

Penerapan Metode KnA (Kombinasi *K-Means* dan *Agglomerative Hierarchical Clustering*) dengan Pendekatan *Single Linkage* untuk Menentukan Status Gizi pada Balita

Intan Alpiana, Lilik Anifah

Jurusan Teknik Informatika, Teknik Informatika, Universitas Negeri Surabaya

Intanalpiana@mhs.unesa.ac.id

Lilikanifah@unesa.ac.id

Abstract— *The nutrition fulfillment for children under five years old is a factor that must be considered in maintaining their health, because toddlers are in a developmental period that is vulnerable to nutrition. The attention of parents and officers of the Public Health Service Center (PUSKESMAS) is very necessary to monitor toddler nutrition. Therefore we need a system that can help to classify the nutritional status of children. The system is designed using the KnA method (a combination of K-means and Agglomerative Hierarchical Clustering) so that clustering results are better by combining the advantages of bottom-up clustering (agglomerative) and top-down clustering (k-means). This study attempts to categorize the data from Pulung Ponorogo health center with 7336 data (testing 3641 and learning 3695) with variables of sex, age, weight and height into 4 clusters of each toddler's nutritional status, namely for BB / U there 4 clusters BB S.Krg, BB Krg, BB Normal, BB Lbh, 4 clusters TB / U consisting of S. Short, Short, Normal, High and BB / TB 4 clusters S.Kurus, Skinny, Normal and Fat. In this study, the KnA method was able to properly classify toddlers' data in large numbers. The learning data is as much as 3695 produces an accuracy of BBU = 94%, TBU = 98%, and BBTB = 93%. From the results of the grouping conducted, it is used to determine the nutritional status of status according to BB / U, TB / U, and BB / TB, each of which is equipped with a handling solution. From the testing with the black-box method as a whole the system is running well, as well as the questionnaire method and as well as validation of nutritional status, a system that made already good and appropriate in terms of determining the nutritional status of children.*

Keywords— *Data Mining, Agglomerative Hierarchical Clustering, K-Means, Nutritional Status Of Toddlers.*

Abstrak— Pemenuhan gizi pada anak usia dibawah lima tahun merupakan faktor yang harus diperhatikan dalam menjaga kesehatan, karena balita merupakan masa perkembangan yang rentan terhadap gizi. Kepedulian orang tua dan petugas Pusat Pelayanan Kesehatan Masyarakat (PUSKESMAS) sangat diperlukan untuk memantau gizi balita. Oleh karena itu diperlukan sebuah sistem yang dapat membantu mengelompokkan status gizi balita. Sistem yang dirancang menggunakan metode KnA (kombinasi *K-means* dan *Agglomerative Hierarchical Clustering*) agar hasil *clustering* lebih baik dengan menggabungkan kelebihan metode *bottom-up clustering* (*agglomerative*) dan *top-down clustering* (*k-means*). Penelitian ini mencoba untuk mengelompokkan data balita dari puskesmas Pulung Ponorogo dengan jumlah 7336 data (testing 3641 dan learning 3695 data) dengan variable jenis kelamin, umur, berat badan dan tinggi badan kedalam 4 cluster dari masing-masing status gizi balita yaitu untuk BB/U terdiri dari 4 cluster BB S.Krg, BB Krg, BB Normal, BB Lbh, 4 cluster TB/U yang terdiri dari S.Pendek, Pendek, Normal, Tinggi dan BB/TB 4 cluster S.Kurus, Kurus, Normal dan Gemuk. Pada penelitian ini metode KnA mampu dengan baik mengelompokkan data balita dalam jumlah yang besar. Dengan data learning sebanyak 3695 menghasilkan akurasi sebesar BBU = 94%, TBU = 98%, dan BBTB = 93%. Dari hasil pengelompokan yang dilakukan, selanjutnya akan digunakan untuk menentukan status gizi balita baik itu menurut BB/U, TB/U, dan BB/TB, dimana masing-masing dilengkapi dengan solusi penanganan. Dari pengujian dengan metode *black-box* secara keseluruhan sistem ini sudah berjalan dengan baik, begitu juga dengan metode kuisioner dan validasi status gizi, sistem yang dibuat sudah baik dan sesuai dalam hal menentukan status gizi balita.

Kata Kunci— *Data mining, Agglomerative Hierarchical Clustering, K-Means, Status Gizi Balita.*

I. PENDAHULUAN

Kekurangan gizi merupakan masalah kesehatan yang sering menimpa balita-balita di Indonesia. Berdasarkan Pantauan Status Gizi (PSG) 2017 yang dilakukan Kementerian Kesehatan, bayi usia di bawah lima tahun

(Balita) yang mengalami masalah gizi pada 2017 mencapai 17,8%. Jumlah tersebut terdiri dari Balita yang mengalami gizi buruk 3,8% dan 14% gizi kurang Grafik masalah gizi balita yang terjadi di Indonesia tahun 2016-2017 seperti pada Gambar 1.

Pemenuhan gizi pada anak usia dibawah lima tahun (balita) merupakan faktor yang perlu diperhatikan dalam menjaga kesehatan, karena balita merupakan masa perkembangan yang rentan terhadap gizi dan membutuhkan zat-zat gizi yang lebih besar dari kelompok lain seperti anak-anak maupun remaja, sehingga balita paling mudah menderita kelainan gizi [11]. Gizi kurang atau malnutrisi pada balita membawa dampak negatif terhadap pertumbuhan fisik maupun mental, yang selanjutnya akan menghambat beberapa proses belajar yang dilakukan oleh balita seperti belajar berbicara, berjalan, makan. Kepedulian orang tua dan petugas Pusat Pelayanan Kesehatan Masyarakat (PUSKESMAS) sangat diperlukan untuk memantau gizi balita. Karena malnutrisi pada balita tidak terjadi secara tiba-tiba seperti penyakit pada umumnya. Tanda-tanda seperti berat badan kurang dari standar, stunting bisa menjadi indikator awal terjadinya malnutrisi pada balita [10].



Gbr. 1 Masalah Gizi Balita Indonesia (2016-2017)

Saat ini parameter yang umum digunakan di puskesmas dalam penentuan status gizi pada balita hanya berdasarkan Berat Badan menurut Umur (BB/U), yang terdapat pada Kartu Menuju Sehat (KMS) dan pengolahan data balita belum spesifik berbasis program, adapun prosedur yang dilakukan adalah petugas mencatat di formulir pemantauan status gizi balita, kemudian di cocokkan status gizi balita tersebut berdasarkan tabel baku rujukan WHO/NCHS. Tetapi Berat Badan menurut Umur (BB/U) tidak spesifik menunjukkan apakah balita tergolong kurus, gemuk, tinggi atau pendek. Sehingga petugas puskesmas maupun orang tua tidak mengetahui apakah balita tergolong dalam anak yang kurus, gemuk, normal, tinggi atau pendek [7]. Supaya lebih efektif dan efisien dalam menentukan status gizi balita, diperlukan sebuah sistem yang dapat membantu untuk mengetahui status gizi balita yang mencakup jenis antropometri yang berpengaruh terhadap penentuan status gizi balita.

Pada tahun 2013 telah dilakukan penelitian oleh Sudirman dkk, dengan judul penelitian "Analisa Klasifikasi status gizi dengan metode fuzzy K-means menggunakan aplikasi berbasis android". Dari penelitian yang dilakukan terhadap 114 data sampel, metode fuzzy c-means

menghasilkan jumlah kesamaan hasil klasifikasi terhadap perhitungan sebanyak 26 hingga 32 data sampel. Dan memiliki kesamaan hasil klasifikasi yang dihasilkan oleh sistem berkisar 22,81 % hingga 28,07 % [8]. Selanjutnya pada tahun 2015 dilakukan penelitian kembali mengenai status gizi balita oleh Windha Mega Pradnya Dhuhita dengan judul "Clustering Menggunakan Metode *K-Means* Untuk Menentukan Status Gizi Balita". Penelitian ini menyatakan bahwa algoritma *K-Means* hanya memiliki nilai akurasi 34% benar untuk mengelompokkan gizi balita [10]. Kemudian tahun 2018 penelitian kembali dilakukan oleh Muhammad hasan wahyudi dengan judul "Sistem Pendukung Keputusan Penentuan Status Gizi Balita Menggunakan Metode Naive Bayes". Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa metode Naive Bayes dapat diimplementasikan pada sistem pendukung keputusan penentuan status gizi balita dengan nilai kinerja sistem sebesar 0,867 atau 86,7% [12]. Kemudian Dewi ayu dan Arfhan Prasetyo melakukan penelitian kembali dengan judul "Sistem Penunjang Keputusan Untuk Menentukan Status Gizi Balita Menggunakan metode Fuzzy Tsukamoto". Uji sample terhadap data yang diperoleh, didapatkan hasil bahwa tingkat keberhasilan aturan fuzzy tsukamoto yang telah terbentuk dalam penentuan status gizi balita adalah 82,35% [3]. Dari beberapa penelitian tersebut perlu adanya penelitian kembali dalam penentuan status gizi balita dengan menggunakan metode dan parameter yang berbeda sehingga akan menghasilkan nilai akurasi yang lebih baik dalam mengelompokkan status gizi balita. Selain itu juga diperlukan output yang tidak hanya fokus terhadap status gizinya saja, melainkan masukan dan solusi untuk penanganan dari masing-masing hasil status gizi balita.

Berdasarkan permasalahan yang telah dijelaskan, penelitian ini mencoba merancang sistem yang dapat mempermudah proses penentuan status gizi pada balita. Sistem yang dirancang akan menggunakan kombinasi metode *Agglomerative Hierarchical Clustering* dengan *K-Means clustering*. Penelitian ini menggabungkan metode *agglomerative Hierarchical clustering* dengan *Kmeans clustering* dimaksudkan agar hasil *clustering* lebih baik, disini akan dikombinasikan kelebihan metode *bottom-up clustering (agglomerative)* dan *top-down clustering (k-means)*. Algoritma *bottom-up* baik dalam mengidentifikasi kelompok kecil sedangkan algoritma *top-down* baik dalam mengidentifikasi kelompok besar. Kombinasi antara metode *Hierarchical Clustering* dan *K-Means Clustering* ini telah diuji dan terbukti bahwa kombinasi ini lebih baik dibandingkan *K-Means* biasa. Pada sistem ini pengguna akan menginputkan data antropometri yang telah diketahui. Indikator yang digunakan dalam penelitian ini diantaranya Berat Badan menurut Umur (BB/U), Tinggi Badan menurut Umur (TB/U), dan Berat Badan menurut Tinggi Badan (BB/TB).

II. METODOLOGI PENELITIAN

Berikut ini merupakan alur jalannya diagram penelitian untuk pembuatan sistem “Penentuan Status Gizi Balita Menggunakan Metode KnA (kombinasi *k-means* dan *agglomerative hierarchical clustering*) :



Gbr. 1 Alur jalannya penelitian

Penelitian ini bertujuan untuk menerapkan salah satu metode dari data mining *clustering* yaitu *KnA* (kombinasi *K-means* dan *agglomerative hierarchical clustering*) sebagai dasar untuk mengetahui status gizi pada balita beserta solusi penanganannya dari masing-masing status gizi yang dihasilkan untuk memudahkan orang tua dan puskesmas dalam menangani masalah gizi pada balita. *Hybrid hierarchical* disini akan mengkombinasikan kelebihan metode *bottom-up clustering* (*agglomerative*) dan *top-down clustering* (*k-means*). Algoritma *bottom-up* baik dalam mengidentifikasi kelompok kecil sedangkan algoritma *top-down* baik dalam mengidentifikasi kelompok besar. *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Akan tetapi, hasil *clustering* dengan *K-Means* sangat bergantung pada pusat awal *cluster* [1]. Hasil *clustering* dengan metode *K-Means* baik jika penentuan pusat *cluster* tepat. Metode *Hierarchical Clustering* dapat digunakan untuk mengatasi masalah penentuan pusat *cluster* pada *K-Means*. Penggabungan metode *Hierarchical clustering* dengan *Kmeans* dimaksudkan agar hasil *clustering* lebih baik. Hasil dari

metode *Hierarchical clustering* digunakan untuk menentukan pusat *cluster*. Pusat *cluster* yang dihasilkan *Hierarchical clustering* selanjutnya digunakan sebagai pusat *cluster* awal pada perhitungan *k-means Clustering* [6]. Variabel dependen yang digunakan dalam penelitian ini adalah jenis kelamin, umur, berat badan dan tinggi badan, sedangkan variabel independennya adalah metode *KnA*.

Sebelum dilakukan pengelompokan, data balita harus di normalisasi terlebih dahulu. Data yang ada tidak dapat langsung dilakukan pemrosesan dikarenakan sering terdapat besaran angka yang cukup jauh antara variabel tinggi badan dan berat badan. Perbedaan jarak atau besaran angka yang cukup jauh ini dapat menyulitkan dalam proses pengelompokan. Sehingga untuk memperkecil besaran angka antar variabel adalah dilakukan normalisasi pada angka-angka yang ada di variabel tinggi badan dan berat badan menggunakan persamaan [10] :

$$\text{Nilai Normalisasi} \equiv \frac{\text{Nilai awal} - \text{Nilai minimal}}{\text{Nilai maksimal} - \text{Nilai minimal}} \quad (1)$$

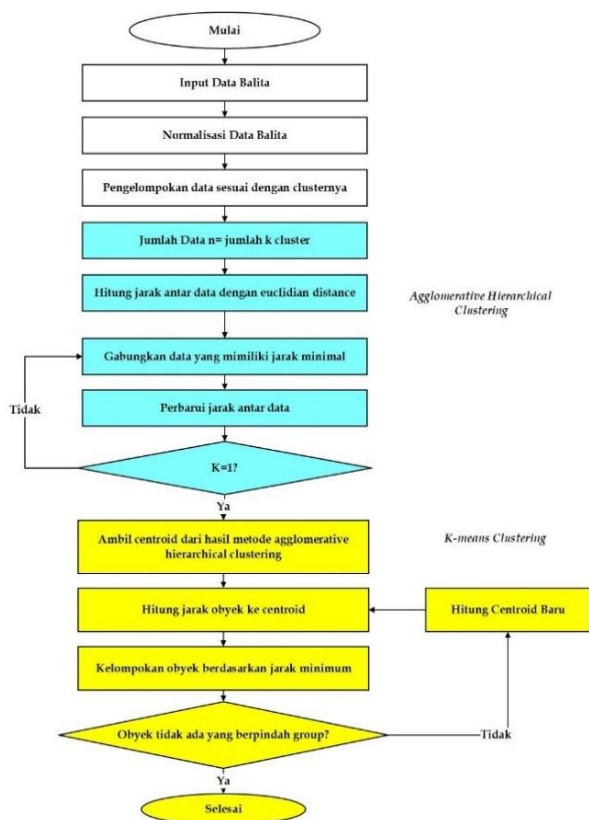
Nilai variabel tinggi badan dan berat badan akan dinormalisasi ke dalam rentang 0 – 1. Adapun tahapan yang dilakukan untuk proses normalisasi adalah Mencari nilai maksimum dan minimum untuk variabel tinggi badan (X) [10]. Setelah itu dimasukkan ke dalam persamaan dan akan menghasilkan data yang siap digunakan untuk pengelompokan data balita.

Alur jalannya perhitungan data balita menggunakan metode *KnA* untuk penentuan status gizi balita dalam penelitian ini adalah sebagai berikut :

1. Normalisasi data balita.
2. Pada penelitian ini clustering akan dilakukan menjadi tiga kelompok bagian yaitu :
 - a. Clustering 1 untuk BB/U (indeks berat badan menurut umur) dengan inputan data jenis kelamin, umur dan BB (berat badan) yang akan menghasilkan output berupa gizi baik, gizi buruk, gizi sedang dan obesitas.
 - b. Clustering 2 untuk TB/U (indeks tinggi badan menurut umur) dengan inputan jenis kelamin, umur dan TB (tinggi badan) yang akan menghasilkan output berupa sangat pendek, pendek, normal dan tinggi.
 - c. Clustering 3 untuk BB/TB (indeks berat badan menurut tinggi badan) dengan inputan jenis kelamin, umur, BB (berat badan), TB (tinggi badan) yang akan menghasilkan output berupa sangat kurus, kurus, normal, dan gemuk.
3. Data balita yang telah diinputkan sesuai dengan clusteringnya akan diproses menggunakan algoritma *agglomerative hierarchical clustering* dengan masing-masing teknik pendekatan yang ada pada *agglomerative hierarchical clustering*, diantaranya metode *single linkage*, *complete linkage* dan *average linkage*.

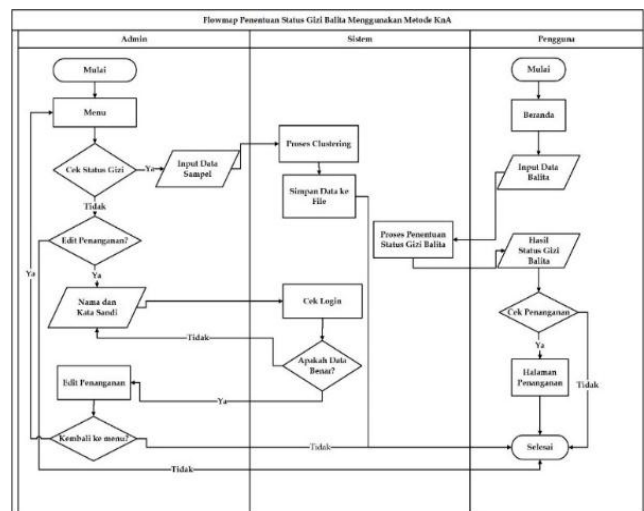
4. Penelitian ini menggunakan *encludiean distance* untuk menghitung jarak antar cluster.
5. Hasil dari metode *agglomerative Hierarchical clustering* akan digunakan untuk menentukan pusat *cluster*. Pusat *cluster* yang dihasilkan selanjutnya digunakan sebagai pusat *cluster* awal (centroid) pada perhitungan *K-means*.
6. Selanjutnya dilakukan perhitungan menggunakan algoritma *K-means clustering*.
7. Penelitian ini akan mengambil hasil terbaik dari metode *agglomerative hierarchical clustering* yang dikombinasikan dengan *K-means* dengan tujuan untuk mendapatkan solusi cluster terbaik dalam kasus pengelompokan status gizi balita.
8. Selanjutnya hasil pengelompokan terbaik dari metode KnA akan digunakan untuk menentukan status gizi pada balita.

Alur jalannya penentuan status gizi balita menggunakan algoritma KnA adalah sebagai berikut :



Gbr. 3 Alur Algoritma KnA

Sistem penentuan status gizi balita menggunakan metode KnA ini akan dirancang menggunakan bahasa pemrograman python dengan *framework flask* sebagai *server side* dan untuk *client side* menggunakan html, serta Json sebagai format penyimpanan datanya. Berikut ini merupakan alur jalannya sistem penentuan status gizi balita menggunakan metode KnA (kombinasi *K-means* dan *agglomerative hierarchical clustering*):



Gbr.4 Flowmap Sistem Penentuan Status Gizi Balita

III. HASIL DAN PEMBAHASAN

Tahap-tahap yang dilakukan dalam pengelompokan data balita adalah sebagai berikut :

A. Filterasi Data Balita



Gbr. 5 Peta kecamatan pulung

Untuk menentukan status gizi balita menggunakan metode KnA penulis menggunakan data balita yang telah dikumpulkan dari puskesmas Kecamatan Pulung Kabupaten Ponorogo. Kecamatan pulung memiliki 18 desa, namun data yang dipakai hanya data desa yang ada dipuskesmas pulung saja. Data yang digunakan terdiri dari 11 desa yang ada dikecamatan pulung yaitu desa banaran, bedrug, bekiring, karang patihan, patik, pulung merdiko, pulung, sidoharjo, singgahan, tegalrejo dan desa wagir kidul. Setelah itu dilakukan filter terhadap data balita di file *excel* untuk membersihkan data-data yang tidak memiliki variabel lengkap. Dari jumlah data mentah 8984 diperoleh sebanyak 7336 data lengkap yang kemudian dibagi rata untuk testing 3641 dan learning 3695. Kemudian data yang telah disimpan di file excel di simpan kembali dalam bentuk .csv yaitu learning.csv dan testing. Csv

B. Balancing Data

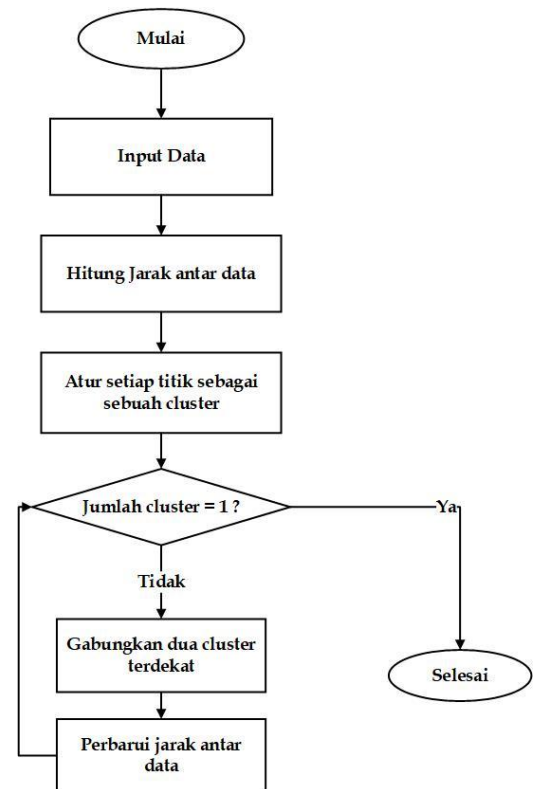
Pada proses penentuan status gizi balita menggunakan metode KnA, terkadang harus dilakukan proses *balancing* terlebih dahulu terhadap data balita yang memiliki jumlah data yang tidak seimbang, agar memiliki jumlah yang sama antar kategori status gizi, dengan tujuan untuk mendapatkan hasil pengelompokan yang lebih baik. Dataset disebut *balance* jika kategori klasifikasi direpresentasikan secara tepat dan sebaliknya disebut *imbalance*, jika kategori klasifikasi tidak direpresentasikan secara tepat. Sampling terdiri atas dua macam yaitu *oversampling* dan *undersampling*. *Oversampling* adalah proses menduplikasi data dari kelas minor agar diperoleh data yang *balance*, sedangkan *undersampling* adalah proses membuang sebagian data dari kelas mayor agar diperoleh data yang seimbang.

C. Normalisasi Data Balita

Setelah melewati proses *filterisasi*, data diolah ke tahap normalisasi untuk memperkecil besaran angka antar variabel dikarenakan sering terdapat besaran angka yang cukup jauh antara variabel tinggi badan dan berat badan [10]. Perbedaan jarak atau besaran angka yang cukup jauh nantinya dapat menyulitkan proses pengelompokan. Nilai variabel tinggi badan dan berat badan akan dinormalisasi ke dalam rentang 0 – 1. Adapun tahapan yang dilakukan untuk proses normalisasi adalah Mencari nilai maksimum dan minimum dari data *learning.csv* dengan jumlah 3695 data untuk variabel berat badan, tinggi badan, umur dan jenis kelamin. Setelah angka pada masing-masing variabel dilakukan normalisasi, langkah selanjutnya adalah menentukan jumlah kelompok atau cluster. Ke 3695 data balita yang ada akan dikelompokkan masing-masing status gizi balita ke dalam 4 cluster .

D. Proses Algoritma Agglomerative Hierarchical clustering

Setelah melalui proses normalisasi, maka data balita akan dikelompokkan menggunakan metode *agglomerative hierarchical clustering*. *Agglomerative hierarchical clustering* merupakan suatu metode *hierarchical clustering* yang bersifat *bottom-up* yaitu menggabungkan n buah kelompok menjadi satu kelompok tunggal [2]. Adapun proses algoritma AHC adalah sebagai berikut :



Gbr. 6 Alur Algoritma AHC

Langkah-langkah metode *agglomerative hierarchical clustering* adalah sebagai berikut [2] :

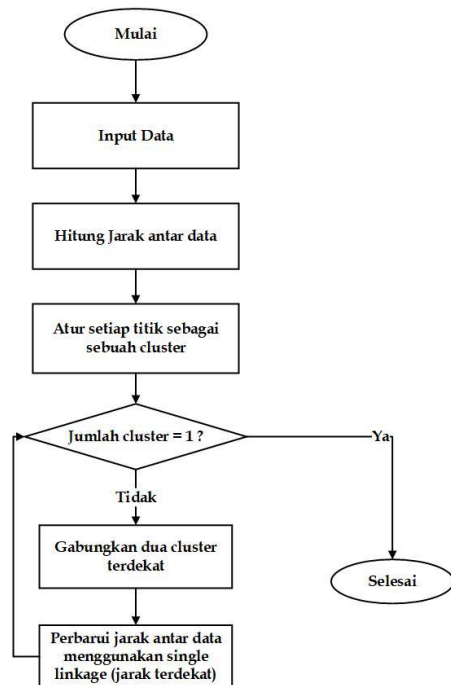
1. Mulai dengan N cluster, setiap cluster mengandung kesatuan yang tunggal dan sebuah matriks simetris $N \times N$ dari jarak (atau kemiripan) $D = \{dik\}$.
2. Cari matriks jarak untuk pasangan cluster yang terdekat. Misalkan jarak antara cluster U dan V yang paling mirip dinotasikan dengan d_{UV} .
3. Gabungkan cluster U dan V . Labeli cluster baru yang terbentuk dengan (UV) , kemudian perbarui entri-entri pada matriks jarak dengan cara :
 1. Menghapus baris-baris dan kolom-kolom yang bersesuaian dengan cluster U dan V .
 2. Menambahkan sebuah baris dan kolom yang memberikan jarak-jarak antara cluster (UV) dan cluster-cluster yang tersisa.
4. Ulangi langkah 2 dan 3 sebanyak $N-1$ kali. Semua objek akan berada dalam cluster tunggal setelah algoritma berakhir. Kemudian catat identitas dari *cluster* yang digabungkan dan level-levelnya (jarak atau kemiripan) dimana gabungannya ditempatkan.

Terdapat tiga teknik kedekatan yang sering digunakan dalam *agglomerative hierarchical clustering* yaitu *single linkage*, *average linkage* dan *complete linkage*. Penelitian ini mencoba untuk menggunakan pendekatan *single linkage* yaitu pengelompokan *agglomerative* yang didasarkan pada jarak minimum antar objek. Pengelompokan *single linkage*

awalnya dipilih jarak terkecil dalam $D = \{d_{ij}\}$ dan menggabungkan objek-objek yang bersesuaian misalnya U dan V untuk mendapatkan *cluster* (UV). Langkah berikutnya, jarak di antara (UV) dan *cluster* lainnya, misalnya W .

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

Dimana, d_{UW} = jarak antara tetangga terdekat dari *cluster* U dan W , dan d_{VW} = jarak antara tetangga terdekat dari *cluster* V [2]. Perhitungan dari *AHC single linkage* adalah sebagai berikut :



Gbr.7 Alur perhitungan AHC single linkage

E. Proses K-means Clustering

Perhitungan *k-means clustering* dimulai dengan mengambil centroid awal yang telah tersimpan di folder dataset dari hasil proses *agglomerative hierarchical clustering*. Setelah itu akan dihitung dengan metode *k-means* dengan langkah-langkah sebagai berikut [5]:

1. Mengambil centroid awal dari hasil perhitungan *agglomerative hierarchical clustering*.
2. Hitung jarak setiap objek ke masing-masing centroid dari masing-masing cluster menggunakan *Euclidian Distance* dengan rumus sebagai berikut ini :

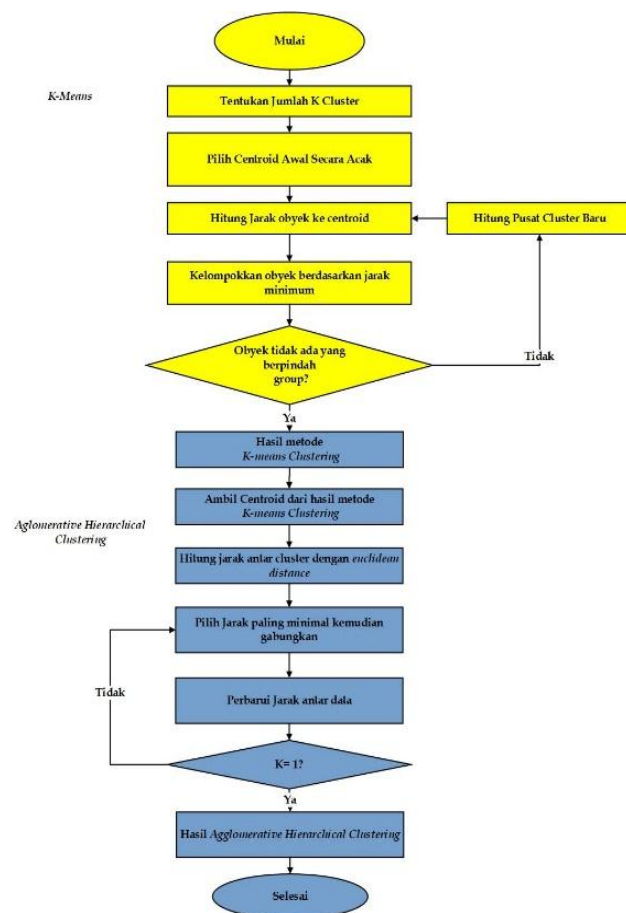
$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Keterangan :
 x_i : objek x ke-i
 y_i : daya y ke-i
 n : banyaknya objek

3. Alokasikan masing-masing objek ke dalam centroid yang paling dekat. Untuk melakukan pengalokasian objek kedalam masing-masing cluster pada saat iterasi secara umum dapat dilakukan dengan cara hard *k-means* dimana secara tegas setiap objek dinyatakan sebagai anggota cluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat cluster tersebut.
4. Lakukan iterasi, kemudian tentukan posisi centroid baru. Ulangi langkah 3 jika posisi centroid baru tidak sama

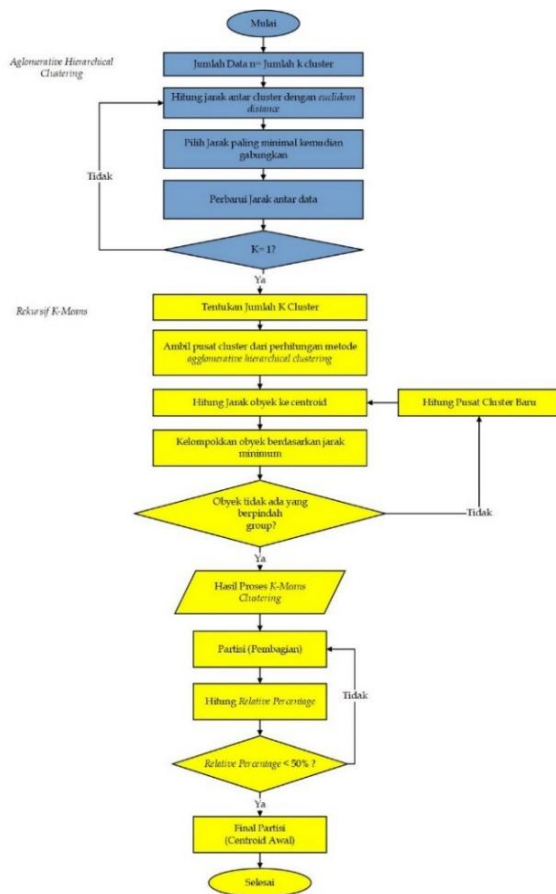
F. Proses KnA

Penelitian ini akan mencoba melakukan pengelompokan menggunakan metode KnA ke dalam tiga versi yaitu KnA menggunakan *agglomerative hierarchical clustering* ke *k-means* dengan alur perhitungan yang telah dijelaskan pada gambar 3. Kemudian KnA menggunakan *k-means* ke *agglomerative hierarchical clustering* dengan alur perhitungan sebagai berikut :



Gbr.8 Alur KnA menggunakan K-means ke AHC

Dan yang terakhir adalah KnA menggunakan *agglomerative hierarchical clustering* ke *Rekursif K-means* dengan alur perhitungan sebagai berikut :



Gbr.9 KnA menggunakan AHC ke rekursif K-means

Setelah dilakukan pengelompokan, maka akan diambil hasil akurasi terbaik yang akan digunakan untuk menentukan status gizi balita pada sistem. Masing-masing akurasi yang didapat dari pengelompokan adalah sebagai berikut :

1. Akurasi KnA menggunakan AHC ke *k-means* adalah sebagai berikut :

Status Gizi	KnA Single Linkage
BBU	86.90121786197564%
TBU	63.05818673883626%
BBTB	80.81190798376184%

2. Akurasi KnA menggunakan *k-means* ke AHC adalah sebagai berikut:

Status Gizi	KnA Single Linkage
BBU	0.00703 %
TBU	0.2830 %

BBTB	0.4516 %
------	----------

3. KnA versi *agglomerative hierarchical clustering* ke *k-means* (rekursif *k-means*)

Status Gizi	KnA	KnA normalisasi
BBU	92.8%	93.4%
TBU	94.9%	97.0%
BBTB	92.1%	92.8%

Dari pengelompokan yang telah dilakukan, maka akan diambil hasil terbaik dari masing-masing status gizi balita BB/U, TB/U dan BB/TB untuk sistem penentuan status gizi balita menggunakan metode KnA.

TABEL I
PENGELOMPOKAN STATUS GIZI OLEH AHLI GIZI

Jenkel	Umur	BB	TB	Status Gizi		
				BB/U	TB/U	BB/TB
2	27	9.5	79	BB Normal	Pendek	Normal
2	14	11.8	76	BB Normal	Normal	Gemuk
2	20	11	83	BB Normal	Normal	Normal
1	26	9.6	81	BB Krg	Pendek	Normal
1	35	22	94	BB Lbh	Normal	Gemuk

TABEL III
PENGELOMPOKAN STATUS GIZI OLEH METODE KNA

Jenkel	Umur	BB	TB	Status Gizi		
				BB/U	TB/U	BB/TB
2	27	9.5	79	BB Normal	Pendek	Normal
2	14	11.8	76	BB Normal	Normal	Gemuk
2	20	11	83	BB Normal	Normal	Normal
1	26	9.6	81	BB Krg	Pendek	Normal
1	35	22	94	BB Lbh	Normal	Gemuk

Berdasarkan tabel pengelompokan data balita yang dilakukan oleh ahli gizi di puskesmas menggunakan buku sk antropometri dengan tabel pengelompokan data balita yang dihasilkan oleh sistem menggunakan metode KnA hasilnya adalah valid. Hal ini menunjukkan bahwa sistem penentuan status gizi balita menggunakan metode KnA berjalan dengan baik dan data yang dikeluarkan sudah sesuai dengan yang diharapkan.

IV. KESIMPULAN

Berdasarkan hasil percobaan mengenai penentuan status gizi balita menggunakan metode KnA dengan parameter data balita berupa umur, jenis kelamin, BB (berat badan) dan TB (tinggi badan), metode KnA (kombinasi *k-means* dan *agglomerative hierarchical clustering*) mampu dengan baik mengelompokkan data balita dalam jumlah yang besar. Dengan data learning sebanyak 3695 data balita menghasilkan masing-masing hasil dengan nilai terendah adalah metode KnA menggunakan *k-means clustering* ke metode *agglomerative hierarchical clustering* dengan nilai akurasi BBU = 0,07%, TBU = 0,28% dan BBTB = 0,45%. Sedangkan untuk KnA menggunakan *agglomerative hierarchical clustering* ke *k-means* menghasilkan akurasi BBU = 87%, TBU = 63% dan BBTB = 80%. Dan yang terakhir untuk KnA menggunakan *agglomerative hierarchical clustering* ke *rekursif k-means* menghasilkan akurasi sebesar BBU = 94%, TBU = 98%, dan BBTB = 93%. Nilai-nilai ini bisa berubah seiring dengan penambahan data latih. Selanjutnya hasil terbaik dari pengelompokan, akan digunakan untuk sistem penentuan status gizi balita menggunakan metode KnA.

REFERENSI

- [1] Agustina, Mitakda, dan Solimun. 2013. "Pemilihan Metode Pengelompokan Terbaik Kabupaten/Kota Berdasarkan Indikator Pendidikan Menggunakan Hybrid Melalui Mutual Cluster, Bottom-up, dan Top-down", Jurnal Mahasiswa Statistik Universitas Brawijaya-Malang Vol. 1 (3): hal. 205-208.
- [2] Alfi, Fadliana dan Fachrur, Rozi. 2015. "Penerapan Metode Agglomerative Hierarchical Clustering Untuk Klasifikasi Kabupaten/Kota Di Provinsi Jawa Timur Berdasarkan Kualitas Pelayanan Keluarga Berencana". Jurnal Matematika Murni dan Aplikasi. Vol.(4): hal. 36-40.
- [3] Wulandari, Dewi dan Prasetyo, Arfhan. 2018. "Sistem Penunjang Keputusan Untuk Menentukan Status Gizi Balita Menggunakan metode Fuzzy Tsukamoto. Jurnal Informatika. Vol.(5): hal. 22-33.
- [4] Astrid, Alfira, dkk. 2016. "Analisis Hybrid Mutual Clustering menggunakan Jarak Square Euclidean". Jurnal Matematika dan Terapan. Vol.(2): hal. 15-21.
- [5] Athman, Bouguettaya, dkk. 2014. "Jurnal Efficient Agglomerative Hierarchical Clustering". Journal ScienceDirect. Vol.(42): hal. 1-13.
- [6] Bikriyah. 2014. "Analisis Hybrid Hierarchical Clustering Melalui Mutual Cluster, Bottomup, dan Top-down Menggunakan Jarak Euclidean dan Mahalanobis", Jurnal Mahasiswa Statistik Universitas Brawijaya-Malang Vol. 2 (5): hal. 397-400.
- [7] Febrealiti, Eka Rahmanurul. 2010. *Sistem Penentuan Status Gizi Balita Menggunakan Metode K-NN (K-NEAREST NEIGHBOR)*. Jurusan Teknik Informatika Fakultas Sains Dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau Pekanbaru. Skripsi tidak diterbitkan. Riau : PPs Universitas Islam Negeri Sultan Syarif Kasim.
- [8] Sudirman, dkk. 2013. "Analisa Klasifikasi status gizi dengan metode fuzzy K-means menggunakan aplikasi berbasis android". Jurnal Informatika. hal 1-9.
- [9] Kaesmetan, Yampi dan Johannis, Jusrianto . 2017. "Jurnal Klasifikasi Status Gizi Balita Di Kelurahan Oesapa Barat menggunakan Metode K-Nearest Neighbor". Jurnal Ilmiah Multitekt Indonesia. Vol(11): hal. 42-50.
- [10] Pradnya, Dhuhiha. 2015. "Jurnal Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita". Jurnal Informatika. Vol.(15): hal. 160-174.
- [11] Prihastari, Lisa. 2018. Malnutrisi. (online) (<https://www.academia.edu/9858871/Malnutrisi>).
- [12] Hasan, Muhammad. 2018. "Sistem Pendukung Keputusan Penentuan Status Gizi Balita Menggunakan Metode Naive Bayes". Seminar Nasional Teknologi Informasi dan Multimedia 2018. ISSN : 2302-3805, 25-30.