



Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita

Hajar Izzatul Islam¹, Muhamad Khandava Mulyadien², Ultach Enri³

^{1,2,3}Universitas Singaperbangsa Karawang

Received: 21 Juni 2022

Revised: 24 Juni 2022

Accepted: 26 Juni 2022

Abstract

Posyandu in Dawuan Barat determines the nutritional status of children by looking at the growth chart in KIA and calculating the z-score manually and then matching the results to the category table and threshold, this takes a long time and is at risk of being inaccurate. The formulation of the problem in this study is how to do classification of data mining to determine the nutritional status of toddlers and how the evaluation results from the classification model. This study uses the C4.5 algorithm with the CRISP-DM methodology to classify the nutritional status of toddlers and uses a confusion matrix to determine the accuracy, precision, recall and f1-score values of the classification model and then the model is implemented into an application. The evaluation results of the 3 model scenarios in this study stated that scenario 1 produced the best performance among other models with 90% accuracy and 87% value of precision, recall and f1-score.

Keywords: Data Mining, C4.5, Toddlers Nutritional Status, CRISP-DM

(*) Corresponding Author: Hajarizzatul@gmail.com

How to Cite: Islam, H., Mulyadien, M., & Enri, U. (2022). Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita. *Jurnal Ilmiah Wahana Pendidikan*, 8(10), 116-125.
<https://doi.org/10.5281/zenodo.6791722>

PENDAHULUAN

Kelompok usia anak balita dikatakan sebagai “usia emas” karena pada usia ini anak sedang membangun pondasi untuk tubuhnya di masa mendatang. Pada usia ini anak sangat rentan terhadap permasalahan gizi karena anak sangat memerlukan asupan gizi yang cukup dalam masa tumbuh kembang (Marsita, 2018). Kekurangan gizi pada anak dapat menyebabkan *stunting*, *wasting*, mudah terkena penyakit serta penurunan kecerdasan otak dan mental (Marniati, 2021). Sehingga pengecekan status gizi balita sangat perlu dilakukan untuk memantau kesehatan pada balita guna menghindari permasalahan terkait gizi pada balita. Kesehatan pada balita selain dipengaruhi asupan gizi juga dipengaruhi kebersihan lingkungan, pola asuh dan kondisi ibu (Bappenas, 2018).

Berdasarkan Riset Kesehatan Dasar (Riskesdas) tahun 2018 menyatakan 17,7% balita mengalami masalah gizi dengan rincian 3,9% mengalami gizi buruk dan 13,8% mengalami gizi kurang. Untuk menangani permasalahan *stunting* di Indonesia, pemerintah membagi daerah-daerah berdasarkan tingginya kasus *stunting*. Kabupaten Karawang termasuk ke dalam 100 Kota/Kabupaten prioritas *stunting* sejak tahun 2018 berdasarkan data Riskesdas tahun 2013 dengan prevalensi *stunting* sebesar 34,87% (Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K), 2017). Menurut Dinas Kesehatan Karawang (2021), pada tahun 2020 prevalensi *stunting* di Kecamatan Cikampek adalah sebesar 2,3% ini



menjadikan Kecamatan Cikampek termasuk kecamatan dengan prevalensi *stunting* tertinggi urutan ke-13 dari 30 kecamatan yang ada di Kabupaten Karawang. Desa Dawuan Barat adalah salah satu desa yang terletak di Kecamatan Cikampek, di Desa ini juga terjadi kenaikan kasus *stunting* pada tahun 2020. Prevalensi *stunting* di Desa Dawuan Barat pada tahun 2019 adalah sebesar 2,54% kemudian pada tahun 2020 menjadi sebesar 5,53%. Kenaikan kasus *stunting* ini disebabkan oleh pandemi yang terjadi pada tahun 2020. Akibat pandemi kegiatan masyarakat terganggu, harga-harga yang meningkat dan tidak sedikit orang yang mengalami PHK, kejadian ini menyebabkan sulitnya memenuhi kebutuhan gizi balita dan menjaga lingkungan sekitar tetap higienis. Salah satu kegiatan yang terganggu akibat pandemi adalah Posyandu. Posyandu Desa Dawuan Barat sempat terhenti dan tidak dapat berjalan secara maksimal dengan jam operasional yang terbatas dan jumlah pengunjung yang dibatasi. Hal ini menyebabkan balita dan ibu hamil tidak terpantau secara rutin seperti sebelumnya.

Posyandu Desa Dawuan Barat menentukan status gizi dengan perhitungan z-score secara manual kemudian dicocokkan dengan tabel kategori dan ambang batas pada Peraturan Menteri Kesehatan Republik Indonesia. Dengan jumlah balita yang semakin bertambah dan kasus *stunting* yang meningkat membuat proses penentuan z-score tidak dapat langsung selesai dengan cepat. Terlebih lagi, tidak semua petugas Posyandu Desa Dawuan Barat mampu melakukan perhitungan untuk menentukan status gizi balita. Perhitungan yang dilakukan secara manual juga dapat berisiko menghasilkan hasil yang tidak akurat.

Data mining adalah proses mengekstraksi data dari sekumpulan data untuk menemukan sebuah pola hubungan antar data yang dapat digunakan sebagai informasi untuk menyelesaikan masalah tertentu (Riani et al., 2019). Salah satu teknik *data mining* adalah klasifikasi. Teknik klasifikasi digunakan untuk mengelompokkan atau mengklasifikasi ke dalam suatu kelas berdasarkan label tertentu (Wanto & Anjar, 2020). Penelitian yang dilakukan (Wahyudi et al., 2021) dalam klasifikasi status gizi balita menjadi gizi baik, gizi kurang, gizi buruk dan obesitas menggunakan algoritma K-Nearest Neighbors (KNN) menghasilkan nilai akurasi sebesar 88%. Penelitian oleh (Ridwan et al., 2018) menggunakan algoritma Naïve Bayes dengan tambahan Adaboost untuk mengklasifikasikan status gizi anak menghasilkan nilai akurasi sebesar 88,84%. Kemudian penelitian (Wahyudin, 2020) melakukan klasifikasi status gizi menjadi 5 kelas dengan algoritma C4.5 menghasilkan tingkat akurasi sebesar 90,23%. Penelitian (Molina & Soetanto, 2021) membandingkan algoritma Naïve Bayes dan C4.5 dalam klasifikasi status gizi menghasilkan algoritma C4.5 lebih unggul dengan nilai akurasi sebesar 80,04%.

Berdasarkan permasalahan yang ada dan penelitian sebelumnya, penelitian ini akan berfokus untuk mengklasifikasikan status gizi balita berdasarkan indeks Berat Badan menurut Tinggi Badan (BB/TB) sesuai standar antropometri pada Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020. Menggunakan teknik klasifikasi *data mining*, penelitian ini menggunakan algoritma C4.5 dan CRISP-DM sebagai metodologi penelitian untuk mengklasifikasikan status gizi balita berdasarkan indeks BB/TB. Indeks BB/TB mengklasifikasikan status gizi balita menjadi 6 kategori yaitu gizi normal, gizi kurang, gizi buruk, risiko gizi lebih, gizi lebih dan obesitas.

METODOLOGI PENELITIAN

Cross Industry Standard Process for Data Mining (CRISP-DM) adalah standar model proses yang sering digunakan dalam *data mining*. CRISP-DM terbagi menjadi 6 tahapan yaitu (Schröer et al., 2021):

1. *Business Understanding*

Pada tahap ini bertujuan untuk memahami situasi dan permasalahan yang ada pada objek untuk menentukan tujuan.

2. *Data Understanding*

Pada tahap ini bertujuan untuk melakukan pengumpulan dan eksplorasi data untuk melihat kualitas data. Eksplorasi data dilakukan dengan pengecekan *missing value*, *duplicate data* dan *outliers* pada data yang akan digunakan.

3. *Data Preparation*

Pada tahap ini dilakukan pembersihan data dari *missing value*, *duplicate data* dan *outliers* pada data yang akan digunakan. Setelah itu dilakukan *data selection* yaitu menghapus atribut yang tidak diperlukan untuk tahap pembuatan model klasifikasi. Kemudian dilakukan *data transformation* yaitu mengubah data nominal menjadi data numerik dan melakukan normalisasi pada atribut tertentu.

4. *Modeling*

Pada tahap ini dilakukan pemodelan terhadap data yang ada. Pembuatan model dilakukan menggunakan algoritma C4.5. Model dibuat dengan 3 skenario dengan jumlah *data training* dan *data testing* yang berbeda. Model skenario 1 menggunakan 80% *data training* dan 20% *data testing*. Model skenario 2 menggunakan 70% *data training* dan 30% *data testing*. Model skenario 3 menggunakan 50% *data training* dan 50% *data testing*.

5. *Evaluation*

Pada tahap ini dilakukan evaluasi dari setiap model skenario yang telah dibuat. Menggunakan *confusion matrix* dapat diketahui nilai *true positive*, *true negative*, *false positive* dan *false negative* pada suatu model. Hasil dari evaluasi setiap model skenario berupa nilai *accuracy*, *precision*, *recall* dan *f1-score*. Model skenario dengan kinerja terbaik akan diimplementasikan ke dalam sebuah aplikasi dan di-*deploy* ke dalam *cloud platform*.

6. *Deployment*

Pada tahap ini model skenario akan diimplementasikan ke dalam sebuah aplikasi berbasis *web* untuk memudahkan penggunaan petugas Posyandu dalam menentukan status gizi balita secara cepat dan tepat. Aplikasi berbasis *web* tersebut di-*deploy* ke dalam sebuah *cloud platform* sehingga dapat diakses dimana saja dan kapan saja selama perangkat terhubung dengan jaringan internet.

KAJIAN TEORI

Data Mining

Data mining adalah proses mengekstraksi data dari sekumpulan data untuk menemukan sebuah pola hubungan antar data yang dapat digunakan sebagai informasi untuk menyelesaikan masalah tertentu (Riani et al., 2019). Berdasarkan

fungsionalitas data mining dibagi menjadi 6 yaitu klasifikasi, *clustering*, regresi, asosiasi, deteksi anomali dan *summarization*. Klasifikasi digunakan untuk mengelompokkan data ke dalam suatu kelas berdasarkan label tertentu. *Clustering* digunakan untuk mengelompokkan data yang belum memiliki label ke dalam suatu kelompok berdasarkan kesamaan karakteristik tertentu. Regresi digunakan untuk membuat permodelan dari suatu data yang berguna untuk memprediksi sesuatu. Asosiasi digunakan untuk menemukan korelasi dalam sekumpulan item. Deteksi anomali digunakan untuk mendeteksi data *abnormal* dalam suatu kumpulan data yang besar. *Summarization* digunakan untuk visualisasi data sebagai pendukung dari suatu informasi (Wanto & Anjar, 2020)

Algoritma C4.5

Algoritma C4.5 juga dikenal sebagai *decision tree*. Algoritma ini adalah salah satu algoritma yang biasa digunakan dalam teknik klasifikasi *data mining*. Algoritma C4.5 menghasilkan sebuah pohon keputusan atau *decision tree* yang digunakan sebagai acuan untuk menentukan label dari *input* yang diberikan (Lukhayu Pritalia, 2018). Tahapan dalam algoritma C4.5 yang pertama adalah menghitung nilai *entropy* total dari setiap label, kemudian menghitung nilai *entropy* dari masing-masing atribut. Setelah mendapatkan nilai *entropy* dari masing-masing atribut dapat diketahui nilai *gain*. Nilai *gain* dari setiap atribut inilah yang menentukan *node* dalam pohon keputusan (Saleh et al., 2020).

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

Keterangan:

S = Himpunan kasus

n = Jumlah partisi S

p_i = Proporsi dari S_i dan S

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

Keterangan:

S = Himpunan kasus

A = Atribut

n = Jumlah partisi atribut A

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

$|S|$ = Jumlah kasus dalam S

Python

Python adalah salah satu bahasa pemrograman komputer yang berfokus pada penulisan *code* yang mudah dipahami. Secara umum Python berbentuk pemrograman berorientasi objek, pemrograman *imperative* dan pemrograman fungsional (Enterprise, 2019). Python dilengkapi dengan sekumpulan modul berisi kode-kode yang disebut *library*. Adanya *library* Python berguna untuk mengatasi masalah terkait *programming* dalam kehidupan sehari-hari (Python, n.d.). Python memiliki *editor* bawaan yaitu *Integrated Development and Learning Environment* (IDLE) namun Python tetap dapat berjalan di *editor* lain dari pihak ketiga seperti Visual Studio Code, Sublime, dan lain-lain (Enterprise, 2019).

Confusion Matrix

Confusion matrix digunakan untuk merepresentasikan kinerja model dalam 4 parameter yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP) dan *false negative* (FN). Keempat parameter ini digunakan sebagai landasan dalam rumus untuk menentukan *performance metrics* seperti *accuracy*, *precision*, *recall* dan *f1-score*. *Accuracy* adalah nilai prediksi benar atas keseluruhan data, *precision* adalah nilai prediksi benar positif dibandingkan keseluruhan hasil yang diprediksi positif, *recall* adalah nilai prediksi benar positif dibandingkan keseluruhan data yang benar positif, *f1-score* adalah rata-rata perbandingan dari nilai *precision* dan *recall* (Nugroho, 2019).

$$Accuracy = \frac{TP+TN}{TP + TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1-Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (6)$$

Flask

Flask adalah *web framework* yang dikategorikan sebagai *micro framework* karena sebagian fungsi dan komponennya telah dimuat oleh *third party*. Flask difungsikan untuk membuat *core* yang sederhana untuk aplikasi dan dapat dengan mudah dimanipulasi. Flask dikatakan memiliki fleksibilitas dan skalabilitas yang tinggi dibandingkan *framework* lain (Irsyad, 2018).

Heroku

Heroku adalah sebuah *cloud platform as a service* (PaaS) yang memungkinkan pengguna untuk membangun, menjalankan dan mengoperasikan sebuah aplikasi secara keseluruhan (Heroku, n.d.). Sebagai PaaS, Heroku memungkinkan pengguna untuk hanya berfokus pada aplikasi tanpa mengelola infrastruktur atau *server* dan *hardware* (Suprijono et al., n.d.), untuk men-*deploy* aplikasi, pengguna hanya perlu melakukan konfigurasi dengan *platform* yang didukung oleh Heroku (Syaputro & Widya, 2022).

HASIL DAN PEMBAHASAN

Business Understanding

Pada tahap ini diketahui permasalahan dalam objek untuk menentukan tujuan. Objek pada penelitian ini adalah anak usia 0-24 bulan di Posyandu Desa Dawuan Barat. Permasalahan yang ada pada objek adalah perhitungan z-score untuk menentukan status gizi balita masih dilakukan secara manual dan tidak semua petugas Posyandu dapat menentukan status gizi balita dengan perhitungan z-score.

Maka tujuan dalam penelitian ini adalah untuk melakukan klasifikasi status gizi balita secara cepat dan tepat menggunakan algoritma C4.5.

Data Understanding

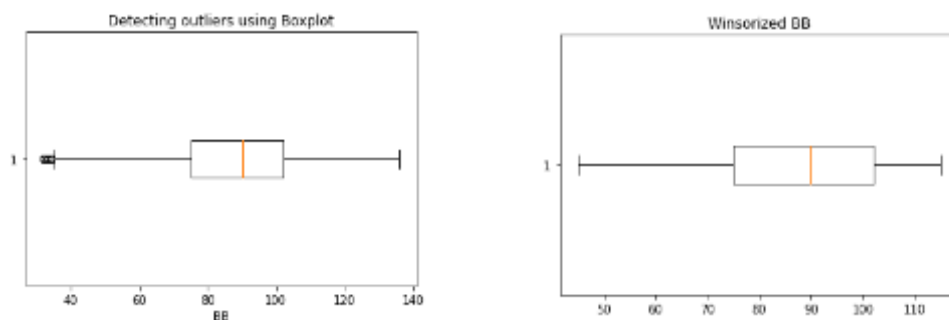
Hasil dari tahap ini adalah pengumpulan dan eksplorasi data. Data yang digunakan dalam penelitian ini berasal dari 13 Posyandu Desa Dawuan Barat dengan jumlah *records* sebanyak 1340 dan 15 atribut. Eksplorasi data untuk menemukan *missing value* dan *duplicated value* dilakukan menggunakan bantuan library Pandas. Pada *dataset* yang digunakan tidak terdapat *missing value* dan *duplicated value*, selanjutnya dilakukan pengecekan *outlier* menggunakan rumus IQR, hasilnya ditemukan sejumlah *outlier* pada atribut BB.

NAMA	RT/RTW	JK	BLN/TGL/THN	UMUR	TB	BB	GIZI	KONSELING	KUNJUNGAN RUMAH	KEPEMILIKAN AIR BERSIH	KEPEMILIHAN JAMBAN	AKTA LAHIR	JAKES	PAUD
Ampari	1/2	P	1/25/2021	11	89.0	8.6	N	Y	T	Y	Y	Y	Y	Y
Astari	2/1	L	4/25/2021	9	67.5	9.0	N	Y	T	Y	Y	Y	Y	Y
Azi-Zam	1/2	L	5/20/2021	8	66.2	8.2	N	Y	T	Y	Y	Y	Y	Y
Abdul P	2/1	L	9/30/2021	4	61.1	8.7	N	T5	T	Y	Y	Y	Y	T5
Kayla	1/2	P	9/30/2021	4	60.0	8.0	N	T5	T	Y	Y	Y	Y	T5

Gambar 1. Raw Dataset

Data Preparation

Pertama-tama dilakukan pembersihan data, dari tahap sebelumnya diketahui bahwa terdapat sejumlah *outlier* pada atribut BB. *Outlier* tersebut kemudian akan ditangani dengan winsorization yaitu proses mengganti sebuah nilai untuk membatasi efek *outlier* dalam sebuah data. Proses winsorization dilakukan menggunakan bantuan library Scipy.



Gambar 2. Visualisasi *outlier* dengan boxplot sebelum dan sesudah dilakukan proses winsorization

Tahap selanjutnya dilakukan *data selection* yaitu menghapus atribut yang tidak dibutuhkan untuk pembuatan model. Pada *dataset* yang digunakan hanya atribut “JK”, “TB”, “BB” dan “GIZI”. Kemudian mengganti *value* pada atribut “JK” dan “GIZI”. Pada atribut “JK” *value* “L” diganti menjadi “0” dan “P” menjadi “1”. Untuk atribut “GIZI” *value* N (gizi normal) diubah menjadi “0”, kemudian “K” (gizi kurang) menjadi “1”, “B” (gizi buruk) menjadi “2”, “RGL” (risiko gizi lebih) menjadi “3”, “GL” (gizi lebih) menjadi “4” dan “O” (obesitas) menjadi “5”.

JK	UMUR	TB	BB	GIZI
1	0.022272	0.026143	0.023530	0
0	0.018223	0.025574	0.025641	0
0	0.016198	0.025840	0.024133	0
0	0.008099	0.023150	0.019910	0
1	0.008099	0.022733	0.017497	0

Gambar 3. *Dataset* setelah dilakukan *data transformation*

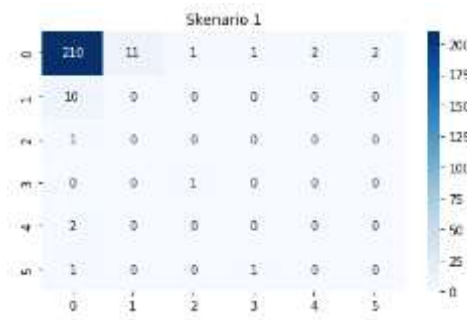
Modeling

Pada tahap ini dilakukan proses pemodelan algoritma C4.5 dengan bantuan *library* Sklearn. Pembuatan model dilakukan dengan 3 skenario. Menggunakan fungsi `train_test_split` dari *library* Sklearn data dibagi menjadi *data training* dan *data testing* berdasarkan skenario yang ada. Kemudian mendefinisikan atribut “Umur”, “JK”, “TB” dan “BB” sebagai *feature* dan atribut “GIZI” sebagai *target*. Setelah itu membuat objek dari *class* `DecisionTreeClassifier` dari *library* Sklearn sebagai model algoritma dan *mem-fit data training* ke dalam objek tersebut sebagai proses *training* model. Setelah itu model dari *data training* diuji dengan *data testing* yang sebelumnya telah dipisahkan.

Evaluation

Evaluasi menggunakan *confusion matrix* untuk model skenario 1 terdapat pada Gambar 4 diketahui TP pada *class* 0 (gizi normal) adalah sebanyak 210 data, sementara *class* 1-5 adalah sebanyak 0. Dengan kata lain, model skenario 1 tidak mampu memprediksi benar atas *class* 1-5. Hal ini terjadi karena *dataset* yang digunakan tidak seimbang atau disebut *imbalanced data*. Jika dilihat menggunakan *classification report* seperti pada Gambar 5 diketahui dari 243 *data testing* jumlah data untuk *class* 0 adalah sebanyak 227 data, untuk *class* 1 (gizi kurang) adalah sebanyak 10 data, *class* 2 (gizi buruk) dan *class* 3 (risiko gizi lebih) adalah sebanyak 1 data, dan *class* 4 (gizi lebih) dan *class* 5 (obesitas) adalah sebanyak 2 data. Untuk mengatasi hal tersebut digunakan teknik SMOTETomek yaitu gabungan teknik SMOTE dan Tomek-Links. Setelah diterapkan teknik SMOTETomek, model skenario 1 memiliki data yang lebih seimbang seperti pada Gambar 6 jumlah *data testing* menjadi 1424 dengan jumlah data pada *class* 0 sebanyak 232, pada *class* 1 sebanyak 243, *class* 2 sebanyak 248, *class* 3 sebanyak 233, *class* 4 sebanyak 240, dan *class* 5 sebanyak 228.

Pada model skenario 2 dan 3 juga ditemukan permasalahan serupa yaitu data yang tidak seimbang. Sehingga pada model skenario 2 dan 3 juga diterapkan teknik SMOTETomek untuk mengatasi jumlah data yang tidak seimbang pada setiap *class*. Hasilnya dari evaluasi seluruh model skenario seperti pada tabel 1.



Gambar 4. Confusion matrix model skenario 1

	precision	recall	f1-score	support
0	0.93	0.92	0.93	227
1	0.00	0.00	0.00	10
2	0.00	0.00	0.00	1
3	0.00	0.00	0.00	1
4	0.00	0.00	0.00	2
5	0.00	0.00	0.00	2
accuracy			0.86	243
macro avg	0.16	0.15	0.15	243
weighted avg	0.87	0.86	0.87	243

Gambar 5. Classification report model skenario 1

	precision	recall	f1-score	support
0	0.87	0.87	0.87	232
1	0.83	0.83	0.83	243
2	0.93	0.92	0.93	248
3	0.91	0.91	0.91	233
4	0.94	0.95	0.95	240
5	0.89	0.90	0.90	228
accuracy			0.90	1424
macro avg	0.90	0.90	0.90	1424
weighted avg	0.90	0.90	0.90	1424

Gambar 6. Classification report model skenario 1 setelah dilakukan teknik SMOTETomek

Tabel 1. Hasil evaluasi seluruh model skenario

Model Klasifikasi		<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>Imbalanced data</i>	Skenario 1	86%	93%	92%	98%
	Skenario 2	85%	93%	92%	98%
	Skenario 3	84%	92%	91%	96%
<i>Balanced data dengan SMOTETomek</i>	Skenario 1	90%	87%	87%	87%
	Skenario 2	89%	86%	89%	87%
	Skenario 3	88%	82%	89%	85%

Deployment

Pada tahap sebelumnya, diketahui model skenario 1 memiliki performa terbaik diantara model skenario lainnya. Model skenario 1 tersebut kemudian diimplementasikan ke dalam aplikasi berbasis *web* untuk memudahkan penggunaan

petugas Posyandu dalam menentukan status gizi balita. Pada tahap ini dibuat laman sederhana menggunakan HTML dan CSS kemudian model klasifikasi diubah menjadi format pickle dan dimasukkan ke dalam Flask agar API model dapat diintegrasikan dengan laman *web* yang telah dibuat. Kemudian setelah API dan laman *web* telah terintegrasi, aplikasi tersebut di-*deploy* ke dalam *cloud platform* Heroku agar dapat diakses melalui jaringan internet.



Gambar 7. Laman *web* sebelum dan sesudah ditekan tombol submit

KESIMPULAN

Pada penelitian ini diketahui bahwa model klasifikasi yang telah dibuat menggunakan algoritma C4.5 dapat mengklasifikasikan status gizi balita menjadi 6 kategori berdasarkan indeks BB/TB yaitu gizi normal, gizi kurang, gizi buruk, risiko gizi lebih, gizi lebih dan obesitas. Pada evaluasi model diketahui *dataset* yang digunakan memiliki *class* yang tidak seimbang satu sama lain, namun *imbalanced data* tersebut dapat diatasi dengan teknik SMOTETomek sehingga jumlah data pada setiap *class* menjadi seimbang. Model skenario 1 menghasilkan kinerja terbaik diantara model skenario lainnya yaitu dengan nilai *accuracy* sebesar 90% dan nilai *precision*, *recall* dan *f1-score* sebesar 87%. Model skenario tersebut diimplementasikan ke dalam aplikasi berbasis *web* untuk memudahkan penggunaan petugas Posyandu dalam menentukan status gizi balita.

DAFTAR PUSTAKA

- Bappenas. (2018). Pedoman Pelaksanaan Intervensi Penurunan Stunting Terintegrasi Di Kabupaten/ Kota. Kementerian Perencanaan Pembangunan Nasional.
- Enterprise, J. (2019). Python untuk Programmer Pemula. Elex Media Komputindo.
- Heroku. (n.d.). *Heroku*. Heroku.Com. Retrieved April 13, 2022, from <https://www.heroku.com/>
- Irsyad, R. (2018). Penggunaan Flask untuk Pemula. 1–4.
- Lukhayu Pritalia, G. (2018). Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce. *Indonesian Journal of Information Systems, 1*(1). <https://doi.org/10.24002/ijis.v1i1.1727>

- Marniati. (2021). *Manajemen Pemanfaat Posyandu Balita* (T. Hidayati, Ed.). Pena Persada.
- Marsita, M. (2018). Implementasi algoritma decision tree C4.5 untuk mengidentifikasi gizi balita berdasarkan indeks antropometri: studi kasus posyandu Seruni. UIN Sunan Gunung Djati.
- Molina, J., & Soetanto, H. (2021). Optimalisasi Kinerja Hasil Komparasi Algoritma C4.5 dan Naive Bayes Menggunakan Particle Swarm Optimization untuk Deteksi Dini Penentuan Status Gizi pada Balita. Deepublish.
- Nugroho, K. S. (2019). Confusion Matrix untuk Evaluasi Model pada Supervised Learning. Medium.Com.
- Python. (n.d.). *The Python Standard Library*. Docs.Python.Org. Retrieved April 13, 2022, from <https://docs.python.org/3/library/>
- Riani, A., Susianto, Y., & Rahman, N. (2019). Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode Naive Bayes. *Journal of Innovation Information Technology and Application (JINITA)*, 1(01). <https://doi.org/10.35970/jinita.v1i01.64>
- Ridwan, A., Andono, P., & Supriyanto, C. (2018). Optimasi Klasifikasi Status Gizi Balita Berdasarkan Indeks Antropometri Menggunakan Algoritma Naive Bayes Classification Adaboost. *Jurnal Teknologi Informasi*. <http://research.pps.dinus.ac.id/index.php/Cyberku/article/download/76/72>
- Saleh, H., Informatika, J. T., & Komputer, F. I. (2020). Analisa Faktor Penyebab Stunting Menggunakan Algoritma C4 . 5. *Scientico: Computer Science and Informatics Journal*, 3(1).
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181. <https://doi.org/10.1016/j.procs.2021.01.199>
- Suprijono, S. A., Setianingsih, C., & Saputra, R. E. (n.d.). Deteksi Tinggi Rendah Gelombang Air Laut Dengan Multisensor Menggunakan Algoritma Fuzzy Sukamoto High And Low Sea Waves Detection With Multisensor Using Fuzzy Sukamoto Algoritm.
- Syaputro, & Widya, R. (2022). Pemanfaatan Bot Api Telegram Pada Sistem Manajemen (Studi Kasus UKM Informatika dan Komputer). *Digilab STMIK AKAKOM*.
- Tim Nasional Percepatan Penanggulangan Kemiskinan. (2017). *100 Kabupaten/Kota Prioritas untuk Intervensi Anak Kerdil (Stunting)*.
- Wahyudi, R., Orisa, M., & Vendyansyah, N. (2021). Penerapan Algoritma K-Nearest Neighbors Pada Klasifikasi Penentuan Gizi Balita (Studi Kasus Di Posyandu Desa Bluto). *Jati (Jurnal Mahasiswa Teknik Informatika)*, 5(2). <https://doi.org/10.36040/jati.v5i2.3738>
- Wahyudin, W. (2020). Klasifikasi Stunting Balita Menggunakan Naive Bayes Dengan Seleksi Fitur Forward Selection. *Jurnal Bisnis Digital Dan Sistem Informasi*, 1. <https://ejr.stikesmuhkudus.ac.id/index.php/jikoma/article/viewFile/1220/760>
- Wanto & Anjar, D. (2020). *Data Mining: Algoritma dan Implementasi - Books*. In *Yayasan kita menulis*.