

# Model Penentuan Potensi Status Gizi Bermasalah Menggunakan *Decision Tree*

Liliana Swastina, Bambang Lareno  
Teknik Informatika  
STMIK Indonesia Banjarmasin  
Banjarmasin, Indonesia  
liliser@gmail.com  
blareno@gmail.com

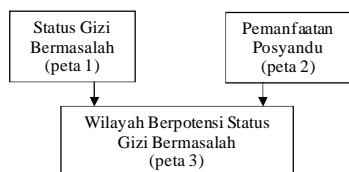
**Abstrak** - Perubahan berat badan balita dari waktu ke waktu merupakan petunjuk awal perubahan status gizi balita. Dalam periode 6 bulan, bayi yang berat badannya tidak naik 2 kali berisiko mengalami gizi kurang 12,6 kali dibandingkan pada balita yang berat badannya naik terus. Dengan demikian perlu informasi mengenai wilayah mana saja yang memiliki kurang gizi, sebagai bahan masukan bagi pemerintah dan penentu kebijakan di bidang kesehatan masyarakat dalam melakukan pencegahan dan penanggulangan gizi buruk. Metode yang diusulkan adalah metode eksperimen untuk menguji algoritma decision tree. Algoritma decision tree yang akan diuji adalah C4.5. Algoritma akan implementasikan dengan menggunakan RapidMiner 5.2. Hasil untuk wilayah berpotensi status gizi bermasalah: Rendah, Sedang, Tinggi. Perhitungan yang berdasarkan pada nilai kecenderungan (P1, P2, P3, P4 dan P5) menghasilkan akurasi 90,9 %. Perhitungan bertahap yaitu menghasilkan peta 3 berdasar hasil peta 1 dan peta 2, hanya mencapai akurasi 54, 5 %. Dengan demikian, untuk menghasilkan peta sebaran potensi gizi bermasalah (peta3), lebih baik jika didasarkan langsung pada data lima nilai kecenderungan.

**Kata kunci**—decision tree, algoritma C4.5, gizi balita.

## A. PENDAHULUAN

Kondisi kesehatan dan status gizi balita merupakan salah satu tolok ukur cerminan keadaan gizi masyarakat secara luas. Kasus gizi buruk tidak hanya menjadi beban keluarga tetapi juga menjadi beban Negara. Kesehatan dan status gizi balita merupakan salah satu tolok ukur yang dapat mencerminkan keadaan gizi masyarakat luas. Pola pengasuhan anak di masyarakat pada umumnya lebih mengutamakan anak balita. Gizi kurang pada balita tidak terjadi secara tiba-tiba, tetapi diawali dengan keterbatasan kenaikan berat badan yang tidak cukup. Perubahan berat badan balita dari waktu ke waktu merupakan petunjuk awal perubahan status gizi balita. Dalam periode 6 bulan, bayi yang berat badannya tidak naik 2 kali berisiko mengalami gizi kurang 12,6 kali dibandingkan pada balita yang berat badannya naik terus [1]. Dengan demikian perlu informasi mengenai wilayah mana saja yang memiliki kurang gizi, sebagai bahan masukan bagi pemerintah dan penentu kebijakan di bidang kesehatan masyarakat dalam melakukan pencegahan dan penanggulangan gizi buruk. Novianti, dkk [2] memetakan sebaran dengan analisa spasial terhadap data Riskesda 2007 menggunakan GIS. Mereka menghasilkan peta wilayah yang berpotensi rawan kurang gizi. Namun proses perhitungan dilakukan terpisah dan mengandalkan data yang telah matang. Yose Rizal [3], mengidentifikasi faktor-faktor risiko penderita gizi kurang dan gizi buruk dengan pemahaman tentang

distribusi spasial untuk mengidentifikasi pengelompokan (Clustering) kerawanan kurang gizi, berkaitan dengan topografi wilayah dan produktivitas lahan pertanian. Namun Rizal lebih fokus ke kaitan antar faktor – faktor risiko berkaitan dengan karakteristik keluarga dan distribusi spasial tempat tinggal keluarga yang menderita gizi kurang dan gizi buruk. Sementara Zhiwu Liu [4], menggunakan aturan klasifikasi dan model prediksi untuk menandai siswa, menggambarkan bagaimanakah hasil analisis prediksi dapat digunakan untuk mengetahui faktor yang dapat mempengaruhi siswa sehingga beberapa kebiasaan atau perilaku belajar siswa yang negatif dapat diketahui dan diperbaiki tepat pada waktunya. Oded Maimon menggunakan Algoritma C4.5 untuk menganalisis penjualan dan sebarannya [5]. Kemampuan klasifikasi Algoritma C4.5 dapat digunakan untuk membangun pohon keputusan guna pemberian status potensi bagi wilayah tertentu, sehingga terlihat kecenderungan status gizi balita dan dapat ditindaklanjuti pemangku kepentingan sebagaimana seharusnya. Dengan demikian, dapat dikatakan bahwa informasi wilayah sebaran status gizi balita penting bagi pihak terkait, baik secara data maupun label wilayah. Model yang ingin dibangun adalah sebagaimana gambar 1.



Gambar 1 Fakta yang berpotensi mempengaruhi tingkat gizi balita

Penelitian ini berfokus pada bagaimana membuat sebuah model yang dapat digunakan untuk pemberian potensi status gizi bermasalah (peta1) dan tingkat pemanfaatan pos yandu (peta 2). Pemberian Status Gizi dan tingkat pemanfaatan berdasarkan data dari pos yandu (penyedia data lapis terbawah). Keduanya penting guna membangun peta daerah rawan gizi (peta 3).

**B. METODE PENELITIAN**

Penelitian ini memakai metode penelitian eksperimen, yang terdiri: (1)Metode Pengumpulan data dan pengolahan data awal, (2)Metode yang diusulkan, (3)Eksperimen dan pengujian model, (4)Hasil eksperimen dan (5)Evaluasi dan validasi hasil.

**B.1 Metode Pengumpulan data dan pengolahan data awal**

Penelitian ini memakai data SKDN 2013, yang didapatkan dari Posyandu Purnama, Desa Pasarlama, Kecamatan Banjarmasin Tengah, Banjarmasin. Data SKDN posyandu lain dalam wilayah kerja puskesmas yang sama didapatkan dari Puskesmas S.Parman. Data lainnya, didapatkan dari Dinas Kesehatan Kotamadya Banjarmasin [7]. Data SKDN dalam bentuk empat kolom angka yaitu S (Jumlah balita dalam wilayah kerja), K (Jumlah balita yang memiliki Kartu Menuju Sehat), D (Jumlah balita ditimbang), N (Jumlah balita yang ditimbang 2 bulan berturut-turut dan garis pertumbuhannya pada KMS naik tapi tidak pindah ke jalur warna di bawahnya), dan angka BGM, sedangkan hasil pengolahan dalam bentuk proporsi misalnya N/D, dst. Pengolahan data ini untuk analisis kecenderungan:

1. Tingkat Liputan Program, yaitu proporsi Jumlah balita yang memiliki KMS (K) dibagi Jumlah balita dalam wilayah kerja (S). Dari data ini didapat angka Kehilangan Kesempatan (P1) yaitu Jumlah balita dalam wilayah kerja (S) dikurangi Jumlah balita yang memiliki KMS(K) kemudian dibagi S, atau proporsi: (S-K)/S
2. Tingkat Partisipasi Masyarakat dalam penimbangan balita, yaitu proporsi Jumlah balita ditimbang (D) dibagi Jumlah balita dalam wilayah kerja (S). Dari data ini didapat angka Pemanfaatan Posyandu (P2): (S-D)/S
3. Tingkat Gizi, yaitu proporsi jumlah balita yang naik beratnya (N) dibagi jumlah balita yang

ditimbang (D), sehingga yang bermasalah (P3) didapat dari: (D-N)/D

4. Tingkat Dropout (P4), yaitu proporsi jumlah balita memiliki KMS (K) dikurangi jumlah balita yang ditimbang (D) kemudian hasilnya dibagi dengan jumlah balita yang memiliki KMS. Proporsi didapat dari: (K-D)/K

5. Tingkat Bawah Garis Merah (P5), yaitu jumlah balita yang gizinya berada di BGM dibagi Jumlah balita ditimbang. Proporsi didapat dari BGM/D

Contoh data posyandu dan pengolahannya, dapat dilihat pada tabel 1.

Tabel 1 Data SKDN Dan Nilai Kecenderungan Posyandu

Posyandu Purnama	2013		
	Apr	Mei	Jun
S	50	50	50
K	48	50	50
D	43	31	29
N	20	12	13
BGM	0	2	0
K/S	96,0%	100,0%	100,0%
D/S	86,0%	62,0%	58,0%
N/D	46,5%	38,7%	44,8%
(S-K)/S	4,0%	0,0%	0,0%
(S-D)/S	14,0%	38,0%	42,0%
(D-N)/D	53,5%	61,3%	55,2%
(K-D)/K	10,4%	38,0%	42,0%
BGM/D	0,0%	6,5%	0,0%

**B.2. Metode yang diusulkan**

Metode yang diusulkan adalah klasifikasi dengan Algoritma Decision Tree C4.5, yang akan dimodelkan dan diimplementasi menggunakan Rapidminer 5.2.

**Algoritma Decision Tree C4.5**

Sering disebut dengan pohon keputusan (decision tree). Mirip sebuah struktur pohon dimana terdapat node internal yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas.

Tahapan Algoritma Decision Tree C4.5 Secara umum tahapan algoritma C4.5 adalah sebagai berikut [8]:

1. Menyiapkan data training.
2. Menentukan akar dari pohon dihitung dengan menggunakan persamaan 1:
 
$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$
3. Hitung nilai Gain dengan menggunakan persamaan 2:
 
$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$
4. Ulangi langkah ke-2 hingga semua tupel terpartisi
5. Proses partisi pohon keputusan akan berhenti saat:

- Semua tupel dalam node N mendapat kelas yang sama
- Tidak ada atribut di dalam tupel yang dipartisi lagi
- Tidak ada tupel di dalam cabang yang kosong

**B.3. Eksperimen dan Pengujian Model/Metode**

Algoritma akan diuji menggunakan data bulan Desember 2013 dan di cek dengan bulan lainnya. Untuk menaikkan akurasi dilakukan beberapa perubahan parameter C4.5 atau modifikasi bentuk data. Dalam setiap perubahan parameter, hasil akan dicatat dan divalidasi. Hasil untuk status gizi bermasalah: Rendah, Sedang, Tinggi. Hasil untuk Pemanfaatan Posyandu (Rasio masyarakat yang tidak memanfaatkan posyandu): Rendah, Sedang, dan Tinggi.

**C. HASIL DAN PEMBAHASAN**

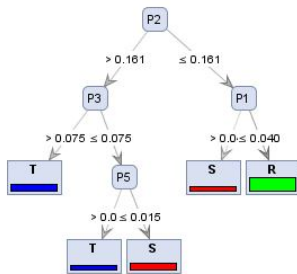
**C.1. Hasil Eksperimen Per Posyandu**

Data SKDN posyandu dihitung untuk mendapatkan nilai kecenderungan. Data input diperlihatkan pada tabel 2 berikut:

Tabel 2 Data Kecenderungan Per Posyandu

NO	POSYANDU	S	K	D	N	BGM		(S-K)/S		(S-D)/S		(D-N)/D		(K-D)/K		BGM/D	Hasil
						P1	P2	P3	P4	P5	P1	P2	P3	P4	P5		
1	Purnama Sari	50	50	17	10	0	0,00%	66,00%	41,18%	66,00%	0,00%	T					
2	Mulia Sari	101	98	93	88	0	2,97%	7,92%	5,38%	5,10%	0,00%	R					
3	Mekar Sari	84	84	66	65	2	0,00%	21,43%	1,52%	21,43%	3,03%	T					
4	Indah Sari	100	80	83	81	0	20,00%	17,00%	2,41%	-3,75%	0,00%	S					
5	Mayang Sari	97	97	80	70	1	0,00%	17,53%	12,50%	17,53%	1,25%	S					
6	Maya Sari	99	99	77	77	0	0,00%	22,22%	0,00%	22,22%	0,00%	T					
7	Restu Ibu	72	72	61	61	1	0,00%	15,28%	0,00%	15,28%	1,64%	S					
8	Antasan Indah I	121	121	103	101	0	0,00%	14,88%	1,94%	14,88%	0,00%	S					
9	Antasan Indah II	154	154	131	127	0	0,00%	14,94%	3,05%	14,94%	0,00%	S					
10	Antasan Indah III	158	150	140	130	0	5,06%	11,39%	7,14%	6,67%	0,00%	S					
11	Antasan Indah IV	133	133	120	116	0	0,00%	9,77%	3,33%	9,77%	0,00%	R					

Data pada tabel 3, diimplementasikan ke dalam algoritma C4.5 menggunakan rapidminer, menghasilkan pohon keputusan dengan akurasi 63,6 sebagaimana terlihat pada gambar 2.



Gambar 2 Diagram hasil data tabel 2, akurasi 63,6

Karena dengan nilai kecenderungan secara langsung masih belum memberikan hasil yang diharapkan, maka data dimodifikasi. Data diberi nilai bobot. Pembobotan masing-masing kecenderungan ditampilkan pada tabel 3.

Tabel 3 Label Kategori Nilai Kecenderungan

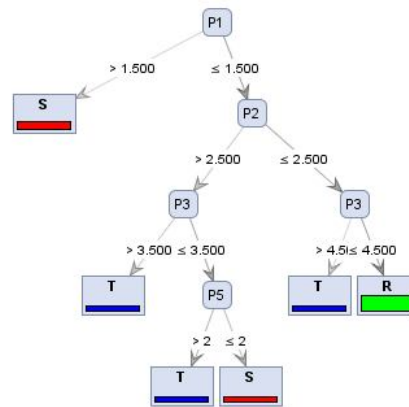
No	Variabel	Rentang	Bobot	Label
P1	Kehilangan kesempatan: (S-K)/S	< 5	1	Rendah
		5 - 10	2	Sedang
		> 10	3	Tinggi
P2	Pemanfaatan Posyandu: (S-D)/S	< 10	1	Rendah
		10 - 20	2	Sedang
		> 20	3	Tinggi
P3	Status Gizi Bermasalah: (D-N)/D	< 5	1	Rendah
		5 - 10	3	Sedang
		> 10	6	Tinggi
P4	Dropout: (K-D)/K	< 10	1	Rendah
		10 - 20	2	Sedang
		> 20	3	Tinggi
P5	TBGM: BGM/D	< 3	1	Rendah
		3-6	3	Sedang
		> 6	6	Tinggi

Data pada tabel 2 yang telah mengalami pembobotan ditunjukkan pada tabel 4. Dimana R= Rendah, S= Sedang, T = Tinggi.

Tabel 4 Pembobotan Data Kecenderungan Per Posyandu

POSYANDU	P1	P2	P3	P4	P5	Hasil
Purnama Sari	1	3	6	3	1	T
Mulia Sari	1	1	3	1	1	R
Mekar Sari	1	3	1	3	3	T
Indah Sari	3	2	1	1	1	S
Mayang Sari	1	2	6	2	1	T
Maya Sari	1	3	1	3	1	S
Restu Ibu	1	2	1	2	1	R
Antasan Indah I	1	2	1	2	1	R
Antasan Indah II	1	2	1	2	1	R
Antasan Indah III	2	2	3	1	1	S
Antasan Indah IV	1	1	1	1	1	R

Data pada tabel 3, diimplementasikan ke dalam algoritma C4.5, menghasilkan pohon keputusan dengan akurasi 72,7 pada gambar 3.



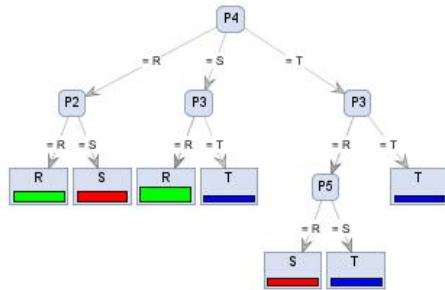
Gambar 3 Diagram hasil data tabel 4, akurasi 72,7

Karena masih belum memberikan hasil yang diharapkan, maka data kembali dimodifikasi. Data diberi nilai label sesuai bobot pada tabel 3. Pembobotan masing-masing kecenderungan ditampilkan pada tabel 5.

Tabel 5 Pelabelan Data Per Posyandu

POSYANDU	P1	P2	P3	P4	P5	Hasil
Purnama Sari	R	T	T	T	R	T
Mulia Sari	R	R	S	R	R	R
Mekar Sari	R	T	R	T	S	T
Indah Sari	T	S	R	R	R	S
Mayang Sari	R	S	T	S	R	T
Maya Sari	R	T	R	T	R	S
Restu Ibu	R	S	R	S	R	R
Antasan Indah I	R	S	R	S	R	R
Antasan Indah II	R	S	R	S	R	R
Antasan Indah III	S	S	S	R	R	S
Antasan Indah IV	R	R	R	R	R	R

Data pada tabel 5, menghasilkan pohon keputusan dengan akurasi yang cukup baik, yaitu 92,3. Diagram ditunjukkan pada gambar 4.



Gambar 4 Diagram hasil data tabel 5, akurasi 90,9

Hasil ini menjadi dasar untuk penggambaran sebaran potensi gizi bermasalah (peta 3).

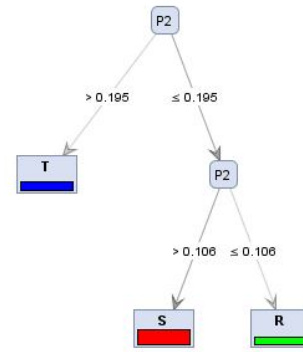
Per Peta Bertahap Posyandu

Data kecenderungan posyandu untuk peta 1 diperlihatkan pada tabel 6 berikut:

Tabel 6 Data Kecenderungan untuk peta 1

NO	POSYANDU	(S-K)/S (S-D)/S (K-D)/K			Hasil
		P1	P2	P4	
1	Purnama Sari	0,00%	66,00%	66,00%	T
2	Mulia Sari	2,97%	7,92%	5,10%	R
3	Mekar Sari	0,00%	21,43%	21,43%	T
4	Indah Sari	20,00%	17,00%	-3,75%	S
5	Mayang Sari	0,00%	17,53%	17,53%	S
6	Maya Sari	0,00%	22,22%	22,22%	T
7	Restu Ibu	0,00%	15,28%	15,28%	S
8	Antasan Indah I	0,00%	14,88%	14,88%	S
9	Antasan Indah II	0,00%	14,94%	14,94%	S
10	Antasan Indah III	5,06%	11,39%	6,67%	S
11	Antasan Indah IV	0,00%	9,77%	9,77%	R

Data pada tabel 6, diimplementasikan ke dalam algoritma C4.5 menggunakan rapidminer, menghasilkan pohon keputusan dengan akurasi 81,8 sebagaimana terlihat pada gambar 5.



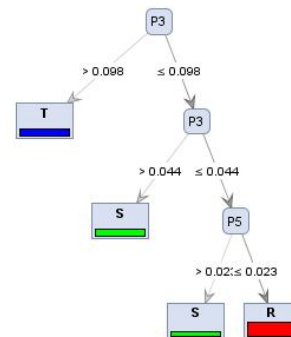
Gambar 5 Diagram hasil data tabel 6, akurasi 81,8

Data kecenderungan posyandu untuk peta 2 diperlihatkan pada tabel 7 berikut:

Tabel 7 Data Kecenderungan untuk peta 2

NO	POSYANDU	(D-N)/D (BGM)/D		Hasil
		P3	P5	
1	Purnama Sari	41,18%	0,00%	T
2	Mulia Sari	5,38%	0,00%	S
3	Mekar Sari	1,52%	3,03%	S
4	Indah Sari	2,41%	0,00%	R
5	Mayang Sari	12,50%	1,25%	T
6	Maya Sari	0,00%	0,00%	R
7	Restu Ibu	0,00%	1,64%	R
8	Antasan Indah I	1,94%	0,00%	R
9	Antasan Indah II	3,05%	0,00%	R
10	Antasan Indah III	7,14%	0,00%	S
11	Antasan Indah IV	3,33%	0,00%	R

Data pada tabel 7, menghasilkan pohon keputusan dengan akurasi 72,7 sebagaimana terlihat pada gambar 6.



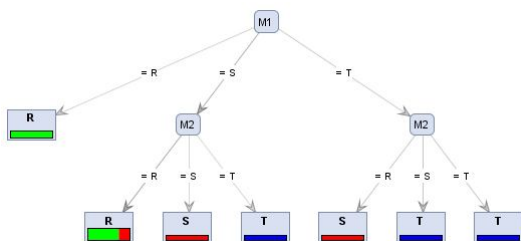
Gambar 6 Diagram hasil data tabel 7, akurasi 72,2

Data peta 1 dan peta 2 dipergunakan sebagai input untuk peta 3. Data input diperlihatkan pada tabel 8 berikut:

Tabel 8 Data Peta 1 dan Peta 2 untuk peta 3

POSYANDU	M1	M2	Hasil
Purnama Sari	T	T	T
Mulia Sari	R	S	S
Mekar Sari	T	S	T
Indah Sari	S	R	S
Mayang Sari	S	T	T
Maya Sari	T	R	S
Restu Ibu	S	R	S
Antasan Indah I	S	R	S
Antasan Indah II	S	R	S
Antasan Indah III	S	S	S
Antasan Indah IV	R	R	R

Data pada tabel 8, menghasilkan pohon keputusan dengan akurasi hanya 54,5 sebagaimana terlihat pada gambar 7.



Gambar 7 Diagram hasil data tabel 8, akurasi 54,5

C.2. Evaluasi dan validasi hasil

Pelabelan yang dilakukan oleh C4.5 dibandingkan dengan pelabelan yang dihasilkan dengan penghitungan yang merujuk pada standart WHO yang diadopsi oleh Kementerian Kesehatan dalam standart antropometri penilai status gizi anak [9]. Hasilnya tabel 9.

Tabel 9 Perbandingan Hasil dan Nilai Rujukan

POSYANDU	Rujukan	C4.5 (Langsung)	C4.5 (Bertahap)
Purnama Sari	T	T	T
Mulia Sari	R	R	S
Mekar Sari	T	S	T
Indah Sari	S	S	S
Mayang Sari	T	T	T
Maya Sari	S	S	S
Restu Ibu	R	R	S
Antasan Indah I	R	R	S
Antasan Indah II	R	R	S
Antasan Indah III	S	S	S
Antasan Indah IV	R	R	R

Perhitungan yang berdasarkan pada nilai kecenderungan (P1, P2, P3, P4 dan P5) menghasilkan akurasi 90,9 %. Perhitungan bertahap yaitu menghasilkan peta 3 berdasar hasil peta 1 dan peta 2, hanya mencapai akurasi 54,5 %.

Dengan demikian, untuk menghasilkan peta 1, tetap berbasis P1, P2, P4 karena telah memiliki akurasi 81,8 %. Peta 2, berbasis P3 dan P5, yang dapat mencapai akurasi 72,7. Sedangkan untuk menghasilkan peta sebaran potensi gizi bermasalah

(peta3), lebih baik jika dasarkan langsung pada data lima nilai kecenderungan.

D. KESIMPULAN

Berdasarkan hasil penelitian dan pengukuran, dapat disimpulkan:

1. Pohon keputusan berbasis data P1, P2, P4 tetap dapat dipakai untuk peta 1, karena telah memiliki akurasi 81,8 %. Peta 2, berbasis P3 dan P5, yang dapat mencapai akurasi 72,7.
2. Peta wilayah berpotensi gizi bermasalah lebih baik dihitung langsung dari lima nilai kecenderungan SKDN, dengan akurasi 90,9 %
3. Algoritma decision tree C4.5 dapat dipakai untuk menghasilkan peta sebaran potensi gizi bermasalah.

E. SARAN

Beberapa hal perlu disampaikan untuk pengujian yang lebih lanjut:

1. Selain data kecenderungan SKDN, dapat dipakai data hasil pencatatan antropometri.
2. Hasil Algoritma C4.5 dapat diuji untuk data per puskesmas dalam kecamatan, atau seluruh puskesmas dalam satu kota/kabupaten.
3. Peta dapat dibuat berdasarkan parameter target dan pencapaian tiap posyandu atau puskesmas.

DAFTAR PUSTAKA

- [1] Depkes RI, Pencegahan Dan Penanggulangan Gizi Buruk.: Direktorat Bina Gizi Masyarakat, 2006.
- [2] Noviatu Fuada, Sri Muljati, and Tjetjep S Hidayat, "Penentuan Daerah Rawan Gizi Berdasarkan Analisis Spasial," Media Litbang Kesehatan, vol. 22, no. 1, pp. 18-29, Maret 2012.
- [3] Yose Rizal, "Distribusi Spasial kasus Gizi Buruk Dan Gizi Kurang Pada Balita Di Kecamatan Mapat Tunggal Kabupaten Pasaman Tahun 2007," Sekolah Pasca Sarjana UGM, Yogyakarta, Tesis 2008.
- [4] Zhiwu Liu and Xiuzhi Zhang, "Prediction and Analysis for Students' Marks Based on Decision Tree Algorithm," in 2010 Third International Conference on Intelligent Networks and Intelligent Systems, 2010.
- [5] Rong Cao and Lizhen Xu, "Improved C4.5 Algorithm for the Analysis of Sales," in 2009 Sixth Web Information Systems and Applications Conference, 2009.
- [6] Oded Maimon Lior Rokach, Data Mining and Knowledge Discovery Handbook., 2010.

- [7] Seksi Kesehatan Gizi, Bidang Kesehatan Keluarga, "Perbaikan Gizi Masyarakat Kota Banjarmasin 2013," Dinas Kesehatan Kotamadya Banjarmasin, Banjarmasin, Laporan Tahunan 2013.
- [8] Kusrini and Emha Taufiq Lutfhfi, *Algoritma Data Mining*.: ANDI Yogyakarta, 2009.
- [9] Kementerian Kesehatan RI, SK Menkes RI No. 1995/Menkes/SK/XII/2010 Tentang Standar Antropometri Penilaian Status Gizi Anak.: Direktorat Bina Gizi, Direktorat Jenderal Bina Gizi Dan Kesehatan Ibu dan Anak, 2011.