

KLASTERISASI PASIEN STUNTING MENGGUNAKAN ALGORITHMMA K-MEANS

Ulumuddin ¹, hasbi firmansyah ²

Universitas Bhamada Slawi, Universitas Panca Sakti Tegal

E-mail: ulumudinboy@gmail.com¹, hasbifirmansyah@upstegal.ac.id²

ABSTRAK

Stunting merupakan gangguan pertumbuhan linier pada bayi baru lahir yang di sebabkan oleh beberapa factor yaitu antara lain factor BBLR berat badan lahir rendah, pendidikan ibu, pendapatan rumah tangga, dan lain sebagainya. Sehingga diperlukannya pengklasteran pada kasus ini agar mengetahui stunting dengan cluster ringan, berat ataupun sedang. Pada penelitian ini peneliti menggunakan algorithmma k-means dalam melakukan pengklasteran. Sehingga mendapatkan hasil cluster berat sebanyak 41, cluster sedang sebanyak 109 sedangkan cluster ringan sebanyak 50

Kata Kunci : Stunting, algorithmma K-Means

LATAR BELAKANG

Stunting adalah gangguan pertumbuhna linier yang di sebabkan kurang gizi pada balita sehingga berefek pada gangguan pertumbuhan pada anak balita terutama pada pertumbuhan tinggi badan. Stunting juga menjadi penyebab utama anak – anak balita pada meninggal pada seluruh duina hingga 2,2 juta jiwa.

Dalam kasus ini peneliti melakukan penelitian dengan mengklaster pada pasien yang terkena stunting, Dalam kasus ini peneliti menggunakan algorithmma k-means untuk dilakukannya klasterisasi pada jenis kelamin pasien stunting.

Menurut Han dan Kamber secara umum metode klastering dapat diklasifikasikan menjadi empat kategori dasar yaitu metode partisi, metode hirarki, metode berbasis jarak, dan metode berbasis *grid*. Metode klastering yang paling populer adalah metode hirarki dan metode partisi metode klastering adalah proses pengelompokan sekumpulan objek data ke dalam beberapa kelompok atau klaster sehingga data dalam sebuah klaster memiliki kemiripan yang maksimum, dan data antar klaster memiliki tingkat kemiripan yang minimum Aplikasi pengklasteran dipakai di berbagai bidang, seperti data mining, pengenalan pola, klasifikasi citra, ilmu-ilmu yang terkait dengan biologi, pemasaran, perencanaan kota, *document retrieval*, dan lain-lain.

K-means merupakan sebuah metode penganalisaan data atau salah satu metode data mining yang merupakan proses permodelan tanpa supervisi (*unsupervised*) serta merupakan salah satu metode yang melakukan permodelan data dengan mengelompokan data naik skala kecil atau besar dengan sistem partisi

TINJAUAN PUSTAKA

- a. Gladys Apriluana* dan Sandra Fikawati 2018 *Analisis Faktor-Faktor Risiko terhadap Kejadian Stunting pada Balita (0-59 Bulan) di Negara Berkembang dan Asia Tenggara* Stunting adalah gangguan pertumbuhan linier yang disebabkan kurang gizi yang berlangsung kronis. Keadaan gizi balita pendek menjadi penyebab 2,2 juta dari seluruh

penyebab kematian balita di seluruh dunia. Tujuan dari review literatur ini adalah menganalisa efek dari faktor-faktor risiko determinan terhadap kejadian stunting pada balita. Desain penelitian ini adalah literature review. Artikel-artikel yang dipilih dengan search engine adalah artikel correlation research yang menggunakan study cross-sectional dengan respondennya adalah anak dengan stunting usia 0-59 bulan. Kriteria inklusi artikel yang dipilih adalah anak dengan stunting, berusia 0-59 bulan, wilayah negara berkembang (termasuk wilayah Asia Tenggara), memiliki KMS, masih memiliki orang tua lengkap. Proses pencarian hingga pengeksklusian artikel-artikel yang digunakan untuk review literatur ini menggunakan metode PRISMA. Hasil penelitian menunjukkan faktor status gizi dengan berat badan lahir < 2.500 gram memiliki pengaruh secara bermakna terhadap kejadian stunting pada anak dan memiliki risiko mengalami stunting sebesar 3,82 kali. Faktor pendidikan ibu rendah memiliki pengaruh secara bermakna terhadap kejadian stunting pada anak dan memiliki risiko mengalami stunting sebanyak 1,67 kali. Faktor pendapatan rumah tangga yang rendah diidentifikasi sebagai predictor signifikan untuk stunting pada balita sebesar 2,1 kali. Faktor sanitasi yang tidak baik memiliki pengaruh yang signifikan terhadap kejadian stunting pada balita dan memiliki risiko mengalami stunting hingga sebesar 5,0 kali. Kesimpulan penelitian ini adalah semakin rendahnya berat badan lahir (BBLR), tingkat pendidikan ibu, pendapatan rumah tangga, dan kurangnya hygiene sanitasi rumah maka risiko balita menjadi stunting semakin besar.

b. Uliyatul Laili dan Ratna Ariesta Dwi Andriani (2019) *pemberdayaan masyarakat terhadap stunting*

Pada tahun 2017 pemerintah telah meluncurkan program Rencana Aksi Nasional Penanganan stunting pada tingkat nasional, daerah terutama desa. Salah satu bagian dari program tersebut adalah pencegahan stunting yang terjadi di masyarakat. Karena sebagian besar masyarakat belum paham dengan benar mengenai stunting, dan beranggapan bahwa stunting atau kerdil sebutan yang biasa digunakan di masyarakat adalah faktor keturunan. Kegiatan pengabdian pada masyarakat ini dilakukan untuk meningkatkan pengetahuan dan pemahaman serta peran serta masyarakat dalam program pencegahan dan deteksi dini stunting pada balita yang diharapkan secara langsung dapat memotivasi masyarakat untuk ikut serta memperhatikan pertumbuhan dan perkembangan pada anaknya sehingga pertumbuhan dan perkembangannya dapat optimal. Metode yang dilakukan adalah menilai pengetahuan masyarakat serta permasalahan yang dihadapi tentang pencegahan dan cara menilai/ deteksi dini stunting pada balita di RW 2 Kelurahan Wonokromo Kecamatan Wonokromo Surabaya. Pengetahuan ibu diukur dengan menggunakan menggunakan pre test sebelum kegiatan dan post test setelah diberikan pengetahuan. Berdasarkan hasil pretest dan post test yang diikuti oleh 35 responden dapat dinyatakan bahwa hasil pre test tentang tingkat pengetahuan responden mengenai program pencegahan stunting sebesar 14 responden (40%) mengerti tentang program pencegahan stunting sedangkan berdasarkan hasil post test terdapat 27 responden (77,1%) yang mengerti tentang program pencegahan stunting.

c. Awa Ramdhani, Hani Handayani, Asep Setiawan (2020) *Hubungan Pengetahuan Ibu Dengan Stunting;*

Stunting adalah suatu kondisi kekurangan gizi kronis yang terjadi pada saat periode kritis dari proses tumbuh dan kembang mulai janin. Untuk Indonesia, saat ini diperkirakan ada 37,2% dari anak usia 0-59 bulan atau sekitar 9 juta anak dengan kondisi stunting, yang berlanjut sampai usia sekolah 6-18 tahun. Stunting didefinisikan

sebagai kondisi anak usia 0 – 59 bulan, dimana tinggi badan menurut umur berada di bawah minus 2 Standar Deviasi (<-2SD) dari standar median WHO. Stunting akan berdampak dan dikaitkan dengan proses kembang otak yang terganggu, dimana dalam jangka pendek berpengaruh pada kemampuan kognitif. Jangka panjang mengurangi kapasitas untuk berpendidikan lebih baik dan hilangnya kesempatan untuk peluang kerja dengan pendapatan lebih baik (Risksedas, Prevalensi Stunting, 2018). Salah satu penyebab stunting adalah pola asuh ibu terhadap balitanya. Pola asuh erat kaitannya dengan tingkat pengetahuan ibu. Pengetahuan yang kurang dapat menjadikan pola asuh ibu kurang sehingga memengaruhi kejadian stunting pada balita. Tujuan penelitian ini untuk mengetahui hubungan pengetahuan ibu dengan kejadian stunting pada balita. Desain penelitian ini menggunakan metode literature review. Penelusuran artikel dilakukan dengan menggunakan search engine google scholar dan diperoleh data jurnal sebanyak 89 jurnal yang terdiri dari 5 judul dari proquest dan 84 judul dari google scholar. Di dapatkan sebanyak 10 jurnal full text yang sesuai kriteria inklusi, yaitu data berasal dari artikel jurnal nasional dan merupakan hasil penelitian primer, berbahasa Indonesia, serta terbit dalam 10 tahun terakhir. Berdasarkan hasil analisis dari jurnal menunjukkan tingkat pengetahuan ibu tentang stunting masih kurang, dan terdapat hubungan antara pengetahuan ibu dengan kejadian stunting. Penyebab kurangnya pengetahuan ibu tentang Stunting adalah karena tidak semua ibu balita melakukan kunjungan ke Posyandu. Disarankan untuk tenaga Kesehatan dapat memberikan pendidikan Kesehatan dengan metode yang berbeda untuk meningkatkan pengetahuan ibu tentang stunting

- d. Nielza Atthina 2014, Klasterisasi Data Kesehatan Penduduk Untuk Menentukan Rentang Drajat Kesehatan Daerah Dengan Metode K-Means

Salah satu indikator untuk menilai tingkat keberhasilan pembangunan negara adalah tingkat capaian negara dalam memberikan jaminan bidang kesehatan untuk masyarakatnya. Pemerintah Indonesia, melalui Departemen Kesehatan, menetapkan sejumlah indikator sebagai tolak ukur kemajuan pembangunan kesehatan. Setiap tahun Departemen Kesehatan mengumpulkan data kesehatan penduduk di berbagai level unit daerah untuk diolah sehingga dihasilkan ranking Provinsi dan Kabupaten/Kota Sehat. Metode pengolahan terhadap data kesehatan yang sudah terkumpul tersebut masih berbasis pada teknik statistik dasar, dimana memiliki kelemahan dalam hal konsistensi data dan minimnya informasi tentang hubungan antar data. Data mining dilihat dari sisi teknik pengolahan data menyediakan sejumlah algoritma yang dapat digunakan untuk menggali informasi tersembunyi dari kumpulan data yang multidimensi. Penelitian ini membahas tentang pemanfaatan algoritma k-Means untuk mengklaster atau mengelompokkan kabupaten-kabupaten di Provinsi Jawa Tengah dan DIY berdasarkan kemiripan nilai lima buah ukuran indikator mortalitas derajat kesehatan daerah, yaitu angka kelahiran kasar, angka kematian kasar, angka kematian bayi, angka kematian balita, dan angka kematian ibu. Secara umum, output penelitian berupa hasil klasterisasi dapat digunakan sebagai sebuah acuan untuk menggambarkan distribusi pengelompokkan kabupaten berdasarkan kondisi profil kesehatan daerah tersebut

METODE PENELITIAN

Desain Penelitian

Dalam penelitian ini peneliti menggunakan metode algoritma k-means guna mengkluster jenis kelamin pada pasien stunting. Adapun alur berjalannya penelitian yang diusulkan adalah sebagai berikut.

- a. Mengumpulkan data di puskesmas terdekat
- b. Melakukan pengolahan data awal
- c. Menerapkan desain penelitian yang akan diusulkan
- d. Eksperimen metode yang diusulkan
- e. Evaluasi dan validasi data yang akan di hasilkan

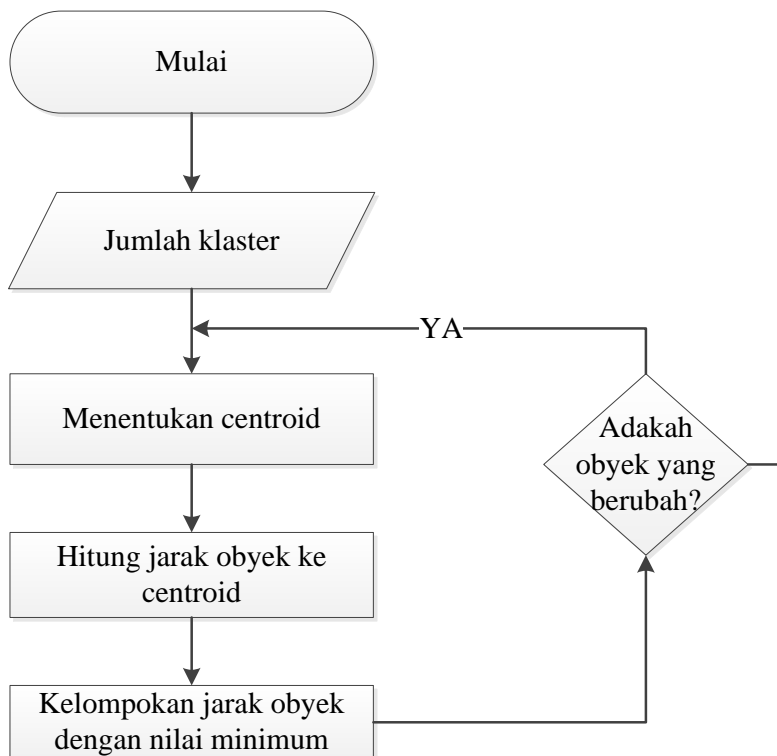
Pengumpulan Data

Dalam penelitian ini data yang digunakan adalah data yang didapatkan puskesmas desa jatibogor kecamatan suradadi kabupaten tegal dan data ini bersifat private. Di dalam data ini ada beberapa atribut yang peneliti gunakan guna untuk mengelompokkan pasien stunting yang ada di puskesmas desa jatibogor kecamatan suradadi kabupaten tegal. Adapun atribut – atribut yang peneliti gunakan adalah sebagi berikut

No	Atribut
1	BBLR Berat Badan Lahir Rendah
2	Pendidikan Ibu
3	Pendapatan rumah tangga
5	Jenis Kelamin

Metode yang diusulkan

Dalam penelitian ini peneliti menggunakan metode algoritma k-means untuk mengelompokkan pasien stunting. Berikut adalah alur kerja dari algoritma k-means.



Pada gambar diatas menjelaskan alur kerja k-means. Berikut penjelasan pada alur kerja k-means.

1. Tentukan banyaknya kluster, pada penelitian ini peneliti menggunakan 3 kluster ($k=3$). Banyaknya kluster harus lebih kecil dari pada banyaknya data. Adapun label yang dijadikan sebagai cluster yaitu seperti yang ada pada tabel dibawah ini,

tabel 3.2. label cluster

Cluster 1	Berat
Cluster 2	Ringan
Cluster 3	Sedang

2. Tentukan centroid pada setiap cluster. Untuk menentukan centroid pada kluster. Banyak metode yang bisa digunakan salah satunya adalah dengan cara radom / acak yang diambil dari dataset yang digunakan. Pada tabel

dibawah ini merupakan contohnya yang digunakan sebagai titik centroid yang diambil dari 200 dataset.

Tabel 3.3. Titik Centroid

BBLR	PENDIDIKAN IBU	PENDAPATAN RUMAH TANGGA	JENIS KELAMIN
66	32	117	102
80	35	117	100
66	39	118	101

3. Menghitung jarak data dengan centroid, untuk menghitung jarak banyak rumus bisa yang digunakan, ada euclidean, manhatan / city block dan minkowski. Namun disini peneliti menggunakan euclidean untuk menghitung jarak entroid. Berikut rumus dan penjelasannya,

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}.$$

d = jarak j = banyaknya data c = centroid x = data c = centroid

jarak data dengan cluster 1, 2 dan 3 pada implementasi penelitian ini sebagai berikut,

cluster 1

$$d(x_1, C_1) = \sqrt{(66 - 80)^2 + (32 - 35)^2 + (117 - 117)^2 + (102 - 100)^2} = 0,0$$

$$d(x_1, C_1) = \sqrt{(66 - 75)^2 + (32 - 37)^2 + (117 - 117)^2 + (102 - 104)^2} = 2,6$$

$$d(x_1, C_1) = \sqrt{(66 - 70)^2 + (32 - 33)^2 + (117 - 117)^2 + (102 - 100)^2} = 5,2$$

Cluster 2

$$d(x_1, C_1) = \sqrt{(80 - 65)^2 + (35 - 30)^2 + (117 - 118)^2 + (100 - 103)^2} = 8,6$$

$$d(x_1, C_1) = \sqrt{(80 - 66)^2 + (35 - 36)^2 + (117 - 117)^2 + (100 - 100)^2} = 6,7$$

$$d(x_1, C_1) = \sqrt{(80 - 78)^2 + (35 - 35)^2 + (117 - 117)^2 + (100 - 107)^2} = 8,6$$

Cluster 3

$$d(x_1, C_1) = \sqrt{(66 - 75)^2 + (39 - 35)^2 + (118 - 117)^2 + (101 - 107)^2} = 3,3$$

$$d(x_1, C_1) = \sqrt{(66 - 69)^2 + (39 - 32)^2 + (118 - 117)^2 + (101 - 100)^2} = 8,3$$

$$d(x_1, C_1) = \sqrt{(66 - 82)^2 + (39 - 25)^2 + (118 - 117)^2 + (101 - 104)^2} = 6,1$$

4. Kelompokan data sesuai dengan clusternya, yaitu data yang memiliki jarak terdekat, seperti tabel di bawah ini ;

CENTROID 1			CENTROID 2				CENTROID 3				C1	C2	C3
32.0	117.0	102.0	80	35	117	100	66.0	39.0	118.0	101.0			
0.00			14.46				7.14				YES		
15.00			8.60				17.72					YES	
2.65			16.12				9.27				YES		
10.49			6.71				9.75					YES	
4.47			14.04				3.32						YES
10.72			8.60				11.58					YES	
13.34			9.64				17.46					YES	
5.20			13.19				9.90				YES		
4.58			10.20				7.35				YES		
13.23			3.46				14.07					YES	
16.12			9.33				19.00					YES	
8.77			12.96				8.37						YES

5. Jika nilai centroid pada klaster masih ada yang berubah maka ulangi langkah 3 dan 4 sampai tidak ada perubahan.

HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Dalam penelitian ini peneliti mendapatkan data dari pusat kesehatan masyarakat di desa jatibogor kecamatan suradadi kabupaten tegal dengan data yang ada sebanyak 200 data pasien stunting.

2. Implementasi algoritma k-means

Dalam penerapan algoritma k-means pada penelitian ini ada langkah – langkah yang peneliti lakukan, berikut langkah – langkah yang peneliti lakukan beserta dengan implementasinya dari rumus dibawah ini

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

1. Tentukan nilai K pada jumlah cluster yang akan dibentuk,
2. Tentukan nilai titik pusat cluster dengan cara dirandom/acak,
3. Hitung jarak dari masing – masing centroid dari nilai cluster, untuk menghitung jarak antara objek dengan centroid bisa menggunakan Euclidian Distance,
4. Alokasikan masing – masing objek pada nilai centroid yang paling dekat,
5. Lakukan iterasi kemudian tentukan posisi titik centroid baru dengan persamaan,
6. Ulangi langkah 3 bilamana posisi centroid baru tidak sama

Di bawah ini implementasi serta penjelasan dari langkah langkah k-means diatas,

1. Tentukan nilai K pada jumlah cluster yang akan dibentuk

Dalam penelitian ini peneliti menggunakan 200 data kejahatan. Dari 200 data tersebut peneliti menggunakan 3 cluster untuk pemetaan kejahatan.

tabel 4.1. Penentuan Jumlah K

No	lokasi	umur	j_kelamin	J_kejahatan	cluster
1	66	32	117	102	ringan
2	80	30	117	107	
3	65	30	118	103	
4	75	37	117	104	sedang
5	66	36	117	100	
6	75	35	117	107	
7	78	27	117	105	berat
8	69	31	118	106	
9	70	33	117	100	
10	79	34	118	103	

Keterangan dalam tabel diatas ada 4 atribut , yaitu lokasi, umur, jenis kelamin dan jenis kejahatan, sedangkan untuk labelnya ada 3 yaitu ringan, sedang dan berat. Label digunakan sebagai cluster untuk pengelompokkan tindak pidana kejahatan.

2. Tentukan nilai titik pusat cluster dengan cara dirandom atau acak, Langkah kedua adalah dengan merandom atau mengacak dengan tujuan untuk dijadikan sebagai nilai titik cluster, misalkan dari keseluruhan data diatas diambil atau diacak nilai yang mau dijadikan sebagai titik pusat cluster. Disini peneliti mengambil nilai – nilai yang ada dalam tabel dibawah ini.

tabel 4.2. Titik Cluster

No	lokasi	umur	j_kelamin	J_kejahatan
1	66	32	117	102

3. Hitung jarak dari masing – masing centroid dari nilai cluster, untuk menghitung jarak antara objek dengan centroid bisa menggunakan Euclidian Distance. Euclidian distance merupakan rumus perhitungan jarak antar obyek di bawah ini merupakan rumusnya.

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}.$$

Implementasi dari rumus diatas adalah sebagai berikut

$$d(x_1, C_1) = \sqrt{(66 - 80)^2 + (32 - 35)^2 + (117 - 117)^2 + (102 - 100)^2} = 0.0$$

$$d(x_1, C_2) = \sqrt{(88 - 66)^2 + (39 - 43)^2 + (118 - 117)^2 + (101 - 112)^2} = 8.6$$

tabel 4.3. Jarak Titik Centroid

no	jarak titik centroid			cluster
	centroid 1	centroid 2	centroid 3	
1	0,00	14,46	7,14	1
2	15,00	8,60	17,72	2
3	2,65	16,12	9,27	1
4	10,49	6,71	9,75	2
5	4,49	14,04	3,32	3
6	10,72	8,60	11,58	2
7	13,34	9,64	17,46	2
8	5,20	13,19	9,90	1
9	4,58	10,20	7,35	1
10	13,23	3,46	14,07	2

4. Alokasikan masing masing objek pada nilai centroid paling dekat

no	centroid	cluster	objek / data			
1	0,00	1	66	32	117	102
2	8.6	2	80	30	117	107
3	2,65	1	65	30	118	103
4	6,71	2	75	37	117	104
5	3,32	3	66	36	117	100
6	8,60	2	75	35	117	107
7	9,64	2	78	27	117	105
8	5,20	1	69	31	118	106
9	4,58	1	70	33	117	100
10	3,46	2	79	34	118	103

5. Lakukan iterasi kemudian tentukan posisi titik centroid baru dengan persamaan, yaitu dengan menerapkan rumus persamaan cambera diamana terlebih dahulu menentukan titik centroid.

tabel 4.4. Iterasi Data

no	objek / data				cluster
1	66	32	117	102	1
2	80	30	117	107	2
3	65	30	118	103	1

4	75	37	117	104	2
5	66	36	117	100	3
6	75	35	117	107	2
7	78	27	117	105	2
8	69	31	118	106	1
9	70	33	117	100	1
10	79	34	118	103	2

6. Jika masih ada anggota klaster yang berubah lakukanlah langkah 3 dan 4 sampai anggota klaster tidak ada yang berubah

KESIMPULAN

Berdasarkan dari hasil uji coba yang peneliti lakukan adalah mendapatkan hasil sebagai berikut ;

- label 1 sebagai cluster 1 yang mempunyai nilai 41, dimana label 1 ini merupakan cluster dari stunting berat
- label 2 sebagai cluster 2 yang mempunyai nilai 109, dimana label 2 ini merupakan cluster dari stunting sedang
- label 3 sebagai cluster 3 yang mempunyai nilai 50, dimana label 3 ini merupakan cluster dari stunting ringan

DAFTAR PUSTAKA

- Gladys Apriliana, (2018). *Analisis faktot – factor resiko terhadap kejadian stunting pada balita 0 – 59 bulan di Negara berkembang dan asia tenggara . universitas Indonesia depok .*
- Uliyatul laeli 2019. *Pemberdayaan masyarakat terhadap stunting. Universitas nahdhlotul ulama Surabaya*
- Awa ramdhani (2020) *hubungan pengetahuan ibu dengan stunting. Universitas Muhammadiyah tasikmalaya.*
- Edwin Daniel osie (2018). *“hubungan antara sikap dan pengetahuan ibu terhadap kejadian stunting pada nak baru masuk sekolah dasar di kecamatan nanggolo”*
- Han, Jiawet, Kamber Micheline, and Jian Pei. 2012. *Data Mining Concepts and Techniques.*
- Hartigan, J. a., and M. a. Wong. 1979. “Algorithm AS 136: A K-Means Clustering Algorithm.” *Journal of the Royal Statistical Society C* 28(1): 100–108.