Penerapan Metode Klasifikasi *Decision Tree* Pada Status Gizi Balita Di Kabupaten Simalungun

Hananda Hafizan[1], Anggita Nadia Putri[2]*
Prodi Sistem Informasi, STIKOM Tunas Bangsa
Jln. Jendral Sudirman Blok A No.1/2/3 Pematangsiantar, Medan, Indonesia 21127
*hanandahafizan@gmail.com[*1]*

Abstract

One of the health problems in Indonesia is the problem of nutritional status of children under five years. Cases of malnutrition are not only a family problem, but also a state problem. The nutritional status of children under five years can be assessed by measuring the human body known as "Anthropometry". To be able to carry out anthropometric examinations and measurements in order to find out the nutritional status of children under five, they can go to public health service places such as the Posyandu. We went to the KENANGA Posyandu located in Wonorejo, Kerasaan sub-district, Simalungun district. The purpose of this study will be to test the model for the classification of nutritional status of children under the WHO-2005 reference standard by utilizing data mining techniques using the Decision Tree method C4.5 Algorithm.

***Keywords** : Nutritional Status, Anthropometry, Decision Tree, C4.5 Algorithm;*

Abstrak

Salah satu permasalahan kesehatan di Indonesia adalah masalah status gizi balita. Kasus gizi buruk tidak hanya menjadi beban keluarga, tetapi juga menjadi beban negara. Status gizi balita dapat dinilai dengan mengukur tubuh manusia yang dikenal dengan istilah "Antropometri". Untuk dapat melakukan pemeriksaan dan pengukuran antropometri agar dapat mengetahui status gizi balita dapat mendatangi tempat pelayanan kesehatan masyarakat seperti salah satunya yaitu posyandu. Kami mendatangi Posyandu KENANGA yang berada di kelurahan Wonorejo kecamatan Kerasaan kabupaten Simalungun. Tujuan dari penelitian ini akan menguji coba model untuk klasifikasi status gizi balita berdasarkan baku rujukan WHO-2005 dengan memanfaatkan teknik data mining menggunakan metode *Decision Tree* Algoritma C4.5.

Kata kunci : Status Gizi, Antropometri, *Decision Tree*, Algoritma C4.5;

1. Pendahuluan

Status gizi balita merupakan salah satu permasalahan kesehatan di Indonesia. Kasus gizi buruk tidak hanya menjadi beban bagi keluarga tetapi juga menjadi beban bagi negara. Balita mengalami siklus pertumbuhan dan perkembangan yang membutuhkan zat-zat gizi yang lebih besar dari kelompok unsur lain sehingga balita paling mudah menderita kelainan gizi. Status gizi balita dapat dinilai dapat dinilai dengan mengukur tubuh manusia yang dikenal sebagai "Antropometri". Pengukuran antropometri dapat dilakukan dengan melakukan pengukuran indikator berat badan dan tinggi badan serta memperhatikan umur dan jenis kelamin balita itu sendiri. Ukuran antropometri untuk penilaian status gizi merupakan kombinasi antara masing-masing ukuran indikator antropometri yang umum digunakan untuk menilai status gizi yang umum adalah indeks berat badan terhadap umur (BB/U), indeks berat badan terhadap tinggi badan (BB/TB) dan indeks tinggi badan terhadap umur (TB/U). Untuk anak pada umumnya, indeks BB/U merupakan cara baku yang digunakan untuk mengukur pertumbuhan. Indeks BB/TB merupakan ukuran antropometri yang terbaik karena dapat menggambarkan status gizi saat ini dengan lebih sensitif dan spesifik.

ISSN: 2720-992X
Copyright © KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)

Untuk dapat melakukan pemeriksaan dan pengukuran antropometri agar dapat mengetahui status gizi balita dapat mendatangi tempat pelayanan kesehatan masyarakat seperti salah satunya yaitu posyandu yang berada di wilayah tempat tinggal. Posyandu KENANGA merupakan posyandu yang berada di kelurahan Wonorejo kecamatan Kerasaan kabupaten Simalungun. Di posyandu ini laporan data balita disimpan dalam buku catatan posyandu dan file digital. Data tersebut apabila digali dengan tepat maka dapat diketahui pola atau pengetahuan untuk mengambil keputusan. Serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data disebut dengan Data Mining [1]–[4].

Data Mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut [5]. Data Mining sendiri memiliki beberapa teknik salah satunya klasifikasi. Klasifikasi merupakan salah satu proses pada data mining yang bertujuan untuk menemukan pola yang berharga dari data yang berukuran relatif besar hingga sangat besar. Algoritma data mining C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokan dan bersifat prediktif. Algortima C4.5 sendiri merupakan pengembangan dari algortima ID3. Kemampuan klasifikasi Algoritma C4.5 dapat digunakan untuk membangun pohon keputusan (*Decision Tree*) [1], sehingga terlihat kecenderungan status gizi balita dan dapat ditindaklanjuti pemangku kepentingan sebagaimana seharusnya.

2. Metodologi Penelitian

2.1. Data Mining

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Data mining biasa juga dikenal nama lain seperti : Knowledge discovery (*mining*) in databases (KDD), ekstraksi pengetahuan (*knowledge extraction*) Analisa data/pola dan kecerdasan bisnis (*business intelligence*) [4].

2.2. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat [6]–[8]. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Rumus Algoritma C4.5 terbagi menjadi 2 rumus, yang pertama rumus untuk mencari nilai gain. Adapun rumus gain adalah sebagai berikut:

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^{n} \frac{|S_1|}{|S|} * Entropy\ (Si) \quad (1)$$

Keterangan :
S : Himpunan Kasus
A : Atribut
n : Jumlah Partisi Atribut A
$|S_1|$: Jumlah kasus pada partisi ke-i
$|S|$: Jumlah Kasus dalam S
Dan rumus yang kedua adalah :

$$Entropy(S) = \sum_{i=1}^{n}(-p_i) * \log_2 p_i \quad (2)$$

Keterangan :
S : Himpunan kasus
n : Banyaknya partisi S
pi : probabilitas yang didapat dari kasus i dibagi total kasus.

2.3. RapidMiner

RapidMiner adalah salah satu software untuk pengolahan data mining.Aplikasi RapidMiner bersifat open source (terbuka).RapidMiner dibuat menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. Pekerjaan yang dilakukan oleh RapidMiner text mining adalah berkisar dengan analisis teks, mengekstrak pola-pola dari data set yang besar dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan database. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi bermutu tertinggi dari teks yang diolah.

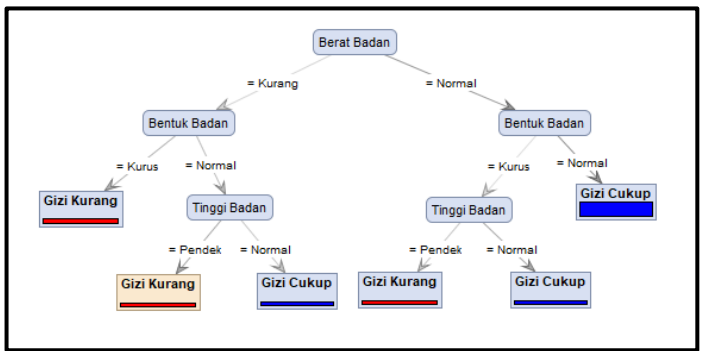
3. Hasil Dan Pembahasan

Dengan menggunakan 3 kriteria antara lain : Berat badan (C1), Tinggi Badan (C2), dan Bentuk Badan (C3) kami memperoleh data dari salah seorang yang bertugas merangkap laporan status gizi balita pada posyandu Kenanga. Beliau memberikan kami 40 data status gizi balita . Data diakumulasi dengan menggunakan bantuan *software microsoft excel*.

Tabel 1. Data Akumulasi

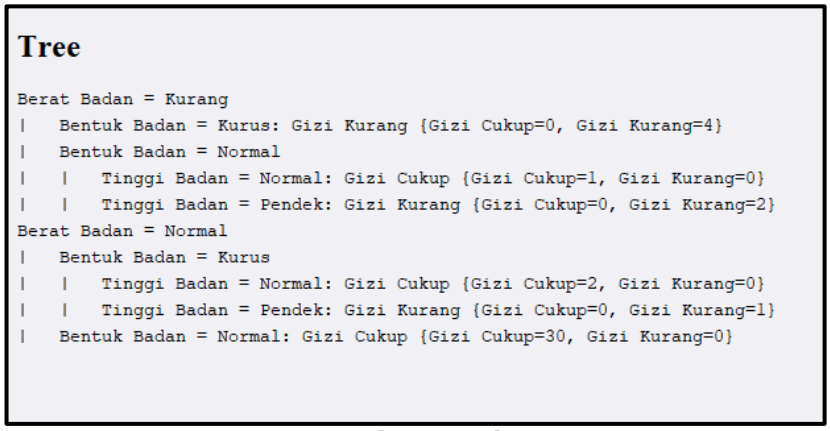
Berat Badan	Tinggi Badan	Bentuk Badan	Status Gizi
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Pendek	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Pendek	Normal	Gizi Cukup
Normal	Pendek	Normal	Gizi Cukup
Kurang	Pendek	Normal	Gizi Kurang
Normal	Normal	Normal	Gizi Cukup
Normal	Pendek	Kurus	Gizi Kurang
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Kurang	Normal	Kurus	Gizi Kurang
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup
Normal	Normal	Normal	Gizi Cukup

Pada tahap selanjutnya data akan di proses menggunakan bantuan Tools RapidMiner maka data akan terlihat pohon keputusan, dan rulenya. Setelah dilakukan perhitungan metode algoritma C4.5 pada RapidMiner, maka terbentuk pohon keputusan seperti gambar di bawah ini :



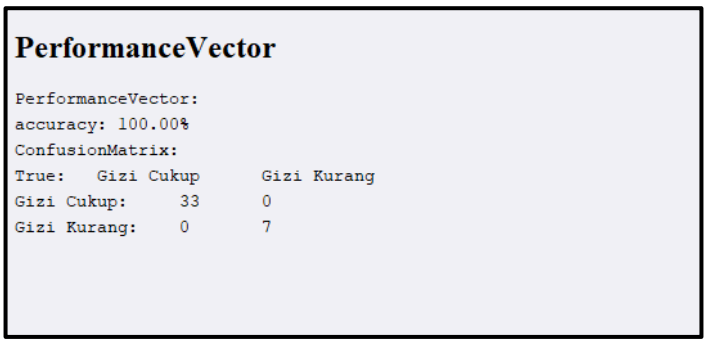
Gambar 1. *Decision Tree*

Pohon Keputusan diatas menghasilkan rule yang akan diimplementasikan pada program. Rule yang dihasilkan adalah sebagai berikut:



Gambar 2. *Rule*

Berdasarkan gambar 2 dapat dijelaskan bahwa dari 3 atribut penilaian terhadap status gizi balita di posyandu kenanga kabupaten Simalungun (Berat badan (C1), Tinggi Badan (C2), dan Bentuk Badan (C3)), diperoleh atribut Berat Badan(C1) dan Bentuk Badan (C3) menjadi atribut yang paling berpengaruh terhadap status gizi balita. Dalam hal ini juga kami menguji kebenaran data yang digunakan, apakah metode klasifikasi C4.5 dapat diterapkan pada status gizi balita dengan bantuan software RapidMiner. Hasil perhitungan software RapidMiner menunjukkan akurasi menggunakan tools performance adalah 100%. Hal ini membuktikan bahwa metode C4.5 dapat digunakan untuk menilai status gizi balita.



Gambar 3. *Performance*

4. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa algoritma C4.5 dapat diterapkan dalam menentukan status gizi balita di posyandu Kenanga kabupaten Simalungun. Dari 3 (tiga) atribut antara lain : Berat badan (C1), Tinggi Badan (C2), dan Bentuk Badan (C3) diperoleh atribut Berat Badan(C1) dan Bentuk Badan (C3) menjadi atribut yang paling berpengaruh terhadap status gizi balita. Dalam hal ini juga kami menguji kebenaran data yang digunakan, apakah metode klasifikasi C4.5 dapat diterapkan pada status gizi balita dengan bantuan software RapidMiner. Hasil perhitungan software RapidMiner menunjukkan akurasi menggunakan tools performance adalah 100%.

Daftar Pustaka

[1] M. Widyastuti, A. G. Fepdiani Simanjuntak, D. Hartama, A. P. Windarto, and A. Wanto, “Classification Model C.45 on Determining the Quality of Custumer

Service in Bank BTN Pematangsiantar Branch," *J. Phys. Conf. Ser.*, vol. 1255, no. 012002, pp. 1–6, 2019, doi: 10.1088/1742-6596/1255/1/012002.

[2] R. W. Sari, A. Wanto, and A. P. Windarto, "Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 224–230, 2018, doi: 10.30865/komik.v2i1.930.

[3] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.

[4] A. P. Windarto *et al.*, "Analysis of the K-Means Algorithm on Clean Water Customers Based on the Province," *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012001.

[5] W. Katrina, H. J. Damanik, F. Parhusip, D. Hartama, A. P. Windarto, and A. Wanto, "C.45 Classification Rules Model for Determining Students Level of Understanding of the Subject," *J. Phys. Conf. Ser.*, vol. 1255, no. 012005, pp. 1–7, 2019, doi: 10.1088/1742-6596/1255/1/012005.

[6] A. Raditya, *Implementasi Data Mining Classification untuk Mencari Pola Prediksi Hujan dengan Menggunakan Algoritma C4.5*. 2015.

[7] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," *Ultim. Vol. VI, No. 1 | Juni 2014*, vol. VI, no. 1, pp. 15–20, 2014.

[8] A. Nurzahputra and M. A. Muslim, "Peningkatan Akurasi Pada Algoritma C4.5 Menggunakan Adaboost Untuk Meminimalkan Resiko Kredit," *Pros. SNATIF ke-4*, pp. 243–247, 2017.