipilih secara random dari data sesuai jumlah kelompok yang akan dibentuk, sehingga menghasilkan kelompok dengan anggota didalamnya.

1. Hasil Implementasi

Tahapan implementasi merupakan tahapan untuk mengimplementasikan perancangan menggunakan *Agglomerative hirarchical clustring* ke dalam *system* yang akan dibuat sehingga sistem tersebut mudah dimengerti oleh *user*. Tahapan pengimplementasian dalam bahasa pemrograman *Python* akan menghasilkan program untuk mengelompokkan data bank sampah. Pada tahapan ini untuk menerapkan tahapan *data mining* kedalam web aplikasi.

1. Pengujian

Setelah system selesai dibuat maka tahapan selanjutnya adalah pengujian terhadap sistem. Pengujian ini bertujuan untuk menguji apakah system sudah berjalan sesuai keinginan atau belum. Pengujian terhadap system dilakukan dengan metode *Agglomerative hirarchical clustring.* Tahap pengujian dengan *Agglomerative hirarchical clustring* untuk mengetahui seberapa akurat hasil dari pengelompokan data bank sampah menggunakan metode *Agglomerative hirarchical clustring.*

1. Hasil Pengujian

Dari hasil pengujian akan menghasilakan kelompok cluster terbaik berdasarkan masing-masing data *cluster*  yang telah di uji. Data  *cluster*  dengan hasil nilai data tertinggi yang akan diterapkan menjadi kelompok data bank sampah yang terbaik berdasarkan hasil pengujian yang terbaik.

# BAB IV HASIL DAN PEMBAHASAN

## Hasil Pengumpulan Data

Berdarkan hasil pengumpulan data yang di dapatkan dari Ibu Utaminingsih Liniarti S.T., M.T. selaku Sekertaris Program Studi Teknik Industri Unversitas Ahmad Dahlan. Data yang digunakan dalam penelitian “Pengelompokan Data Bank Sampah Menggunakan Metode *Agglomerative Hirarchical Clustering (AHC)*” berupa data bank sampah.

Dataset berupa file *excel* dengan nama *file* “Pengelola.xlx” terdapat sebuat sheet dengan nama “*sheet1*”. Didalam *file* tersebut terdapat 46 kolom dan memiliki 126 data. Ke 46 kolom tersebut adalah Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah.

## Tahapan Data Mining

1. ***Load* Data**

Data bank sampah Kota Yogyakarta sudah dikumpulkan lalu di *load* kedalam program agar data bank sampah dapat di olah ke tahap selanjutnya. Hasil *Load* data dapat dilihat pada Tabel 4.1. dengan data yang digunakan 126 data.

Tabel 4. 1 Load Data

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | M01 | M02 | M03 | M04 | M05 | M06 | M07 | M08 | M09 | M010 |
| D1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 3 |
| D3 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 3 |
| D4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 4 |
| D5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 | 3 |
| D7 | 1 | 1 | 1 | 1 | 2 | 1 | 3 | 2 | 3 | 3 |
| D8 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| D9 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 4 |
| D10 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 3 | 3 |

1. ***Selection* Data**

Data yang telah diperoleh kemudian dilakukan proses *selection*. Dari 46 kolom pada data tersebut Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah. Hasil proses *selection* data dapat dilihat pada tabel 4.1. Data yang digunakan dalam penelitian ini menggunakan data bank sampah yang berisi 126 data. Selanjutnya akan di lakukan *cleaning* data pada bagian 4.2.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | M01 | M02 | M03 | M04 | M05 | M06 | M07 | M08 | M09 | M010 |
| D1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 3 |
| D3 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 3 |
| D4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 4 |
| D5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 | 3 |
| D7 | 1 | 1 | 1 | 1 | 2 | 1 | 3 | 2 | 3 | 3 |
| D8 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| D9 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 4 |
| D10 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 3 | 3 |

Tabel 4. 2 Selection Data

1. ***Cleaning* Data**

Setelah data bank sampah berhasil diseleksi maka tahapan selanjutnya adalah *cleaning* data. *Cleanging* data merupakan proses yang digunakan untuk membuang data yang tidak konsisten dan bersifat *noise.* Yang dapat dilihat pada tabel 4.3.

Tabel 4. 3 Cleaning Data

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | M01 | M02 | M03 | M04 | M05 | M06 | M07 | M08 | M09 | M010 |
| D1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 3 |
| D3 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 3 |
| D4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 4 |
| D5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 |
| D6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 | 3 |
| D7 | 1 | 1 | 1 | 1 | 2 | 1 | 3 | 2 | 3 | 3 |
| D8 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| D9 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 4 |
| D10 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 3 | 3 |

1. **Transformasi Data**

Transformasi yaitu proses pengkategorian atau perubahan data ke format yang sesuai untuk proses data mining. Transformasi dilakukan dengan cara mengubah format dan pengkategorian data sehingga lebih mudah untuk diolah. Yang dapat dilihat pada tabel 4.4.

Tabel 4. 4 Transformasi Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID |  |  |  |  |
| D1 | 27.39570479 | 45.44206518 | 20 | 26 |
| D2 | 43.19589629 | 47.75429294 | 24 | 26 |
| D3 | 36.69565612 | 54.90597851 | 18 | 26 |
| D4 | 41.12593562 | 52.05028212 | 8 | 26 |
| D5 | 27.16525253 | 47.33541053 | 24 | 27 |
| D6 | 27.0061271 | 50.21392334 | 24 | 28 |
| D7 | 32.74570496 | 66.56869413 | 27 | 28 |
| D8 | 43.11084251 | 41.35856812 | 16 | 25 |
| D9 | 47.01650688 | 51.4272936 | 16 | 28 |
| D10 | 50.37920813 | 54.00943167 | 18 | 27 |

1. **Tahapan Pengelompokkan *Agglomerative Hirarchical Clustering (AHC)***

Pada tahapan ini, dalam penulisan ini pengelompokan menggunakan *Agglomerative Hirarchical Clustering (AHC)*. Data input berupa excel selanjutnya diproses menggunakan AHC. Data Bank Sampah sejumlah 126 dengan 46 variable yang diproses menggunakan AHC senhingga akan terbentuk kelompok-kelompok. Setiap pembentukkan kelompok akan diuji menggunakan *Silhoutte Coeficient*. Proses pengelompokan dan pengujian akan dilakukan dengan menggunakan sistem yang telah dibuat. Proses AHC terdiri dari tiga metode tetapi pada penelitian ini menggunakan satu metode yaitu *complete linkage .*

Tahapan selanjutnya *Euclidean Distance* digunakan untuk mengukur jarak terdekat pada setiap variable bank smapah. Jarak terdekat iyalah nilai jarak *Euclidean* yang paling kecil atau nilai mendekati 0. Hasil proses *Euclidean Distance* dapat dilihat pada tabel 4.5.

Tabel 4. 5 Euclidean Distance

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Euclidean Distance* | | | | | | | | | | |
| Nama BS | BS 1 | BS 2 | BS 3 | BS 4 | BS 5 | BS 6 | BS 7 | BS 8 | BS 9 | BS 10 |
| BS 1 | 0 | 2 | 2.107 | 1 | 1 | 1 | 2.484 | 3 | 2.625 | 2.213 |
| BS 2 | 2.310 | 0 | 2.599 | 3.009 | 2.310 | 2.615 | 3.535 | 4.096 | 3.391 | 2.397 |
| BS 3 | 2.107 | 2.599 | 0 | 3.203 | 2.107 | 2.408 | 1.952 | 4.376 | 2 | 3.349 |
| BS 4 | 2 | 3.009 | 3.203 | 0 | 1.749 | 2.279 | 3.441 | 2.124 | 2.464 | 2.394 |
| BS 5 | 1 | 2.310 | 2.107 | 1.749 | 0 | 1.082 | 2.484 | 3.023 | 2.625 | 2.213 |
| BS 6 | 1 | 2.615 | 2.408 | 2.279 | 1.082 | 0 | 2.279 | 2.981 | 3.359 | 1.868 |
| BS 7 | 2.484 | 3.535 | 1.952 | 3.441 | 2.482 | 2.279 | 0 | 4.142 | 3 | 3 |
| BS 8 | 2 | 4.096 | 2.376 | 2 | 3.023 | 2.981 | 4.142 | 0 | 3 | 3.907 |
| BS 9 | 2.625 | 3.391 | 2 | 2.464 | 2.625 | 3.359 | 3 | 3 | 0 | 3.684 |
| BS 10 | 2.213 | 2.397 | 3.349 | 2.394 | 2.213 | 1.868 | 3 | 2.907 | 3.684 | 0 |

Rumus :

= = 2

Berdasarkan matrik jarak, selanjutnya dilakukan pengelompokkan data dengan *Agglomerative Hirarchical Clustering (AHC).*  Proses AHC terdiri dari tiga metode tetapi pada penelitian ini menggunakan satu metode yaitu metode *complete linkage .*

Metode complete linkage ( jarak terjauh) adalah jarak antara kelompok didefinisikan sebagai jarak antara pasangan data paling jauh dari setiap kelomppok.

1. Dengan memperlakukan data sebagai kelompok, selanjutnya penulis memlih jarak dua kelompok terkecil.

Rumus :

Keterangan :

= jarak antara data U dan V terdekat dari masing-masing *cluster* U dan V

1. Terpilih kwlompok 1 dan 5, kedua kelompok ini digabungkan untuk melanjutkan pengelompokan.
2. Menghitung jarak antara kelompok 1 dan 5 dengan kelompok lain yang masih tersisa yaitu, 2,3,4,6,7,8,9,10.

D((1,5)2) = max(d(1,2),d(5,2) =max(2.310,2310) = 2.310

1. Dengan menghapus baris-baris dan kolom-kolom matrik jarak yang bersesuaian dengan kelompok 1 dan 5, serta menambahkan baris dan kolom untuk kelompok (1,5).

Tabel 4. 6 Tahap 1 *Complete Linkage*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | D(1,5) | D2 | D3 | D4 | D6 | D7 | D8 | D9 | D10 |
| D1 | 0 | 2.310 | 0 | 2.599 | 0 | 3.441 | 0 | 3.023 | 0 |
| D2 | 2.310 | 0 | 2.599 | 0 | 3.441 | 0 | 3.023 | 0 | 3.359 |
| D3 | 2.107 | 2.599 | 0 | 3.441 | 0 | 3.023 | 0 | 3.359 | 0 |
| D4 | 1.749 | 3.009 | 3.441 | 0 | 3.023 | 0 | 3.359 | 0 | 2.279 |
| D6 | 1 | 2.310 | 2.484 | 3.023 | 0 | 3.359 | 0 | 2.279 | 0 |
| D7 | 1.082 | 2.615 | 2.408 | 2.981 | 3.359 | 0 | 2.279 | 0 | 4.142 |
| D8 | 2.484 | 3.535 | 1.952 | 4.142 | 2.482 | 2.279 | 0 | 4.142 | 3.907 |
| D9 | 3.023 | 4.096 | 4.142 | 2 | 3.023 | 2.981 | 4.142 | 0 | 3.684 |
| D10 | 2.625 | 3.391 | 3 | 2.464 | 2.625 | 3.359 | 3 | 3.684 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil.
2. Menghitung jarak (1,5)6 dengan kelompok lain yang tersisa.

Tabel 4. 7 Tahap 2 *Complete Linkage*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | D(1,5,6) | 2 | 3 | 4 | 7 | 8 | 9 | 10 |
| D1 | 0 | 2.310 | 0 | 2.599 | 0 | 3.441 | 0 | 0 |
| D2 | 1 | 0 | 2.599 | 0 | 3.441 | 0 | 3.023 | 3.359 |
| D3 | 0 | 2.599 | 0 | 3.441 | 0 | 3.023 | 0 | 0 |
| D4 | 2107 | 3.009 | 3.441 | 0 | 3.023 | 0 | 3.359 | 2.279 |
| D7 | 1749 | 2.310 | 2.484 | 3.023 | 0 | 3.359 | 0 | 0 |
| D8 | 0 | 3.535 | 1.952 | 4.142 | 2.279 | 0 | 2.279 | 4.142 |
| D9 | 0 | 4.096 | 4.142 | 2 | 2.981 | 4.142 | 0 | 3.907 |
| D10 | 2279 | 3.391 | 3 | 2.464 | 3.359 | 3 | 3.684 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6)10 dengan kelompok lain yang tersisa

Tabel 4. 8 Tahap 3 *Complete Linkage*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | D(1,5,6,10) | 2 | 4 | 7 | 8 | 9 | 10 |
| D1 | 0 | 2.599 | 3.349 | 3.535 | 4.096 | 0 | 0 |
| D2 | 2.599 | 0 | 3.441 | 4.096 | 3.391 | 3.023 | 3.359 |
| D3 | 3.349 | 3.441 | 0 | 3.391 | 2.484 | 0 | 0 |
| D4 | 3.535 | 3.009 | 3.441 | 0 | 0 | 3.359 | 2.279 |
| D7 | 4.096 | 2.464 | 2.625 | 3.359 | 0 | 0 | 0 |
| D8 | 3.391 | 2.615 | 2.464 | 2.625 | 3.359 | 0 | 4.142 |
| D9 | 2.484 | 3.535 | 2.482 | 3.441 | 2.482 | 4.142 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6,10)9 dengan kelompok lain yang tersisa.

Tabel 4. 9 Tahap 4 *Complete Linkage*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | D(1,5,6,,10,9) | 2 | 4 | 7 | 8 | 9 |
| D1 | 0 | 2.615 | 3.349 | 2.394 | 2.213 | 1.868 |
| D2 | 2.615 | 0 | 2.599 | 3.441 | 2.484 | 3.023 |
| D3 | 3.349 | 2.599 | 0 | 3.441 | 2.615 | 2.981 |
| D4 | 2.394 | 3.441 | 3.441 | 0 | 3.441 | 3.023 |
| D7 | 2.213 | 2.484 | 3.023 | 3.023 | 0 | 3.359 |
| D8 | 1.868 | 2.615 | 2.981 | 3.359 | 3.359 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6,10,9)7 dengan kelompok lain yang tersisa.

Tabel 4. 10 Tahap 5 *Complete Linkage*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | D(1,5,6,10,9,7) | 2 | 4 | 7 | 8 |
| D1 | 0 | 3.391 | 3.349 | 2.464 | 2.625 |
| D2 | 3.391 | 0 | 3.349 | 3.441 | 3.441 |
| D3 | 3.349 | 3.349 | 0 | 2.484 | 3.023 |
| D4 | 2.464 | 3.441 | 3.441 | 0 | 3.023 |
| D8 | 2.625 | 2.484 | 3.023 | 3.023 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6,10,9,7)4 dengan kelompok lain yang tersisa.

Tabel 4. 11 Tahap 6 *Complete Linkage*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | D(1,5,6,10,9,7,4) | 2 | 4 | 8 |
| D1 | 0 | 3.535 | 3.349 | 3.441 |
| D2 | 3.535 | 0 | 4.376 | 3.441 |
| D3 | 3.349 | 4.376 | 0 | 3.441 |
| D8 | 3.441 | 3.441 | 3.441 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6,10,9,7,4)8 dengan kelompok lain yang tersisa.

Tabel 4. 12 Tahap 7 *Complete Linkage*

|  |  |  |  |
| --- | --- | --- | --- |
| ID | D(1,5,6,7,10,9,7,4,8) | 2 | 8 |
| D1 | 0 | 3.535 | 3.349 |
| D2 | 3.535 | 0 | 4.376 |
| D3 | 3.349 | 4.376 | 0 |

1. Selanjutnya dipilih secara random dari kelompok terkecil
2. Menghitung jarak (1,5,6,10,9,7,4,8) dengan kelompok lain yang tersisa.

Tabel 4. 13 Tahap 8 *Complete Linkage*

|  |  |
| --- | --- |
| D(1,5,6,7,10,9,7,4) | 2 |
| 0 | 4.096 |
| 4.096 | 0 |

Pengelompokan menggunakan *complete linkage* memberikan informasi berupa hasil cluster.

Tabel 4. 14 *Cluster* Pengelompokan

|  |  |
| --- | --- |
| Cluster 1 | Cluster 2 |
| Titik pusat BS1 | Titik pusat BS5 |
| BS4 | BS2 |
| BS3 | BS6 |
| BS8 | BS9 |
| BS7 | BS8 |
| BS7 | BS10 |
| BS3 | BS4 |
| BS3 | BS7 |
| BS1 | BS3 |

1. ***Silhoutte Coeficient***

*Silhoutte Coeficient* merupakan metode yang digunakan untuk mellihat kualitas dan kekuatan dari *cluster*. Dengan perhitungan *Silhoutte Coeficient* dapat dilihat hasilnya, jika nilai mendekati 1 maka semakin baik bentuk *cluster*nya.

Langkah-langkah untuk menghitung *Silhoutte Coeficient* sebagai berikut :

1. Menghitung nilai rata-rata jarak objek dengan dokumen yang berada dalam satu cluster dengan menggunakan persamaan

Manghitung jarak antara dokumen satu dengan dokumen yang lain.

a(1) =

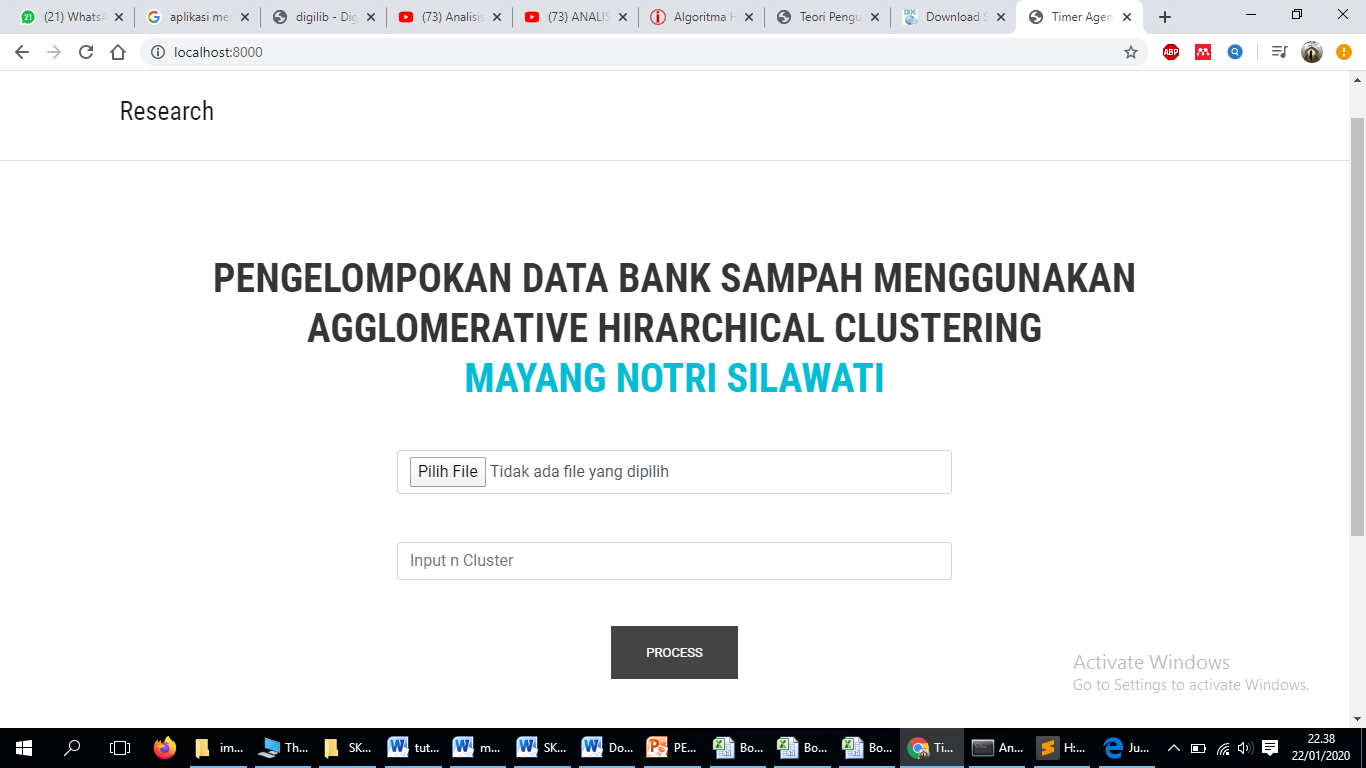
a(2) =

1. **Implementasi**

Implementasi *data mining* menggunakan bahasa Pemrograman *Python* yang digabungkan dengan *Framework Flask.* Menghasilakan aplikasi *data mining* berbasis  *web* dengan bentuk implementasi sebagai berikut :

1. ***Dashboard***

Implementasi untuk dashboard sistem dapat dilihat pada Gambar 4.1. *Interface Dashboard*



Gambar 4. 1 Interface Dashboard

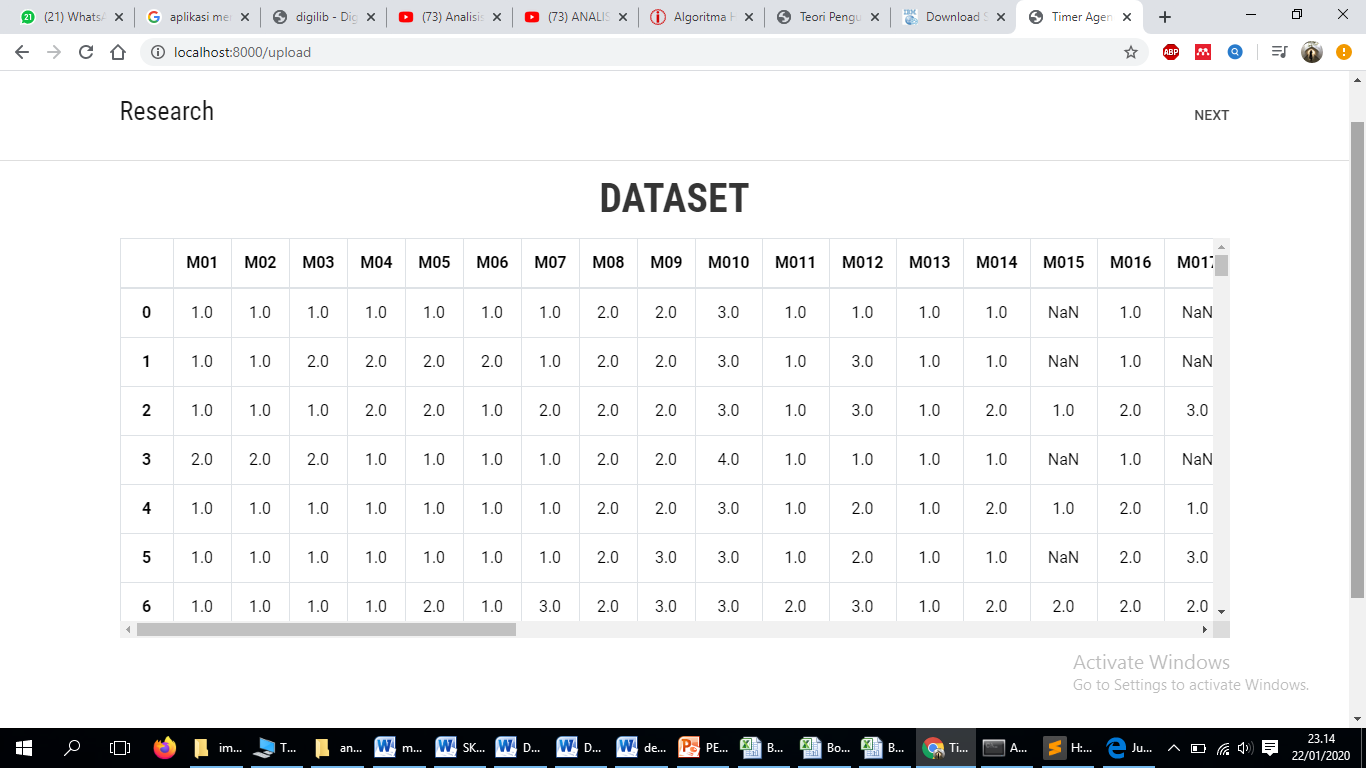
|  |  |
| --- | --- |
|  | <section class="cd-intro"> |
|  | <h1 class="wow fadeInUp animated cd-headline slide" data-wow-delay=".4s" > |
|  | <span>Pengelompokan Data Bank Sampah Menggunakan Agglomerative Hirarchical Clustering</span><br> |
|  | <span class="cd-words-wrapper"> |
|  | <b class="is-visible">MAYANG NOTRI SILAWATI </b> |
|  | <b>TEKNIK INFORMATIKA</b> |
|  | </span> |
|  | </h1> |
|  | </section> <!-- cd-intro --> |
|  | <!-- /.slider --> |
|  | <form method="post" action="upload" role="form" style="width: 50%; margin: 0 auto 0 auto;" enctype="multipart/form-data"> |
|  | <div class="form-group wow fadeInDown animated" data-wow-duration="500ms" data-wow-delay=".6s" style="visibility: visible; animation-duration: 500ms; animation-delay: 0.6s; animation-name: fadeInDown;"> |
|  | <label for="file">Upload File</label> |
|  | <input type="file" class="form-control" name="file" id="file" required> |
|  | </div> |
|  | <div class="form-group wow fadeInDown animated" data-wow-duration="500ms" data-wow-delay=".8s" style="visibility: visible; animation-duration: 500ms; animation-delay: 0.8s; animation-name: fadeInDown;"> |
|  | <label for="cluster">Input n Cluster</label> |
|  | <input type="number" min="0" placeholder="Input n Cluster" class="form-control" name="cluster" id="cluster" required> |
|  | </div> |
|  | <div id="submit" class="wow fadeInDown animated" data-wow-duration="500ms" data-wow-delay="1.4s" style="visibility: visible; animation-duration: 500ms; animation-delay: 1.4s; animation-name: fadeInDown;"> |
|  | <button type="submit" class="btn-lines dark light wow fadeInUp animated btn btn-default btn-green hvr-bounce-to-right" data-wow-delay=".9s">PROCESS</button> |
|  | </div> |
|  | <div id="submit" class="wow fadeInDown animated" data-wow-duration="500ms" data-wow-delay="1.4s" style="visibility: visible; animation-duration: 500ms; animation-delay: 1.4s; animation-name: fadeInDown;"> |
|  | <button type="submit" class="btn-lines dark light wow fadeInUp animated btn btn-default btn-green hvr-bounce-to-right" data-wow-delay=".9s">PROCESS</button> |
|  | </div> |

Listing 4. 1 Listing Dashboard

Keterangan :

1. Baris 1 – 10 : untuk tampilan *Dashboar*
2. Baris 11 – 15 : untuk menambahkan variabel data
3. Baris 16 – 19 : untuk menambahkan input clustering
4. Baris 20 - 25 : untuk memproses data / upload data
5. ***Load Data***

Load data pada sistem yang dapat dilihat pada Gambar 4.2.



Gambar 4. 2 Dataset

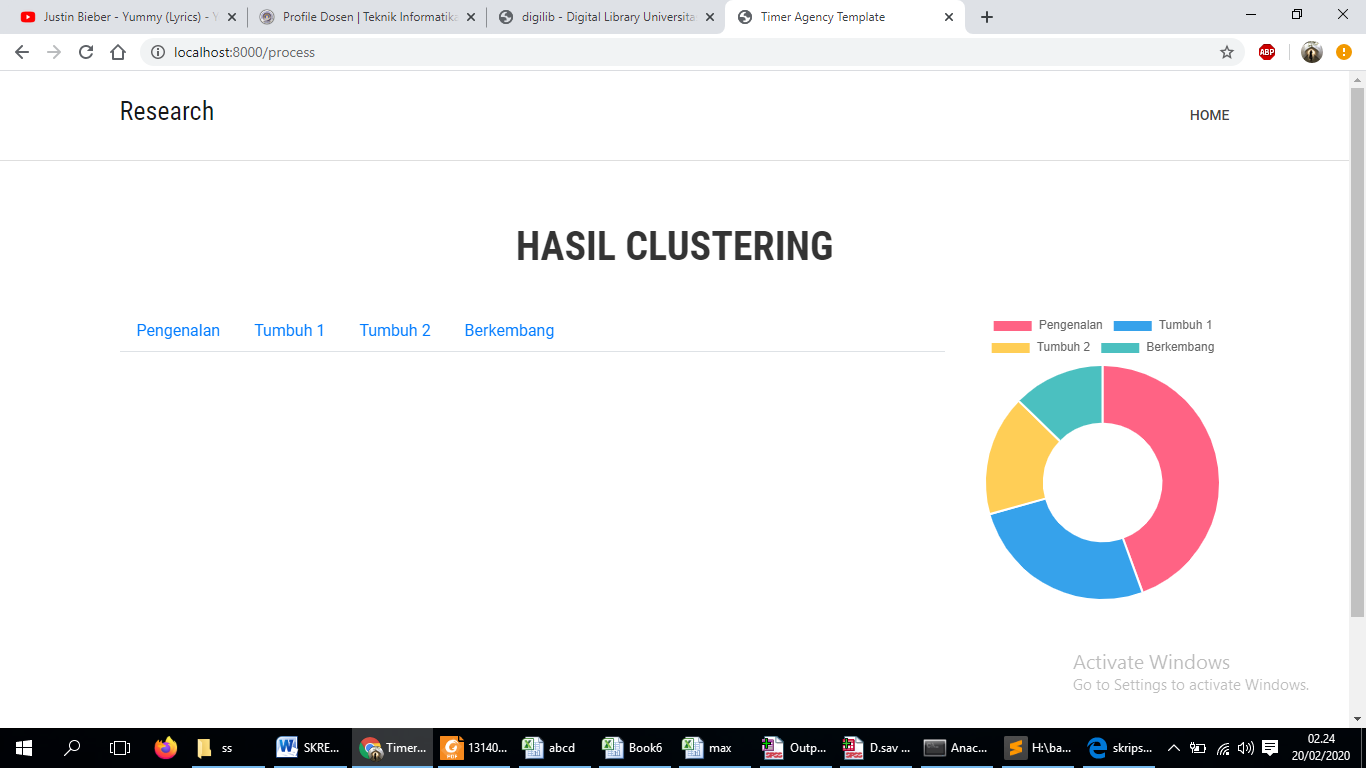
|  |  |
| --- | --- |
|  | class Sweet(object): |
|  |  |
|  | def \_\_init\_\_(self, file, cluster=4): |
|  | wb = pd.read\_excel(file) |
|  | data = pd.DataFrame(wb.values) |
|  | data.columns = ['M01', 'M02', 'M03', 'M04', 'M05', 'M06', 'M07', 'M08', 'M09', 'M010', 'M011', 'M012', 'M013', 'M014', 'M015', 'M016', 'M017', 'M018', 'M019', 'M020', 'M021', 'M022', 'M023', 'M024', 'M025', 'M026', 'M027', 'M028', 'M029', 'M030', 'M031', 'M032', 'M033', 'M034', 'M035', 'M036', 'M037', 'M038', 'M039', 'M040', 'M041', 'M042', 'M043', 'M044', 'M045', 'M046'] |
|  | self.dataset = data |
|  | self.cluster = cluster |
|  | self.X = None |
|  | self.labels = list() |
|  |  |
|  | def get\_data(self): |
|  | return self.dataset |
|  |  |
|  | def transform(self): |
|  | d = self.dataset.fillna(0) |
|  | new\_df = list() |
|  | for i, row in d.iterrows(): |
|  | v1 = 0 |
|  | v2 = 0 |
|  | v3 = 0 |
|  | v4 = 0 |
|  | offset = 0 |
|  | vals = 0 |
|  | for x in range(1, 6): |
|  | vals += row["M0" + str(x + offset)] |
|  |  |
|  | v1 = vals |
|  |  |
|  | offset = 6 |
|  | vals = 0 |
|  | for x in range(1, 6): |
|  | vals += row["M0" + str(x + offset)] |
|  |  |
|  | v2 = vals |
|  |  |
|  | offset = 12 |
|  | vals = 0 |
|  | for x in range(1, 17): |
|  | vals += row["M0" + str(x + offset)] |
|  |  |
|  | v3 = vals |
|  |  |
|  | offset = 29 |
|  | vals = 0 |
|  | for x in range(1, 17): |
|  | vals += row["M0" + str(x + offset)] |
|  |  |
|  | v4 = vals |
|  | new\_df.append([v1 \* (10 + rd.random()) / 2, v2 \* (10 + rd.random()) / 2, v3, v4]) |
|  |  |
|  | X = pd.DataFrame(new\_df).to\_numpy() |
|  | self.X = X |
|  | return X |

Listing 4. 2 Dataset

Keterangan :

1. Baris 1 – 13 : Tampilan hasil dari Dataset
2. Baris 15 – 54 : Menampilkan hasil data *transform* dan *cleaning*
3. ***Agglomerative Hirarcical Clustering***

Implementasi untuk clustering dapat dilihat dalam bentuk tabel dan dounat pada Gambar 4.3. *Interface* hasil *Clustering* dan *Listing* 4.3. *Listing* hasil *Clustering*



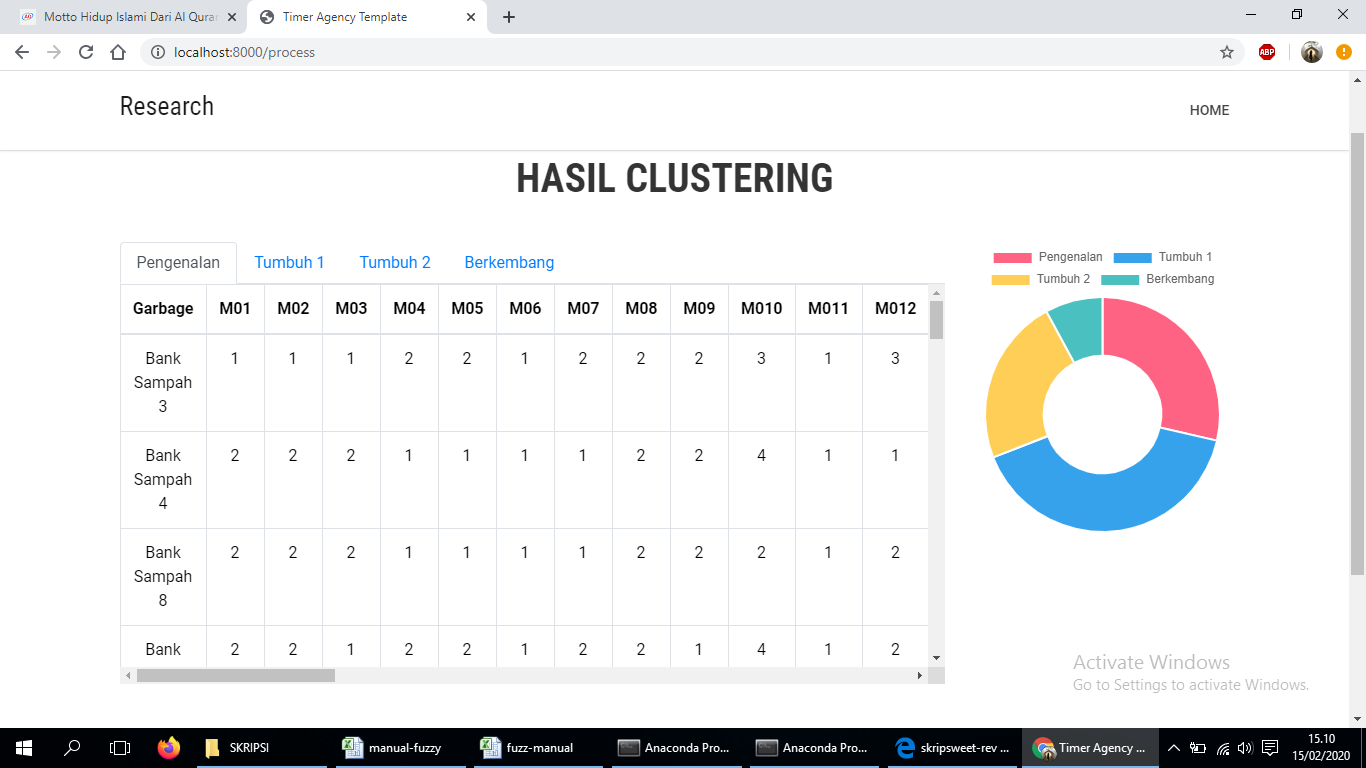
Gambar 4. 3 *Interface* hasil *Clustering*

|  |  |
| --- | --- |
|  | def get\_cluster(self): |
|  | # membuat model |
|  | model = AgglomerativeClustering(n\_clusters=self.cluster, affinity='euclidean', linkage='ward') |
|  | model.fit(self.X) |
|  | self.labels = model.labels\_ |
|  | return model |

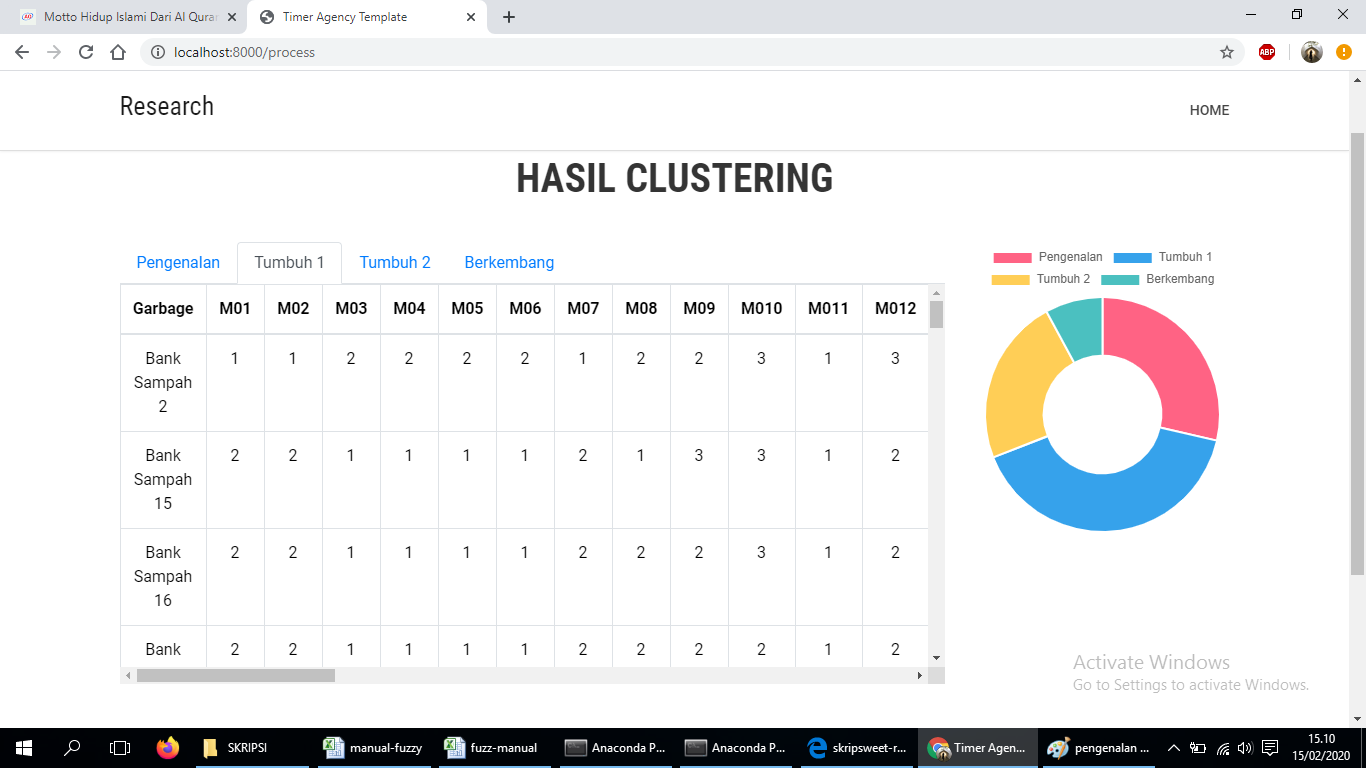
Listing 4. 3 Hasil *Clustering*

Keterangan :

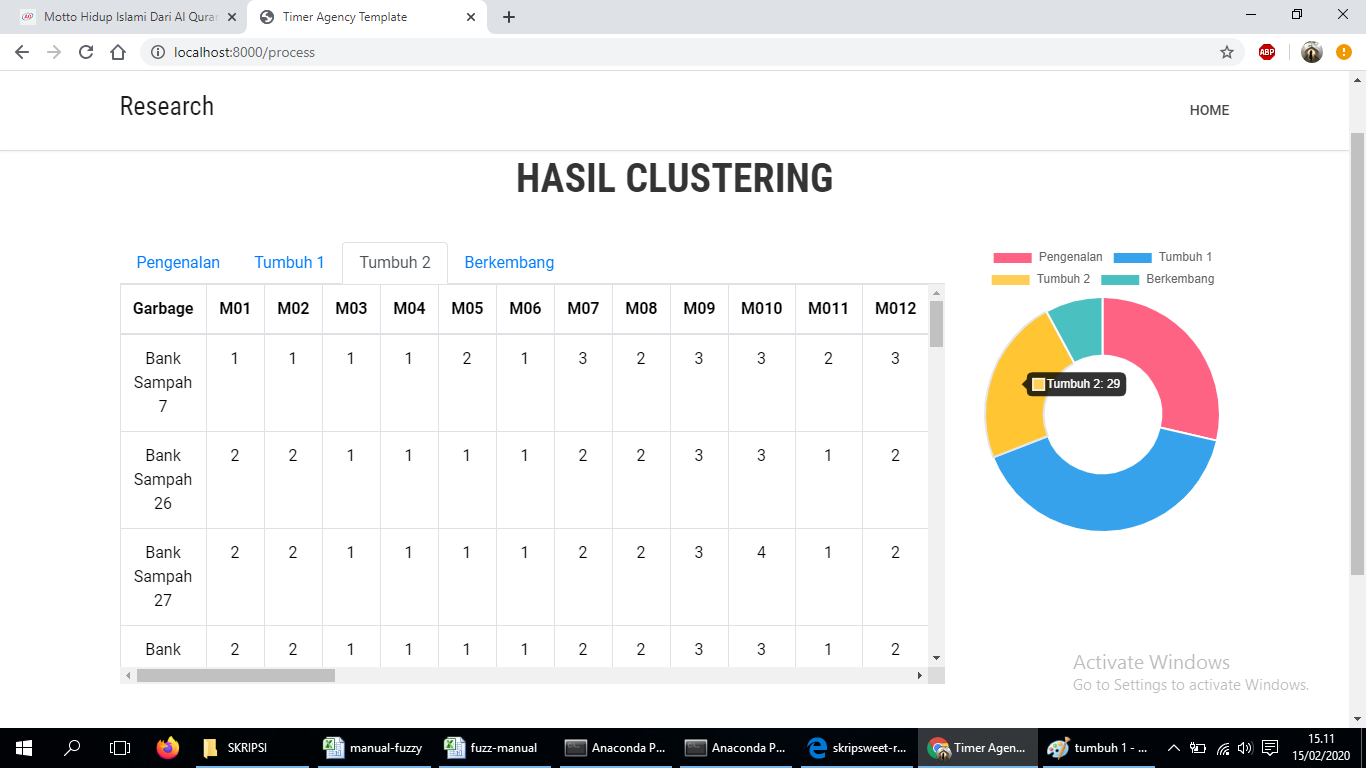
1. Baris 1 – 6 : *funcion Agglomerative Hirarcical Clustering* untuk *variabel* pengelompokan dalam bentuk tabel dan dounat.



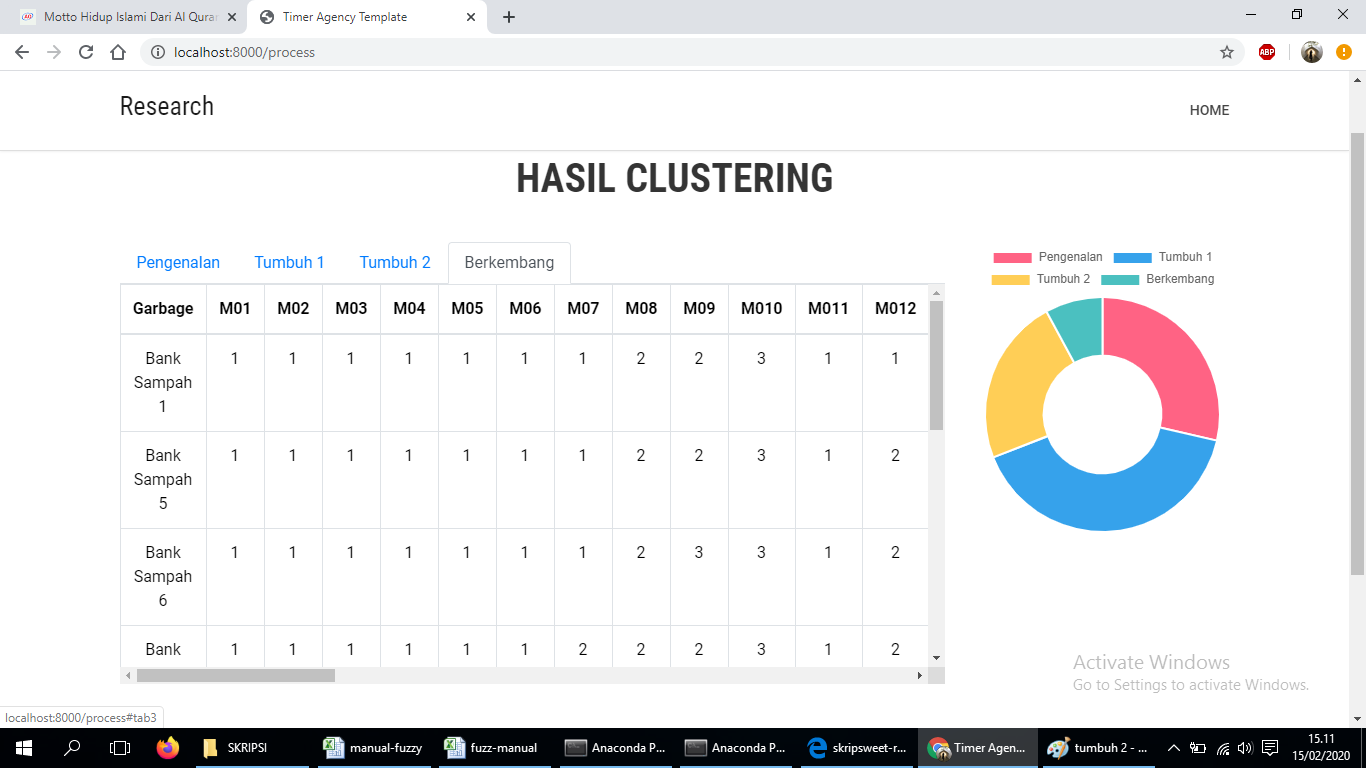
Gambar 4. 4 *Interface* hasil *Clustering* (Pengenalan )



Gambar 4. 5 *Interface* hasil *Clustering* (Tumbuh 1 )



Gambar 4. 6 *Interface* hasil *Clustering* (Tumbuh 2)



Gambar 4. 7 *Interface* hasil *Clustering* (Berkembang)

|  |  |
| --- | --- |
|  | @staticmethod |
|  | def get\_cluster\_label(argument): |
|  | switcher = { |
|  | 0: "Pengenalan", |
|  | 1: "Tumbuh 1", |
|  | 2: "Tumbuh 2", |
|  | 3: "Berkembang" |
|  | } |
|  | return switcher.get(argument, f"Cluster {argument}") |

Listing 4. 4 Hasil Clustering Pengenalan,Tumbuh 1,Tumbuh 2,Berkembang.

Keterangan :

1. Baris 1 – 4 :method argument cluster 0 menjadi Pengenalan
2. Baris 5 : cluster 1 menjadi Tumbuh 1
3. Baris 6 : cluster 2 menjadi Tumbuh 2
4. Baris 7 : cluster 3 menjadi Berkembang
5. **Pengujian Akurasi**

Untuk mengetahui hasil pengujian akurasi yang menggunakan *Silhoutte Coeficent* . Pengujin dilakukan untuk mengetahui kelompok yang baik berdasarkan variabel untuk diterapkan dalam jumlah data yang ada yaitu 126 berdasarkan hasil dari proses *cluster* dengan metode *Agglomerative Hirarcical Clustering*. Jika hasil nilai *Silhoutte Coeficient* mendekati 1 maka semakin baik kualitas *cluste*r, jika semakin mendekati -1 maka kualitas *cluster* kurang baik. Hasil dari *Silhoutte Coeficient*  dapat dilihat pada Tabel 4.8. Hasil *Silhoutte Coeficient.*

Tabel 4. 14 Hasil Silhoutte Coefficient

|  |  |
| --- | --- |
| *cluster* | Hasil pengujian (Si) |
| 2 | 0.187 |
| 3 | 0.295 |
| 4 | 0.343 |
| 5 | 0.298 |
| 6 | 0.314 |
| 7 | 0.300 |
| 8 | 0.349 |
| 9 | 0.316 |
| 10 | 0.238 |

Hitung nilai *Silhoutte Coeficient* dengan persamaan :

Keterangan :

S(i) = *Silhoutte Coeficient*.

a(i) = Rata-rata jarak objek dengan semua objek yang berbeda dalam satu *cluster*.

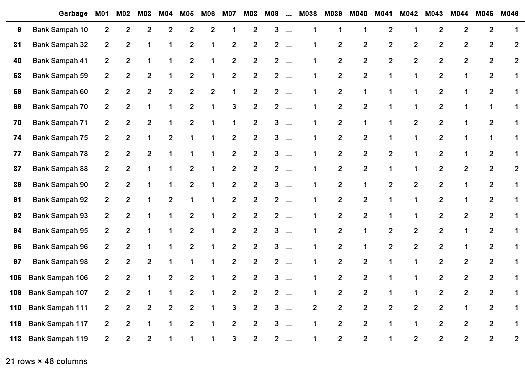
b(i) = Rara-rata jarak objek dengan semua objek lain yang berbeda pada *cluster* lainnya.

Data yang diujikan adalah data yag bersih dalam jumlah 126 data. Dari jumlah data *cluster* telah dilakukan pengujian sejumlah 10 *cluster*. Pengelompokan data dilakukan menggunakan program agar dapat dibandingkan dengan hasil dari pengelompokan menggunakan metode *Agglomerative Hirarcical Clustering*  *(AHC).* Dari perhitungan diatas didapatkan hasil *Silhoutte Coeficient* adalah terbesar 0.349. Hasil tersebut menunjukkan bahwa pengelompokkan dengan data tersebut paling baik apabila menggunakan 8 *cluster*.

Namun pihak peneliti menginginkan hanya menggunakan 4 cluster yang terdiri dari cluster 1, cluster 2, cluster 3, cluster 4 dengan hasil nilai pengujian menggunakan *Silhoutte Coeficient* adalah 0.343 . Dimna cluster 1 adalah cluster pengenalan, cluster 2 adalah cluster tumbuh 1, cluster 3 adalah cluster tumbuh 2, dan cluster 4 adalah cluster berkembang.

1. Cluster 1 (Pengenalan)

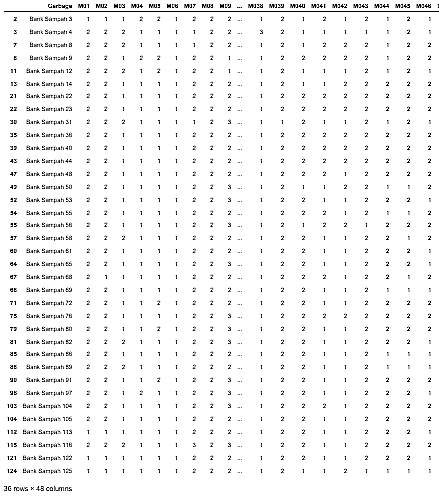
Pada cluster 1 (Pengenalan) terdapat 21 data bank sampah yang terdiri dari 46 variable yaitu variable MO1 sampai variable M046 yang meliputi : Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah.



Gambar 4. 8 Data *Cluster* 1

1. Cluster 2 (Tumbuh 1)

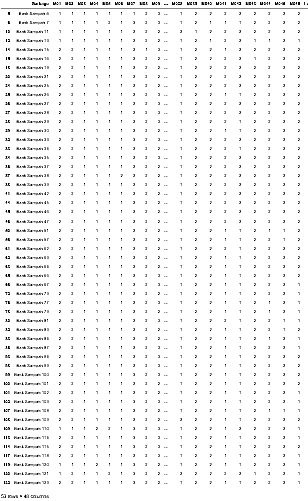
Pada *cluster 1* (Pengenalan) terdapat 36 data bank sampah yang terdiri dari 46 variable yaitu variable MO1 sampai variable M046 yang meliputi : Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah.



Gambar 4. 9 Data *Cluster* 2

1. Cluster 3 (Tumbuh 2)

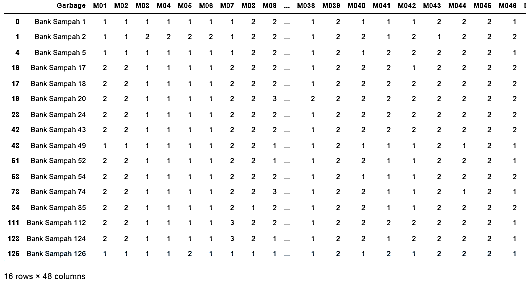
Pada cluster 1 (Pengenalan) terdapat 56 data bank sampah yang terdiri dari 46 variable yaitu variable MO1 sampai variable M046 yang meliputi : Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah.



Gambar 4. 10 Data *Cluster* 3

1. Cluster 4 (Berkembang)

Pada cluster 1 (Pengenalan) terdapat 16 data bank sampah yang terdiri dari 46 variable yaitu variable MO1 sampai variable M046 yang meliputi : Lantai, Dinding,Ventilasi,Pagar, Ruang Pelayanan, Luas Area, Penyuluhan Program Bank Sampah, Tersedia Wadah Sesuai Pemilahan, Jumlah Pengurus, Pendidikan ketua, Sistem Gaji, Pengelola mengikuti Pelatihan, Rekapitulasi Hasil Perolehan, Pengumpulan Sampah Nasabah Terjadwal, Jadwal Pengumpulan Sampah, Pengumpulan Rutin, Aktivitas Penjualan Sampah Nasabah, Bank Sampah Memiliki Kerjasama, Jumlah Pengempul Sampah yang Bekerjasama, Kerjasama Dengan Industri Daur Ulang Sampah, Dijual Semua Ke Pengepul Atau Industri Daur Ulang, Dimanfaatkan Untuk Pembuatan Produk Daur Ulang, Proses Monitoring Dan Evaluasi, Proses Monitoring, Monitoring dan Evaluasi Terhadap Pengurangan Atau Peningkatan Jumlah Nasabah, Monitoring dan Evaluasi Terhadap Jumlah Simpan Uang Nasabah yang Tersimpan, Hasil Monitoring dan Evaluasi Pengelolaan Bank Sampah Tercatat Dengan Baik, Hasil Monitoring dan Evaluasi Pengelolan Bank Sampah dilaporkan Kepada Pemerintah Kota melalui Dinas Lingkungan Hidup, Tidak Lanjut Berdasarkan Hasil Monitoring dan Evaluasi, Jumlah Nasabah, Nasabah Awal yang Aktif, Modal Awal, Kondisi Rata-Rata Berat Sampah Per Bulan, Rata-Rata Berat Sampah Per Bulan, Rata-Rata Hasil Tabungan Nasabah, Rata-Rata Jumlah Penambahan Nasabah, Rata-Rata Waktu Maksimal Nasabah Baru Bertahan, Ketentuan Aturan Hukum Sesuai Bentuk Bank Sampah, Ketetapan Harga Beli Sampah, Aktivitas Simpan-Pinjam, Penetapan Jam Kerja, Sistem Bagi Hasil Penjualan, Ketentuan Jelas Mengenai Kondisi Sampah, Ketentuan Jelas Mengenai Berat Minimum Sampah, Ketentuan Jelas Mengenai Jenis Sampah, Tersedia Jasa Penjemputan Sampah.



Gambar 4. 11 Data *Cluster* 4

**BAB V  
KESIMPULAN DAN SARAN**

1. **Kesimpulan**

Berdasarkan hasil penelitian “Pengelompokan Data Bank Sampah Kota Yogyakata Menggunakan Motode *Agglomerative Hierarchical Clustering (AHC)”* dapat ditarik kesimpulan sebagai berikut :

1. Telah dibuat aplikasi dengan bahasa pemrograman *Python* yang dapat mengelompokkan Data Bank Sampah Kota Yogyakarta.
2. Uji *silhoutte coefficient* pada aplikasi “Pengelompokan Data Bank Sampah Kota Yogyakata Menggunakan Motode *Agglomerative Hierarchical Clustering (AHC)”* dengan 126 data dan dibagi menjadi 4 kelompok sesuai dengan hasil pengujian yang telah dilakukan yaitu sebesar 0.343.
3. Hasil pengujian *silhoutte coefficient* tergolong rendah karena *Agglomerative Hierarchical Clustering (AHC)* merupakan pengelompokan dengan dataset yang kecil dan akan lebih efektif.
4. **Saran**

Penelitian “Pengelompokan Data Bank Sampah Kota Yogyakata Menggunakan Motode *Agglomerative Hierarchical Clustering (AHC)”* masih memiliki kekurangan, maka diharapkan adanya pengembangan lebih lanjut lagi. Saran yang dapat digunakan sebagai acuan untuk pengembangan sellanjutnya, yaitu :

1. Proses upload dilakukan secara manual dengan file yang sebatas berekstensi.xlsx. diharapkan untuk dikembangkan lagi dengan mengambil dari *dataset* suatu *website*.
2. Perlunya peningkatan untuk hasil akurasi yang lebih baik.
3. Hasil pengelompokan yang dihasulkan kurang bagus, sehingga diharapkan menggunakan metode yang cocok untuk data Bank Sampah Kota Yogyakarta agar pengelompokan yang duhasilkan sesuai dengan yang diharapkan.

# DAFTAR PUSTAKA

Ilmiah, J. *et al.* (no date) ‘PENERAPAN DATA MINING UNTUK MEMBENTUK KELOMPOK BELAJAR MENGGUNAKAN METODE CLUSTERING DI SMPN 19 BANDUNG Andre Catur Prasetyo Jurnal Ilmiah Komputer dan Informatika ( KOMPUTA )’.

Indonesia, R. . U. 4. (2012) ‘Presiden republik indonesia’, (1), pp. 1–5. doi: 10.1016/j.aquaculture.2007.03.021.

Kurniawan, A. A. (2017) ‘Implementasi Algoritma Agglomerative Hirarchical Clustering untuk Mengelompokkan Capaian Belajar Siswa SD’.

Pitaloka, N., Maulana, K. and Kurniati, P. A. (2009) ‘Pengelompokan Data Menggunakan Hierarchical Clustering (Ahc)’, p. 1. Available at: http://repository.telkomuniversity.ac.id/pustaka/files/94531/resume/pengelompokan-data-menggunakan-hierarchical-clustering-ahc-.pdf.

Zahrotun, L. (2015) ‘Analisis Pengelompokan Jumlah Penumpang Bus Trans Jogja Menggunakan Metode Clustering K-Means Dan Agglomerative Hierarchical Clustering ( Ahc )’, *Analisis Pengelompokan Jumlah Penumpang Bus Trans Jogja Menggunakan Metode Clustering K-Means Dan Agglomerative Hierarchical Clustering ( Ahc )*, 9(1), pp. 1039–1047.

**LAMPIRAN**