

ANALISA FAKTOR PENYEBAB STUNTING MENGGUNAKAN ALGORITMA C4.5

Analysis Of Stunting Causing Factors Using C4.5 Algorithm

Hamsir Saleh

Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo.

e-mail : hamsir.saleh@gmail.com

Abstract

Stunting or dwarf is a condition where a toddler has a length or height that is less when compared to age. There are five main factors causing stunting, namely poverty, social and culture, increased exposure to infectious diseases, food insecurity and public access to health services. C4.5 algorithm is a data classification algorithm with a decision tree method that has advantages such as being able to process numeric (continuous) and discrete data, rules that are formed based on exiting decision tree, the main factors causing stunting in the Dulupi sub – district, Boalemo Regency, are Health Facilities and exclusive breastfeeding for babies 0 – 6 months.

Keywords – Stunting, C4.5, decision tree.

1. Pendahuluan

Meningkatkan kualitas hidup manusia Indonesia melalui peningkatan kualitas pendidikan dan pelatihan dengan program “Indonesia Pintar” dan membangun Indonesia dari pinggiran dengan memperkuat daerah – daerah dan desa dalam kerangka Negara kesatuan merupakan jbaran dari Nawa Cita ke – 5 dan ke – 3, namun upaya menghadirkan generasi emas Indonesia ini dibayangi kehadiran stunting yang masih mengancam. Stunting merujuk pada kondisi anak yang lebih pendek dari tinggi badan seumurannya, stunting terjadi lantaran kekurangan gizi dalam waktu lama pada masa 1.000 (seribu) hari pertama kehidupan [1].

Gangguan tumbuh kembang pada anak akibat kekurangan gizi bila tidak mendapatkan intervensi sejak dini akan berlanjut hingga dewasa. Stunting pada balita perlu mendapatkan perhatian khusus karena dapat menyebabkan terhambatnya pertumbuhan fisik, perkembangan mental dan status kesehatan pada anak, studi terkini menunjukkan anak yang mengalami stunting berkaitan dengan prestasi di sekolah yang buruk, tingkat pendidikan yang rendah dan pendapatan yang rendah saat dewasa. Kasus stunting pada anak dapat dijadikan indikator rendahnya kualitas sumber daya manusia suatu Negara [2].

Provinsi Gorontalo masuk dalam daftar 1.000 desa prioritas stunting tahun 2018, Kabupaten Gorontalo dan Kabupaten Boalemo menjadi prioritas stunting diantara kabupaten kota yang ada di Provinsi Gorontalo. Khusus untuk Kabupaten Boalemo desa prioritas stunting tersebar di 5 kecamatan diantaranya Kecamatan Paguyaman dengan target desa prioritas stunting desa Bongo Tua. Kecamatan Dulupi terdapat 3 desa prioritas yaitu desa Pangi, desa Tangga Jaya, desa Tanah Putih. Kecamatan Tilamuta sebagai ibukota Kabupaten terdapat 1 desa prioritas yaitu desa Bajo. Kecamatan Botomoito terdapat 2 desa prioritas stunting yaitu desa Hutamonu, dan desa Bolihutuo. Sementara itu Kecamatan Paguyaman Pantai terdapat 3 desa prioritas stunting yaitu desa Bubaa,

desa Lito dan desa Tawayu [3].

Pohon keputusan adalah salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia, algoritma C4.5 adalah algoritma klasifikasi data dengan metode pohon keputusan yang memiliki kelebihan misalnya dapat mengolah data numeric (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan – aturan yang mudah diinterpretasikan dan tercepat diantara algoritma – algoritma yang menggunakan memori utama di komputer [4].

Hampir seluruh penyebab langsung terhadap kejadian stunting telah ditangani dengan baik oleh kebijakan program pemerintah di Indonesia, namun permasalahan dalam menekan angka kejadian stunting adalah mendeteksi faktor utama penyebab stunting dalam wilayah kecamatan baik perkotaan maupun desa tertentu. Oleh sebab itu dalam penelitian ini algoritma C4.5 digunakan untuk menganalisa faktor utama penyebab stunting dalam wilayah Kecamatan Dulupi Kabupaten Boalemo.

2. Tinjauan Pustaka

2.1 Stunting

Stunting atau kerdil adalah kondisi dimana balita memiliki panjang atau tinggi badan yang kurang jika dibandingkan dengan umur. Kondisi ini diukur dengan panjang atau tinggi badan yang lebih dari minus dua standar deviasi median standar pertumbuhan anak dari WHO. Balita stunting termasuk masalah gizi kronik yang disebabkan oleh banyak faktor seperti kondisi sosial ekonomi, gizi ibu saat hamil, kesakitan pada bayi dan kurangnya asupan gizi pada bayi. Balita stunting di masa yang akan datang akan mengalami kesulitan dalam mencapai perkembangan fisik dan kognitif yang optimal [5].

Stunting pada anak balita merupakan konsekuensi dari beberapa faktor yang sering dikaitkan dengan kemiskinan termasuk gizi, kesehatan, sanitasi dan lingkungan. Ada lima faktor utama penyebab stunting yaitu kemiskinan, sosial dan budaya, peningkatan paparan terhadap penyakit infeksi, kerawanan pangan dan akses masyarakat terhadap pelayanan kesehatan [6].

Penentu utama terjadinya stunting pada anak di Indonesia adalah ASI tidak eksklusif pada 6 bulan pertama, status ekonomi keluarga yang rendah, bayi lahir prematur, panjang badan bayi lahir yang pendek, ibu yang pendek, tingkat pendidikan orang tua rendah dan anak yang tinggal di daerah miskin perkotaan dan di daerah pedesaan. Sementara itu faktor lain terjadinya stunting pada anak di Indonesia adalah kecenderungan anak laki – laki lebih berisiko mengalami stunting dari pada anak perempuan, anak – anak dari keluarga dengan jamban yang buruk dan air minum tidak layak meningkatkan risiko terjadinya stunting, serta faktor masyarakat dan sosial seperti akses yang rendah terhadap pelayanan kesehatan[7].

2.2 Algoritma C4.5

Proses pelatihan dalam algoritma C4.5 [8]:

1. Tentukan akar dari pohon dengan menggunakan gain ratio. Atribut dengan nilai gain ratio tertinggi dipilih sebagai atribut test untuk simpul. Langkah menghitung gain ratio :
 - a. Hitung nilai total entropy data set dan masing – masing atribut

- b. Hitung nilai gain masing – masing atribut, dengan metode information gain.
- c. Hitung split info masing – masing atribut.
- d. Setelah mendapatkan nilai diatas hitung gain ratio semua atribut.
2. Menentukan Leaf Node berikutnya :
 - a. Label yang mengandung nilai Positif (+) dan Negatif (-) merupakan leaf node berikutnya
 - b. Node yang terpilih merupakan node yang memiliki nilai gain ratio tertinggi
 - c. Apabila node memiliki nilai gain ratio yang sama maka akan ditentukan oleh pakar (expert)
3. Lakukan langkah yang sama untuk leaf node berikutnya
4. Setelah pohon keputusan terbentuk langkah selanjutnya adalah membuat rule berdasarkan pohon keputusan yang telah ada
5. Aturan – aturan yang diperoleh dari pohon keputusan tersebut untuk membentuk model.
6. Apabila terdapat hasil data yang bertentangan pada proses ini maka diperlukan seorang pakar untuk pengambilan keputusan.

Beberapa tahap dalam menggunakan algoritma C4.5 [9] yaitu :

1. Menyiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas – kelas tertentu.
2. Menentukan akar dari pohon. Dalam pembentukan pohon keputusan menggunakan metode algoritma *decision tree* (C4.5), akar diambil dari atribut yang terpilih dengan cara menghitung nilai gain dari masing – masing atribut, sebelum menghitung nilai gain dari atribut, hitung dahulu nilai *entropy* dengan menggunakan rumus :

$$Entropy (S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$

Keterangan :

S = Himpunan kasus

n = Jumlah partisi S

P_i = Proporsi dari S_i terhadap S

3. Kemudian hitung nilai gain dengan metode *information gain* :

$$Gain (S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy (S_i) \quad (2)$$

Keterangan :

S = Himpunan kasus

A = Atribut

n = Jumlah partisi atribut A

$|S_i|$ = Jumlah kasus pada partisi ke- i

$|S|$ = Jumlah kasus dalam S

Information gain mempelajari informasi dari data, dan menggunakan nilai tersebut sebagai acuan dalam menentukan atribut yang akan digunakan dalam menyusun pohon keputusan, atribut yang dipilih akan menghasilkan partisi dengan data yang lebih seragam dan dapat menghasilkan pohon keputusan yang sederhana mungkin dengan perulangan yang sedikit.

Untuk menghitung *gain ratio* perlu diketahui pemisahan informasi atau *split info*, *split info* dihitung dengan cara [4]:

$$SplitInfo (S, A) = - \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (3)$$

Keterangan :

- S = Himpunan kasus
- A = Atribut
- c = Jumlah partisi atribut A
- S_i = Jumlah kasus pada partisi ke- i
- S = Jumlah kasus dalam S

Selanjutnya *gain ratio* dihitung dengan cara :

$$Gain\ Ratio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)} \quad (4)$$

Keterangan :

- S = Himpunan kasus
- A = Atribut

4. Setelah mendapatkan nilai *gain ratio* kita tentukan variabel berikutnya :
 - a) Class yang mengandung nilai Positif (+) dan Negatif (-) merupakan variabel berikutnya
 - b) Variabel yang terpilih merupakan variabel yang memiliki nilai *gain ratio* tertinggi
5. Lakukan langkah yang sama untuk variabel berikutnya
6. Setelah pohon keputusan terbentuk langkah selanjutnya adalah membuat rule berdasarkan pohon keputusan yang telah ada.

3. Hasil dan Pembahasan

Jumlah data anak balita yang mengalami stunting di tiga desa Kecamatan Dulupi Kabupaten Boalemo Provinsi Gorontalo, terindikasi positif stunting sebanyak 111 balita dan negatif stunting sebanyak 19 balita.

Data jenis kelamin terdiri dari laki laki dan perempuan, data status gizi BB/U terdiri dari gizi buruk, gizi kurang dan gizi baik, status gizi TB/U terdiri dari sangat pendek dan pendek, status gizi BB/TB terdiri dari gemuk, kurus dan normal, semntara itu untuk fasilitas kesehatan terdiri dari puskesmas, pustu dan tidak ada fasilitas kesehatan, untuk pemberian asi eksklusif terdiri dari ya dan tidak, penggunaan jamban terdiri dari menumpang dan pribadi, pola hidup bersih dan sehat terdiri dari ya dan tidak, dan saluran pembuangan air limbah (SPAL) terdiri dari kubangan dan got.

Berikut tahap dalam melakukan penentuan faktor – faktor penyebab stunting di Kecamatan Dulupi.

- a. Menentukan akar dari pohon keputusan, akar diambil dari atribut yang terpilih dengan cara menghitung nilai *gain ratio* dari masing – masing atribut, nilai *gain ratio* yang paling tinggi akan menjadi akar pertama.
- b. Menghitung nilai *entropy* total dengan menggunakan persamaan 1 sebagai berikut :

$$Entropy\ Total = -\frac{111}{130} \log_2 \frac{111}{130} + -\frac{19}{130} \log_2 \frac{19}{130} = 0.60$$
 Dimana terdapat 111 data positif dan 19 data negatif.
- c. Lakukan perhitungan *entropy* untuk masing – masing atribut, untuk perhitungan atribut Jenis Kelamin tetap menggunakan persamaan 1 dengan penjelasan sebagai berikut.

Tabel 1. Perhitungan Atribut Jenis Kelamin

JK	Status	Jumlah
Laki – Laki	Positif	60
Laki – Laki	Negatif	9
Perempuan	Positif	51
Perempuan	Negatif	10

$$\text{Laki - Laki} = -\frac{60}{69} \log_2 \frac{60}{69} + -\frac{9}{69} \log_2 \frac{9}{69} = 0.56$$

$$\text{Perempuan} = -\frac{51}{61} \log_2 \frac{51}{61} + -\frac{10}{61} \log_2 \frac{10}{61} = 0.64$$

- d. Setelah mendapatkan hasil *entropy* dari masing – masing nilai atribut jenis kelamin langkah selanjutnya adalah menghitung nilai gain atribut Jenis Kelamin menggunakan persamaan 2 sebagai berikut :

$$\text{Gain(Total,JK)} = 0.60 - \left(\left(\frac{69}{130} * 0.56 \right) + \left(\frac{51}{130} * 0.64 \right) \right) = 0.0016$$

- e. Perhitungan nilai *split info* menggunakan persamaan 3

$$\text{Split info (Total, JK)} = -\frac{69}{130} \log_2 \frac{69}{130} + -\frac{51}{130} \log_2 \frac{51}{130} = 1$$

- f. Menghitung *gain ratio* untuk menentukan atribut sebagai akar dari pohon keputusan menggunakan persamaan 4

$$\text{Gain Ratio (Total, JK)} = \frac{0.0016}{1} = 0.0016$$

Untuk atribut yang lain dilakukan perhitungan yang sama, untuk nilai gain ratio paling tinggi akan menjadi node akar dari pohon. Hasil atribut Fasilitas Kesehatan memiliki nilai gain ratio tertinggi yaitu 0.8889, sehingga atribut Fasilitas Kesehatan menjadi akar dari pohon keputusan sesuai dengan tabel di bawah ini :

Tabel 2. Perhitungan Atribut Node 1 Keseluruhan

Atribut	Gain ratio
Jenis Kelamin	0.0016
BB/U	0.0229
TB/U	0.0235
BB/TB	0.0611
Fas Kes	0.8889
ASI Eksklusif	0.6047
PHBS	0.4611
Jamban	0.0801
SPAL	0.0942

Atribut Fasilitas Kesehatan memiliki tiga buah nilai yaitu Puskesmas, Pustu dan Tidak Ada Fasilitas, dari nilai tersebut puskesmas dan pustu sudah mengklasifikasikan kasus menjadi Positif, sehingga tidak perlu dilakukan partisi atribut lagi. Untuk nilai tidak ada fasilitas perlu dilakukan partisi atribut lagi.

Untuk mendapatkan atribut selanjutnya lakukan perhitungan yang sama untuk 8 atribut lainnya sehingga delapan nilai atribut lainnya dapat di tentukan. Hasil perhitungan nilai gain ratio atribut berikutnya adalah :

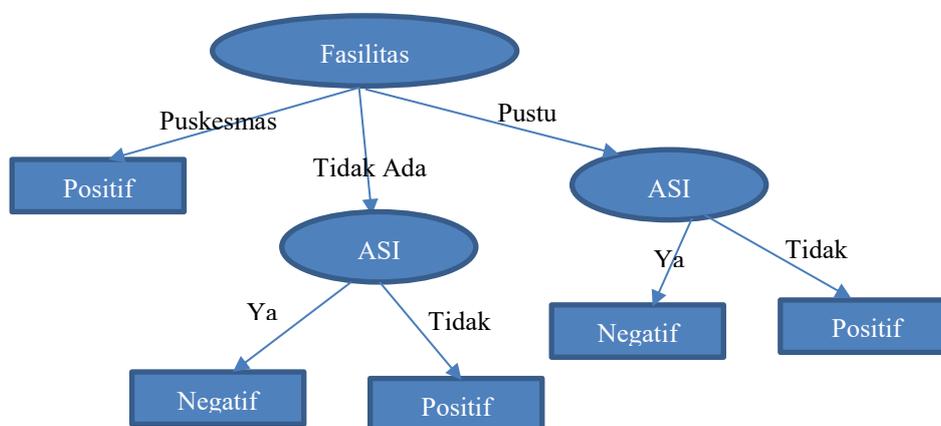
Tabel 3. Perhitungan Atribut Node 2 Keseluruhan

ATRIBUT	GAIN RATIO
Jenis Kelamin	0.0053
BB/U	0.0610
TB/U	0.0143
BB/TB	0.0875
ASI Eksklusif	0.5836
PHBS	0.4464
Jamban	0.0881
SPAL	0.1158

Atribut Asi Eksklusif merupakan nilai gain ratio tertinggi dari tujuh atribut lainnya, Atribut Asi Eksklusif memiliki dua buah nilai yaitu ya dan tidak. Nilai dari tidak sudah diklasifikasikan dalam kategori positif stunting dan ya Asi Eksklusif sudah diklasifikasikan dalam kategori negatif stunting, sehingga tidak perlu lagi dilakukan perhitungan nilai gain ratio.

Hasil perhitungan tertinggi nilai gain ratio menggunakan algoritma C4.5 dengan total data berjumlah 130 balita menghasilkan atribut Fasilitas Kesehatan. Nilai dari atribut Fasilitas Kesehatan 0.8889, Fasilitas Kesehatan memiliki nilai Puskesmas, Pustu dan tidak ada fasilitas menjadi faktor pertama penyebab stunting.

Dengan jumlah data tersisa 76 balita, nilai gain ratio tertinggi menghasilkan atribut Asi Eksklusif menjadi faktor kedua penyebab stunting di keseluruhan data dengan nilai gain ratio 0.5836, atribut Asi Eksklusif nilai Tidak Eksklusif positif stunting. Proses perhitungan yang telah di buat menghasilkan sebuah pohon keputusan sebagai berikut :



Gambar 1. Hasil Pohon Keputusan

Dari pohon keputusan tersebut menghasilkan suatu aturan sebagai berikut :

- 1) Jika Fasilitas Kesehatan = Puskesmas maka Positif stunting
- 2) Jika Fasilitas Kesehatan = Pustu dan Asi Eksklusif = Tidak maka Positif Stunting
- 3) Jika Fasilitas Kesehatan = Pustu dan Asi Eksklusif = ya maka Negatif Stunting.
- 4) Jika Fasilitas Kesehatan = Tidak dan Asi Eksklusif = ya maka Negatif Stunting.
- 5) Jika Fasilitas Kesehatan = Tidak dan Asi Eksklusif = tidak maka Positif Stunting.

Pola data yang telah dihitung mendapatkan hasil bahwa yang menjadi Faktor utama penyebab stunting adalah Fasilitas Kesehatan dan Asi Eksklusif.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan di dalam Kecamatan Dulupi Kabupaten Boalemo dapat disimpulkan bahwa yang menjadi faktor utama penyebab stunting di wilayah tersebut adalah Fasilitas Kesehatan dan pemberian Asi Eksklusif kepada bayi 0 – 6 bulan, fasilitas kesehatan yang kurang serta terpusat di kecamatan mengakibatkan pelayanan kesehatan tidak merata, sehingga diharapkan dengan hasil ini Pemerintah Kecamatan perlu membentuk tim khusus dalam memberikan penyuluhan dan pemahaman tentang pentingnya Asi Eksklusif untuk bayi usia 0 - 6 bulan.

DAFTAR REFERENSI

- [1] E. P. Sandjojo, “Buku Saku Desa dalam Penanganan Stunting,” *Kementeri. Desa Pembang. Drh. Tertinggal Dan Transm.*, p. 42, (2017).
- [2] E. Setiawan, R. Machmud, and M. Masrul, “Faktor-Faktor yang Berhubungan dengan Kejadian Stunting pada Anak Usia 24-59 Bulan di Wilayah Kerja Puskesmas Andalas Kecamatan Padang Timur Kota Padang Tahun 2018,” *J. Kesehat. Andalas*, vol. 7, no. 2, pp. 275–284, (2018).
- [3] TNP2K, “Pemilihan 10 Desa Prioritas di 100 Kabupaten / Kota Prioritas Penanganan Kemiskinan dan Stunting, Versi Ringkas,” Kementerian PPN/Bappenas, Kementerian Koordinator Bidang Pembangunan Manusia dan Kebudayaan, Nov. (2017).
- [4] N. Jayanti, S. Puspitodjati, and T. Elida, “Teknik Klasifikasi Pohon Keputusan Untuk Memprediksi Kebangkrutan Bank Berdasarkan Rasio Keuangan Bank,” presented at the Proceeding Seminar Ilmiah Nasional Komputer dan Sistem Intelejen (KOMMIT), Depok, (2008).
- [5] R. I. Kemenkes, “Situasi Balita Pendek (stunting) di Indonesia,” *Bul. Jendela Data Dan Inf. Kesehat.*, vol. 1, (2018).
- [6] F. O. Aridiyah, N. Rohmawati, and M. Ririanty, “Faktor-faktor yang Mempengaruhi Kejadian Stunting pada Anak Balita di Wilayah Pedesaan dan Perkotaan (The Factors Affecting Stunting on Toddlers in Rural and Urban Areas),” *Pustaka Kesehat.*, vol. 3, no. 1, pp. 163–170, (2015).
- [7] T. Beal, A. Tumilowicz, A. Sutrisna, D. Izwardy, and L. M. Neufeld, “A review of child stunting determinants in Indonesia,” *Matern. Child. Nutr.*, vol. 14, no. 4, p. e12617, 2018.
- [8] H. Saleh, “Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma C4. 5 Berbasis Forward Selection,” *Ilk. J. Ilm.*, vol. 9, no. 2, pp. 173–180, (2017).
- [9] Kusrini and Luthfi, *Algoritma Data Mining*. Yogyakarta: Andi Publishing, (2009).