



Analisis Rekam Medis dengan Metode Data Mining untuk Memprediksi Faktor Risiko Stunting dalam Kesehatan Masyarakat

Nurril Cholifatul Izza^{1*}, Ade Irma Rizmayanti²

^{1,2} Program Studi Manajemen Informasi Kesehatan, Politeknik Karya Husada
nurril.cholifatul@gmail.com, ade.irma.r@gmail.com

Keywords:

*Data Mining,
Medical records,
Public Health,
Random forest,
Stunting*

ABSTRACT

Stunting is a serious condition caused by prolonged malnutrition and recurrent infections, hindering children's physical growth and cognitive development. This condition affects millions of children worldwide, especially in developing countries, and has long-term impacts on their health, education and economic potential. In this study, medical record data on stunting incidents was analyzed using 121,000 datasets. The data is classified into four nutritional statuses: normal, stunted, high, and severely stunted, based on four main variables, namely age, gender, height as a label, and nutritional status as a target variable. The Random forest model was used to test the accuracy of the model in predicting children's nutritional status. The analysis results show that this model is very accurate, with an average accuracy value of 0.9990 after 10 times cross-validation. These findings show that Random forest is an effective model in predicting children's nutritional status and can be used to help better target health interventions. This research also highlights the importance of using comprehensive medical record data in monitoring the health of pediatric populations.

Kata Kunci

*Data Mining,
Kesehatan Masyarakat,
Random Forest,
Rekam Medis,
Stunting*

ABSTRAK

Stunting adalah kondisi serius yang disebabkan oleh malnutrisi berkepanjangan dan infeksi berulang, menghambat pertumbuhan fisik dan perkembangan kognitif anak-anak. Kondisi ini mempengaruhi jutaan anak di seluruh dunia, terutama di negara-negara berkembang, dan memiliki dampak jangka panjang terhadap kesehatan, pendidikan, dan potensi ekonomi mereka. Dalam penelitian ini, data rekam medis terhadap kejadian stunting dianalisis dengan menggunakan 121.000 dataset. Data tersebut diklasifikasikan ke dalam empat status gizi: normal, stunted, tinggi, dan severely stunted, berdasarkan empat variabel utama yaitu usia, jenis kelamin, tinggi badan sebagai label, dan status gizi sebagai variabel target. Model *Random forest* digunakan untuk menguji keakuratan model dalam memprediksi status gizi anak-anak. Hasil analisis menunjukkan bahwa model ini sangat akurat, dengan nilai rata-rata akurasi sebesar 0,9990 setelah dilakukan 10 kali cross-validation. Temuan ini menunjukkan bahwa *Random forest* merupakan model yang efektif dalam memprediksi status gizi anak-anak dan dapat digunakan untuk membantu intervensi kesehatan yang lebih tepat sasaran. Penelitian ini juga menyoroti pentingnya penggunaan data rekam medis yang komprehensif dalam pemantauan kesehatan populasi anak-anak.

Korespondensi Penulis:
Nurril Cholifatul Izza,

Submitted : 21-05-2024; Accepted : 22-05-2024;
Published : 30-05-2024

1. PENDAHULUAN

Stunting adalah kondisi serius yang terjadi akibat malnutrisi berkepanjangan dan infeksi berulang [1], yang menghambat pertumbuhan fisik anak-anak dan dapat mempengaruhi perkembangan kognitif mereka [2]. Menurut Organisasi Kesehatan Dunia (WHO), stunting mempengaruhi lebih dari 150 juta anak di bawah usia lima tahun di seluruh dunia, dengan konsekuensi jangka panjang yang mempengaruhi kesehatan, pendidikan, dan potensi ekonomi mereka di masa dewasa [3]. Di banyak negara berkembang, stunting tidak hanya menjadi masalah kesehatan tetapi juga masalah sosial dan ekonomi [4]. Faktor risiko yang berkontribusi pada stunting sangat beragam, termasuk status gizi ibu, akses terhadap pangan yang bergizi, sanitasi, dan akses ke layanan kesehatan [5]. Dengan kompleksitas ini, pendekatan tradisional seringkali tidak mencukupi untuk mengatasi masalah ini secara efektif [6].

Dalam upaya mengatasi masalah stunting, teknologi informasi dan *big data* menawarkan potensi baru untuk analisis yang mendalam [2]. Data mining, khususnya, memungkinkan para peneliti untuk menganalisis kumpulan data besar dari rekam medis untuk menemukan pola dan hubungan yang tidak terlihat secara manual [7]. Teknik ini dapat membantu mengidentifikasi faktor risiko yang lebih spesifik dan kondisi yang mungkin berkontribusi pada tingginya tingkat stunting dalam populasi tertentu [8].

Penelitian tentang stunting pada jurnal sebelumnya mengkaji faktor-faktor yang berkontribusi terhadap stunting, sebuah kondisi gagal tumbuh pada anak yang terkait dengan kekurangan gizi kronis. Di Indonesia, prevalensi stunting mencapai 21,6% di antara bayi baru lahir, dengan peningkatan kasus yang dikaitkan dengan faktor ekonomi yang mempengaruhi ketersediaan gizi yang adekuat. Penelitian ini menggunakan metode simple random sampling dengan sampel 112 rumah tangga. Faktor-faktor yang dianalisis meliputi pengetahuan ibu tentang ASI eksklusif, sanitasi lingkungan, pemeriksaan kehamilan, dan kunjungan petugas kesehatan. Hasil uji statistik menunjukkan adanya hubungan yang signifikan antara faktor-faktor tersebut dan kejadian stunting, dengan p-value 0,013 yang menunjukkan signifikansi statistik. Penelitian ini menyarankan perlunya edukasi yang lebih baik kepada calon orang tua dan orang tua untuk mengawasi tumbuh kembang anak, dalam rangka mengurangi kasus stunting [1].

Selanjutnya penelitian oleh [6] ini dari penelitian ini adalah mengeksplorasi hubungan antara pola asuh orang tua dan kejadian stunting pada balita di Kecamatan Kuburaya. Stunting, yang didefinisikan sebagai kegagalan pertumbuhan pada balita yang menyebabkan tinggi badan menurut umur kurang dari -2 SD (nilai z-score), sangat dipengaruhi oleh faktor pola asuh, termasuk pola makan. Dalam penelitian ini, ditemukan bahwa prevalensi stunting pada balita usia 24-59 bulan di daerah tersebut mencapai 40,3%. Menggunakan desain penelitian observasional *case-control*, peneliti mengumpulkan data dari 62 ibu dengan balita yang mengalami stunting dan 62 ibu dengan balita yang memiliki status gizi normal, melalui angket yang menilai pola asuh. Hasil menunjukkan bahwa sebagian besar kasus stunting (85,5%) terkait dengan pola asuh yang buruk. Oleh karena itu, penelitian ini menggarisbawahi pentingnya pola asuh yang baik dalam mencegah stunting, mengingat signifikansi hubungan antara pola asuh orang tua dan status gizi balita.

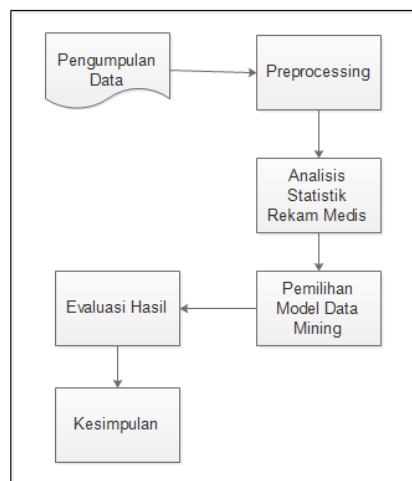
Pada materi yang dibahas dalam penelitian ini dengan menggunakan data mining untuk menganalisis rekam medis dari dataset publik dengan sample 121.000 data. Kumpulan data yang dirancang berdasarkan rumus z-score dari WHO (*World Health Organization*) untuk mendeteksi stunting pada balita (anak di bawah lima tahun). Dataset ini mencakup detail tentang umur, jenis kelamin, tinggi badan, dan status gizi balita. Tujuan dari dataset ini adalah untuk mendukung peneliti, ahli gizi, dan pembuat kebijakan dalam mengidentifikasi dan mengatasi permasalahan stunting pada anak-anak di bawah usia lima tahun dengan mendapatkan hasil klasifikasi pada empat jenis yaitu klasifikasi normal, stunting, tinggi dan sangat terhambat (*severely stunted*) [9]. Tujuannya adalah untuk mengidentifikasi pola yang dapat menginformasikan intervensi yang lebih tepat dan efektif. Dengan mengekstraksi informasi dari data

historis pasien, penelitian ini berharap dapat memberikan wawasan baru yang dapat mengarahkan kebijakan kesehatan publik dan program intervensi yang dirancang untuk mengurangi insiden stunting.

Metodologi yang diterapkan dalam penelitian ini melibatkan algoritma klasifikasi *random forest*, yang bertujuan untuk mengungkap faktor yang berinteraksi dalam kompleksitas kasus stunting. Dengan memahami interaksi ini, dapat dikembangkan strategi yang lebih holistik dan multiaspek dalam penanganan stunting. Melalui analisis data mining, penelitian ini bertujuan untuk memberikan rekomendasi berbasis bukti kepada pemangku kepentingan dalam sektor kesehatan publik. Diharapkan bahwa hasil penelitian ini dapat membantu dalam perancangan program yang lebih efektif untuk mencegah dan mengatasi stunting di masa depan, sehingga memberikan dampak positif pada kesejahteraan anak-anak di seluruh dunia.

2. METODE PENELITIAN

Dalam studi ini, metodologi yang digunakan bertujuan untuk menganalisis dan mengidentifikasi faktor-faktor yang berkontribusi terhadap prevalensi stunting pada balita. Pendekatan ini mengintegrasikan pengumpulan data, analisis statistik, dan penerapan model prediktif yang memanfaatkan dataset "*Stunting Toddler (Balita) Detection*". Berikut gambar 1 metode penelitian.



Gambar 1. Metode Penelitian

2.1 Pengumpulan Data

Pertama, peneliti mengumpulkan data rekam medis dari dataset publik. Dataset ini dibuat dengan menggunakan rumus z-score dari WHO (*World Health Organization*) untuk menentukan kasus stunting pada balita (anak di bawah lima tahun). Terdiri dari 121.000 data, dataset ini meliputi informasi tentang umur, jenis kelamin, tinggi badan, dan status gizi balita dengan empat klasifikasi yaitu 'severely stunting', 'stunting', 'normal', dan 'tinggi' [9].

2.2 Preprocessing

Langkah selanjutnya adalah pembersihan data, yang melibatkan penghapusan entri yang duplikat atau tidak lengkap, serta normalisasi parameter untuk memastikan konsistensi dalam analisis. Hal ini penting untuk meningkatkan keakuratan hasil data mining [10].

2.3 Analisis Statistik Rekam Medis

Untuk mengidentifikasi pola dan tren yang berkaitan dengan stunting pada anak di bawah lima tahun, penelitian akan menerapkan berbagai teknik statistik untuk mengevaluasi distribusi dan hubungan antara variabel seperti umur, jenis kelamin, tinggi badan, dan status gizi. Tujuan utama dari analisis ini adalah untuk mendapatkan wawasan yang dapat membantu dalam penanganan dan pencegahan stunting pada balita. Penelitian ini menggunakan bahasa mesin python dengan *tools* pycaret. PyCaret adalah pustaka pembelajaran mesin *open-source* dan *low-code* dalam Python yang mengotomatisasi alur kerja pembelajaran mesin, Python dirancang untuk memiliki keterbacaan yang tinggi. Bahasa ini sering menggunakan kata kunci bahasa Inggris di mana bahasa lain menggunakan tanda baca, dan memiliki konstruksi sintaksis yang lebih sedikit dibandingkan bahasa lain [11].

2.4 Pemilihan Model Data Mining

Dalam upaya menganalisis dataset "*Stunting Toddler (Balita) Detection*", kami memutuskan untuk menggunakan model *Random forest* dalam pemilihan metode data mining. Pemilihan ini didasarkan pada kemampuan *Random forest* untuk menangani data dengan banyak entri dan variabel [12], seperti dataset yang digunakan pada penelitian ini melibatkan 121.000 data balita. Model ini terkenal akan keakuratannya yang tinggi dan kemampuan untuk mengurangi *overfitting*, yang sangat penting dalam memprediksi stunting pada balita [13]. Dengan menggunakan model *Random forest*, diharapkan mendapatkan pemahaman yang lebih mendalam tentang faktor-faktor yang mempengaruhi stunting serta mengidentifikasi pola yang relevan yang dapat membantu dalam merumuskan strategi intervensi yang efektif.

2.5 Evaluasi Hasil

Model yang dihasilkan dari proses data mining dengan menggunakan model *random forest* divalidasi menggunakan set data tes yang terpisah. Metrik seperti akurasi, presisi, dan recall digunakan untuk mengevaluasi kinerja model. Evaluasi ini penting untuk memastikan bahwa prediksi yang dihasilkan dapat diandalkan dan valid secara statistik [14].

3. HASIL DAN ANALISIS

Pada bagian hasil dan analisis, penelitian ini akan menyajikan temuan dan hasil analisa berdasarkan temuan-temuan statistik dari dataset rekam medis kejadian stunting pada anak 0-60 bulan dan mengeksplorasi seberapa tepat model *random forest* dalam melakukan prediksi terhadap pola-pola hubungan dari data.

3.1 Analisis Dataset

Dataset yang diteliti terdiri dari empat field utama yaitu:

- Umur (Bulan): Menunjukkan usia balita dalam hitungan bulan, yang esensial untuk menilai tahap pertumbuhan mereka dan membandingkannya dengan norma pertumbuhan yang dianggap sehat.
- Jenis Kelamin: Kolom ini membedakan dua kategori, yaitu 'laki-laki' dan 'perempuan'. Jenis kelamin adalah aspek krusial dalam menganalisis tren pertumbuhan dan risiko stunting.
- Tinggi Badan: Diukur dalam centimeter, tinggi badan merupakan parameter kunci untuk evaluasi pertumbuhan fisik balita. Pengukuran ini membantu peneliti mengevaluasi apakah pertumbuhan anak sesuai dengan standar untuk usianya.
- Status Gizi: Terbagi menjadi empat kategori '*severely stunting*', '*stunting*', '*normal*', dan '*tinggi*'. '*Severely stunting*' menggambarkan kondisi yang sangat serius, '*stunting*' untuk keadaan stunting, '*normal*' untuk gizi yang sehat, dan '*tinggi*' untuk pertumbuhan yang lebih dari rata-rata. Klasifikasi ini penting untuk pengidentifikasian dan intervensi cepat pada anak-anak yang berisiko atau menghadapi masalah pertumbuhan.

Tabel 1. Dataset Penelitian

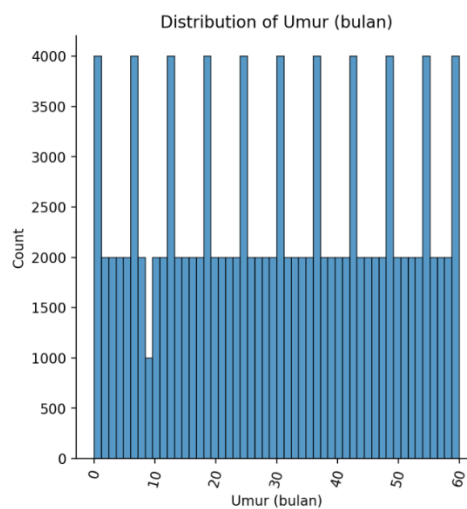
No	Umur (bulan)	Jenis Kelamin	Tinggi Badan (cm)	Status Gizi
0	0	laki-laki	44.59	<i>stunted</i>
1	0	laki-laki	56.71	tinggi
2	0	laki-laki	46.86	normal
3	0	laki-laki	47.51	normal
4	0	laki-laki	42.74	<i>severely stunted</i>
...
120994	60	perempuan	100.60	normal
120995	60	perempuan	98.30	stunted
120996	60	perempuan	121.30	normal
120997	60	perempuan	112.20	normal
120998	60	perempuan	109.80	normal

3.2 Analisis Statistik Rekam Medis

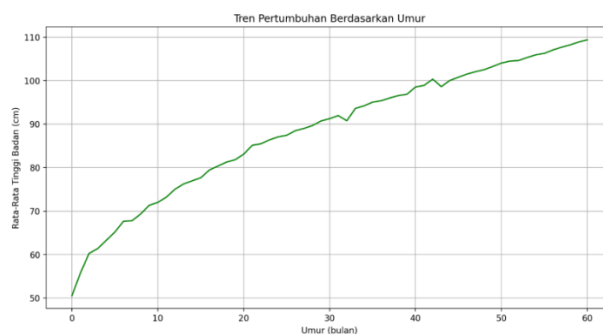
Dalam analisis rekam medis ini, penelitian berfokus pada data yang mencakup usia, jenis kelamin, tinggi badan, dan status gizi balita. Tujuannya adalah untuk mengeksplorasi hubungan antar variabel dan efeknya terhadap status gizi, guna meningkatkan intervensi kesehatan dan strategi pencegahan. Kami akan menggunakan metode statistik untuk mengidentifikasi tren dan pola, serta membahas implikasi temuan ini bagi praktik kesehatan.

a. Distribusi dan tren pertumbuhan berdasarkan usia

Distribusi dan tren pertumbuhan pada dataset yang digunakan pada penelitian ini berdasarkan usia di gambarkan pada gambar 2 dan gambar 3.



Gambar 2. Distribusi Usia



Gambar 3. Tren usia pertumbuhan

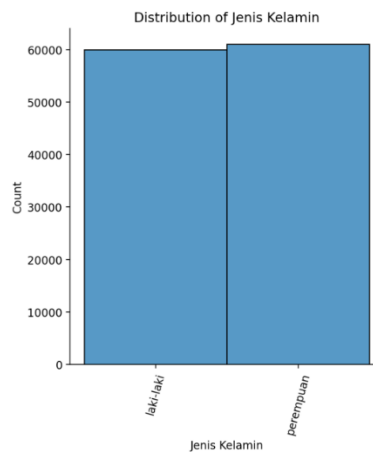
Pada gambar 2 Grafik batang ini menunjukkan distribusi umur dalam bulan untuk dataset stunting. Setiap batang merepresentasikan jumlah balita pada kelompok umur tertentu yang terpisah setiap 10 bulan, mulai dari 0 hingga 60 bulan. Sebagian besar kelompok umur memiliki jumlah yang relatif seragam, dengan jumlah balita di setiap kelompok berkisar sekitar 3500. Selanjutnya melihat tren dari usia untuk mengetahui pola seperti apa yang disajikan pada dataset.

Pada Gambar 3 Grafik mengilustrasikan tren pertumbuhan anak dari usia 0 hingga 60 bulan dengan fokus pada rata-rata tinggi badan dalam cm, menunjukkan peningkatan konsisten sepanjang periode ini dengan pertumbuhan yang lebih cepat di awal kehidupan, khususnya sampai 20 bulan, diikuti oleh penurunan laju pertumbuhan yang masih bertambah secara stabil hingga usia 60 bulan. Grafik ini sangat

bermanfaat dalam pemahaman pola pertumbuhan normal pada anak-anak hingga usia lima tahun, yang esensial untuk praktek pediatri dan pemantauan kesehatan anak agar mereka tumbuh sesuai dengan standar pertumbuhan yang diharapkan.

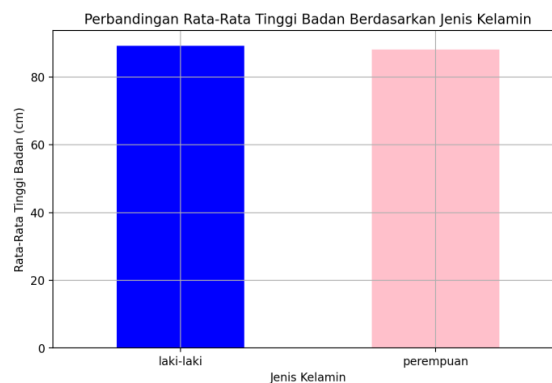
b. Distribusi berdasarkan jenis kelamin

Pada Gambar 4 dapat di lihat bahwa jumlah laki-laki dan perempuan pada dataset yang digunakan sebagai penelitian hampir sama, dengan perempuan sedikit lebih banyak daripada laki-laki. Kedua kolom memiliki jumlah yang sangat serupa, menunjukkan distribusi yang seimbang antara laki-laki dan perempuan.



Gambar 4. Distribusi berdasarkan jenis kelamin

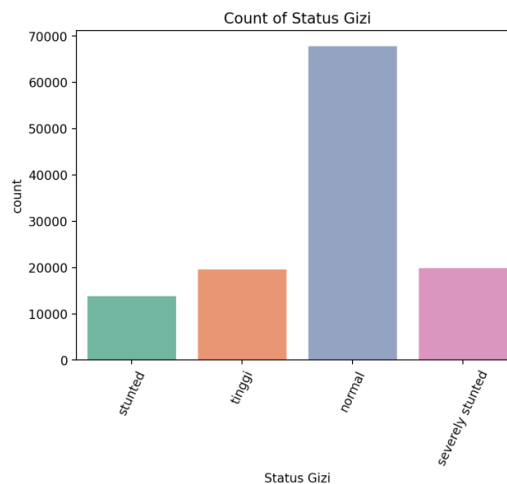
c. Distribusi berdasarkan tinggi badan



Gambar 5. Distribusi berdasarkan tinggi badan

Gambar 5 menunjukkan perbandingan rata-rata tinggi badan antara laki-laki dan perempuan, dengan laki-laki (ditandai dengan warna biru) memiliki rata-rata tinggi badan yang lebih tinggi dibandingkan perempuan (ditandai dengan warna merah muda). Hal ini mengindikasikan adanya perbedaan signifikan antara kedua kelompok yang mungkin disebabkan oleh faktor biologis atau genetik yang mempengaruhi tinggi badan berdasarkan jenis kelamin. Grafik ini sangat relevan untuk analisis dalam bidang demografi, kesehatan, dan ilmu olahraga, membantu dalam memahami bagaimana distribusi tinggi badan bervariasi berdasarkan jenis kelamin.

d. Distribusi berdasarkan frekuensi status gizi



Gambar 6. Distribusi Frekuensi Gizi

Pada gambar 5 Grafik yang disajikan menggambarkan distribusi frekuensi status gizi dibagi ke dalam empat kategori: '*severely stunting*', '*stunting*', 'normal', dan 'tinggi'. Kategori '*severely stunting*' menandakan kondisi yang sangat serius, sementara '*stunting*' merefleksikan kondisi pertumbuhan yang terhambat. 'Normal' menunjukkan gizi yang sehat, dan 'tinggi' menggambarkan pertumbuhan yang lebih dari standar. Klasifikasi ini berguna untuk mengidentifikasi dan melakukan intervensi cepat terhadap anak-anak yang memiliki risiko atau mengalami masalah dalam pertumbuhan.

3.3 Analisis Data Mining

Analisis ini digunakan untuk menilai seberapa efektif model yang dipilih dalam hal ini *random forest* dalam melakukan prediksi terhadap kejadian stunting, pola-pola tersembunyi akan dipelajari sehingga efektif untuk mengukur seberapa besar akurasi yang dihasilkan oleh model yang digunakan. Kelebihan *random forest* ini adalah mampu mengurai data dalam jumlah yang banyak dengan waktu yang efektif singkat [15].

Persamaan *random forest* adalah [16]:

$$Hx(Y) = - \sum p(i)p(j) \log_2 (p(i)) \quad (1)$$

Dimana :

$p(i)$ adalah peluang terjadinya kejadian X,

$p(j)$ apakah peristiwa Y terjadi,

$p(i,j)$ adalah peluang kejadian Y terjadi setelah kejadian X

Hasil analisa dengan menggunakan model *random forest* pada tabel 2.

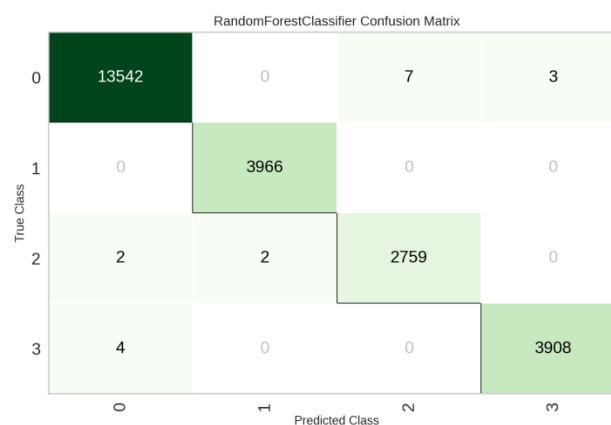
Fold	Akurasi	AUC	Recall	Precision	F1
0	0.9991	1.000	0.9991	0.9991	0.9991
1	0.9994	1.000	0.9994	0.9994	0.9994
2	0.9992	1.000	0.9992	0.9992	0.9992
3	0.9989	1.000	0.9989	0.9989	0.9989
4	0.9990	1.000	0.9990	0.9990	0.9990
5	0.9989	1.000	0.9989	0.9989	0.9989
6	0.9992	1.000	0.9992	0.9992	0.9992
7	0.9993	1.000	0.9993	0.9993	0.9993

Fold	Akurasi	AUC	Recall	Precision	F1
8	0.9989	1.000	0.9989	0.9989	0.9989
9	0.9984	1.000	0.9984	0.9984	0.9984
Mean	0.9990	1.000	0.9990	0.9990	0.9990
Std	0.0003	0.000	0.0003	0.0003	0.0003

Pada tabel 2 model melakukan prediksi yang hampir sempurna terhadap hasil rekam medis kasus stunting dibuktikan dengan 10 kali percobaan (*fold*) menghasilkan nilai rata-rata yang mendekati angka 1 yaitu, 0,9990 dengan demikian Model *random forest* menunjukkan performa yang sangat tinggi di semua metrik.

3.4 Evaluasi Hasil

Pada tahapan ini kesimpulan dapat diperoleh berdasarkan hasil penelitian diantaranya faktor-faktor yang menjadi ciri balita 0-60 terdeteksi sebagai stunting, normal, *severaly* stunting atau tinggi berdasarkan empat faktor atau variabel yang dijadikan patokan yaitu usia, jenis kelamin dan tinggi badan sehingga menghasilkan satu klasifikasi yaitu status gizi sebagai temuan yang digunakan untuk mengidentifikasi status terhadap balita. Hasil evaluasi dengan model *random forest* di tampilkan dalam *confussion matriks* pada gambar 7.



Gambar 7. *Confussion Matriks*

Pada gambar 7 Matriks ini memberikan gambaran yang jelas tentang performa model dalam mengklasifikasikan setiap kelas, dengan angka diagonal utama menunjukkan jumlah prediksi yang benar untuk setiap kelas dan elemen lainnya menunjukkan kesalahan klasifikasi. Sebagai contoh pada kelas 1 (normal) berdasarkan hasil analisa dan prediksi model *random forest* memprediksi dengan tepat terhadap 13.542 dan 0 data untuk tinggi, 7 data yang dikategorikan sebagai normal tetapi hasilnya stunting dan 3 data yang dikategorikan normal tetapi hasilnya *severaly* stunting.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa stunting adalah masalah serius yang disebabkan oleh malnutrisi berkepanjangan dan infeksi berulang, yang menghambat pertumbuhan fisik dan perkembangan kognitif anak-anak. Dengan menganalisis data rekam medis yang mencakup 121.000 dataset, penelitian ini berhasil mengklasifikasikan status gizi balita ke dalam empat kategori: normal, stunted, tinggi, dan severely stunted. Menggunakan model Random Forest, penelitian ini menguji keakuratan model dalam memprediksi status gizi berdasarkan variabel usia, jenis kelamin, tinggi badan, dan status gizi. Hasil analisis menunjukkan bahwa model *Random Forest* sangat efektif dalam melakukan klasifikasi terhadap stunting dengan rata-rata akurasi sebesar 0,9990 setelah dilakukan 10 kali *cross-validation*. Ini menandakan bahwa model ini memiliki kemampuan prediksi yang sangat baik, yang dapat diandalkan untuk membantu dalam mengidentifikasi balita yang berisiko mengalami stunting atau masalah gizi lainnya. Temuan ini memiliki

implikasi penting bagi praktik kesehatan publik, karena akurasi tinggi dari model prediksi dapat memungkinkan intervensi yang lebih tepat waktu dan efektif. Selain itu, penelitian ini menekankan pentingnya penggunaan data rekam medis yang komprehensif untuk pemantauan dan analisis kesehatan populasi anak-anak, yang dapat membantu dalam upaya pencegahan dan penanganan stunting secara lebih efektif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih yang tak terhingga kepada pihak-pihak yang telah membantu dalam penyusunan penelitian ini terutama kepada pihak kampus Politeknik Karya Husada.

REFERENSI

- [1] Nurlela Mariana Nababan, "Factors Influencing Stunting Incidence," *Int. J. Heal. Sci.*, vol. 3, no. 3, pp. 01–06, 2023.
- [2] Y. E. Purnamaningrum, K. D. Pertiwi, M. Margono, and D. Iswara, "Analysis Of Factors Related To Stunting Prevention In Children Aged 2-5 Years," *Interes. J. Ilmu Kesehat.*, vol. 10, no. 2, pp. 262–274, 2022.
- [3] I. Purnamasari, F. Widiyati, and M. Sahli, "Analisis Faktor yang Mempengaruhi Kejadian Stunting pada Balita," *J. Penelit. dan Pengabd. Kpd. Masy. UNSIQ*, vol. 9, no. 1, pp. 48–56, 2022.
- [4] S. Liem, D. R. F. Marta, and P. D. phil. H. Panggabean, "Sanitation Behavior and Risk of Stunting: Understanding the Discourse of a Public Service Announcement," *J. Messenger*, vol. 11, no. 2, p. 168, 2019.
- [5] D. Fifi and S. Hendi, "Pentingnya Kesehatan Masyarakat, Edukasi Dan Pemberdayaan Perempuan Untuk Mengurangi Stunting Di Negara Berkembang," *J. Pros. Semin. Nas. Kesehat.*, vol. 2, no. 01, pp. 16–25, 2020.
- [6] D. Permana, S. Anantanyu, and A. N. Priyatama, "Stunting Incidence in Toddlers Aged 24-59 Months in Kuburaya District Viewed from Feeding Patterns," *Proc. Int. Conf. Nurs. Heal. Sci.*, vol. 4, no. 1, pp. 259–266, 2023.
- [7] T. A. E. Putri, T. Widiariyah, and R. Santoso, "Penerapan Tuning Hyperparameter Randomsearchcv Pada Adaptive Boosting Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *J. Gaussian*, vol. 11, no. 3, pp. 397–406, 2023.
- [8] I. Islamiyah, "Analysis Of Determinant Factors In Stunting Incidence In Toddlers," *Int. J. Transdiscipl. Knowl.*, vol. 3, no. 2, pp. 35–42, 2022.
- [9] R. P. Pradana, "Stunting Toddler (Balita) Detection." .
- [10] R. L. Hasanah, M. Hasan, W. E. Pangesti, F. F. Wati, and W. Gata, "KLASIFIKASI PENERIMA DANA BANTUAN DESA MENGGUNAKAN METODE KNN (K-NEAREST NEIGHBOR)," *J. Techno Nusa Mandiri*, 2019.
- [11] R. Changala, A. Tharun, A. S. Akshith, and K. Karthik, "Analysis and Prediction of Water Quality Data using Machine Learning Approaches and Exploratory Data Analysis," vol. 3307.
- [12] Y. T. Kim, D. K. Kim, H. Kim, and D. J. Kim, "A Comparison of Oversampling Methods for Constructing a Prognostic Model in the Patient with Heart Failure," *Int. Conf. ICT Converg.*, vol. 2020-Octob, pp. 379–383, 2020.
- [13] R. Das, S. K. Biswas, D. Devi, and B. Sarma, "An Oversampling Technique by Integrating Reverse Nearest Neighbor in SMOTE: Reverse-SMOTE," *Proc. - Int. Conf. Smart Electron. Commun. ICOSEC 2020*, no. Icosec, pp. 1239–1244, 2020.
- [14] M. Ohsaki, P. Wang, K. Matsuda, S. Katagiri, H. Watanabe, and A. Ralescu, "Confusion-matrix-based kernel logistic regression for imbalanced data classification," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 9, pp. 1806–1819, 2017.
- [15] L. Abhishek, "Optical character recognition using ensemble of SVM, MLP and extra trees classifier," *2020 Int. Conf. Emerg. Technol. INCET 2020*, pp. 7–10, 2020.
- [16] B. O. Baines, A. Chung, and R. Raval, "Random Forest Classification Algorithm," no. April, pp. 1–10, 2020.