

## ***Decision Tree Regression* untuk Prediksi Prevalensi *Stunting* di Provinsi Nusa Tenggara Timur**

**Irnanda Septian Ika Putri<sup>1)</sup>, Risqy Siwi Pradini<sup>2\*)</sup>, Mochammad Anshori<sup>3)</sup>**

<sup>1)2)3)</sup>Program Studi Informatika, Fakultas Sains dan Teknologi, Institut Teknologi, Sains, dan Kesehatan RS. DR. Soepraoen Kesda V/BRW

<sup>\*)</sup>Correspondence author: [risqypradini@itsk-soepraoen.ac.id](mailto:risqypradini@itsk-soepraoen.ac.id), Malang, Indonesia

DOI: <https://doi.org/10.37012/jtik.v10i2.2179>

### **Abstrak**

*Stunting* adalah kondisi terhambatnya pertumbuhan linier anak-anak karena kekurangan gizi dan perawatan yang tidak memadai sejak dalam kandungan hingga usia dua tahun. *Stunting* disebabkan oleh berbagai faktor, termasuk kurangnya asupan gizi yang memadai, infeksi kronis atau berulang, praktik pemberian makanan yang tidak sesuai, sanitasi yang buruk, serta akses terbatas terhadap layanan kesehatan dan pendidikan gizi. Berdasarkan data *Global Nutrition Report* 2016, Indonesia berada di peringkat 108 dari 132 negara dalam hal prevalensi *stunting* (Kresnawati et al., 2022). Data dari Riset Kesehatan Dasar (Riskesdas) menunjukkan bahwa prevalensi *stunting* pada balita secara nasional menurun selama lima tahun terakhir, yaitu dari 37,2% menjadi 30,8% pada rentang tahun 2013 hingga 2018. Di Indonesia, provinsi yang memiliki prevalensi *stunting* paling tinggi berada di Nusa Tenggara Timur (NTT). Penelitian ini bertujuan untuk membuat model prediksi menggunakan *Decision Tree Regression* untuk memprediksi prevalensi *stunting* di NTT. Dengan demikian, hasil penelitian ini selain menghasilkan model prediksi juga dapat memberikan pemahaman yang lebih komperhensif mengenai faktor-faktor yang mempengaruhi tingkat *stunting* di NTT dan mendukung upaya untuk menurunkan angka prevalensinya di provinsi tersebut. Untuk menguji model prediksi yang dihasilkan, penelitian ini menggunakan metrik RMSE. Hasil pengujian dengan metrik RMSE menunjukkan nilai 0,093. Nilai ini membuktikan bahwa model *Decision Tree Regression* yang digunakan memiliki tingkat kesalahan prediksi yang relatif rendah, sehingga cukup efektif dalam memprediksi prevalensi *stunting* berdasarkan data yang digunakan.

**Kata Kunci:** *Prevalensi, Stunting, Prediksi, Decision Tree Regression, NTT*

### **Abstract**

*Stunting* is a condition where children's linear growth is hampered due to malnutrition and inadequate care from the womb until the age of two years. *Stunting* is caused by various factors, including lack of adequate nutritional intake, chronic or recurrent infections, inappropriate feeding practices, poor sanitation, and limited access to health services and nutrition education. Based on data from the 2016 *Global Nutrition Report*, Indonesia is ranked 108th out of 132 countries in terms of *stunting* prevalence (Kresnawati et al., 2022). Data from *Basic Health Research (Riskesdas)* shows that the prevalence of *stunting* among toddlers nationally has decreased over the last five years, namely from 37.2% to 30.8% in the period 2013 to 2018. In Indonesia, the province has the highest prevalence of *stunting* is in East Nusa Tenggara (NTT). This research aims to create a prediction model using *Decision Tree Regression* to predict the prevalence of *stunting* in NTT. Thus, the results of this research, apart from producing a prediction model, can also provide a more comprehensive understanding of the factors that influence the level of *stunting* in NTT and support efforts to reduce the prevalence rate in the province. To test the resulting prediction model, this research uses the RMSE metric. Test results with the RMSE metric show a value of 0.093. This value proves that the *Decision Tree Regression* model used has a relatively low level of prediction error, so it is quite effective in predicting the prevalence of *stunting* based on the data used.

**Keywords:** *Prevalence, Stunting, Predictions, Decision Tree Regression, NTT*

## PENDAHULUAN

*Stunting* adalah keadaan di mana pertumbuhan anak terhambat karena permasalahan kekurangan gizi (Sofiyya Wahyurin et al., 2019). *Stunting* disebabkan oleh berbagai faktor, termasuk kurangnya asupan gizi yang memadai, infeksi kronis atau berulang, praktik pemberian makanan yang tidak sesuai, sanitasi yang buruk, serta akses terbatas terhadap layanan kesehatan dan pendidikan gizi. Kondisi *stunting* dapat berasal dari permasalahan asupan gizi yang terjadi baik selama masa kehamilan maupun masa balita (Mahanani Mulyaningrum & Mulya Susanti, 2021). *Stunting* yang terjadi pada janin hingga anak usia dua tahun dapat meningkatkan angka kematian bayi dan anak, serta mengurangi efisiensi sistem kekebalan tubuh. Anak yang mengalami *stunting* cenderung rentan terhadap penyakit, menunjukkan postur tubuh yang tidak optimal, dan menunjukkan tingkat produktivitas yang rendah saat mencapai usia dewasa.

Berdasarkan data *Global Nutrition Report 2016*, Indonesia berada di peringkat 108 dari 132 negara dalam hal prevalensi *stunting* (Kresnawati et al., 2022). Data dari Riset Kesehatan Dasar (Riskesdas) menunjukkan bahwa prevalensi *stunting* pada balita secara nasional menurun selama lima tahun terakhir, yaitu dari 37,2% menjadi 30,8% pada rentang tahun 2013 hingga 2018. Di Indonesia, provinsi yang memiliki prevalensi *stunting* paling tinggi berada di Nusa Tenggara Timur (NTT). Berdasarkan data provinsi, NTT memiliki proporsi penduduk dengan status gizi sangat rendah dan *stunting* tertinggi sebesar 42,6%, sementara DKI Jakarta memiliki proporsi terendah sebesar 17,7%. Nilai tersebut masih dikategorikan tinggi karena standar yang ditetapkan oleh *World Health Organization* adalah minimal 20% (Haris et al., 2022). Mengingat potensi dampak serius dari masalah *stunting*, diperlukan prediksi untuk mengantisipasi *stunting* di masa depan dengan tujuan untuk mencegah terjadinya *stunting* dan mengurangi dampaknya pada perkembangan dan kesejahteraan anak.

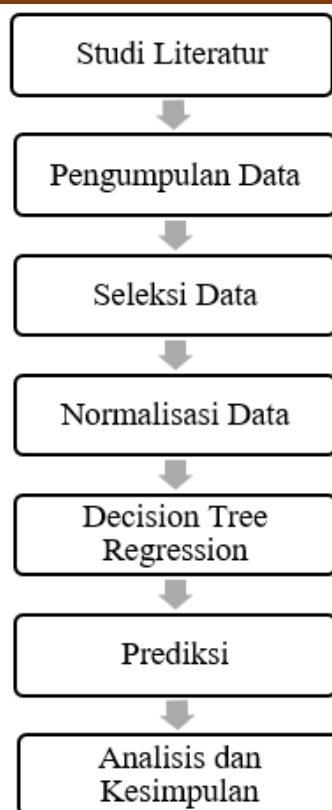
Terdapat beberapa penelitian terdahulu dengan topik *stunting*, seperti penelitian terkait implementasi *Decision Tree* untuk prediksi status gizi balita dengan tingkat akurasi mencapai 92,73% (Nike Aria Kurniawan, 2023). Selain itu, terdapat penelitian mengenai prediksi *stunting* pada anak balita menggunakan algoritma *Random Forest* dengan tingkat

akurasi rata-rata sekitar 97.87% . Terdapat pula penerapan algoritma SVM dan KNN menggunakan *Feature Selection Backward Elimination* untuk memprediksi status balita penderita *stunting* dengan hasil metode KNN menghasilkan nilai terkecil RMSE 2,476 (Labolo et al., 2022). Terdapat penelitian lainnya yang hampir serupa, yaitu penerapan *Decision Tree* untuk klasifikasi status gizi balita dan diperoleh nilai akurasi sebesar 99,18% (Bulkisah et al., 2024).

Berdasarkan penelitian-penelitian di atas, metode *Decision Tree* mendapatkan hasil akurasi paling tinggi dibandingkan algoritma lainnya dalam memprediksi status gizi balita. Selain itu menurut survei yang dilakukan pada tahun 2022, Provinsi NTT merupakan provinsi dengan tingkat prevalensi *stunting* tertinggi di Indonesia (Kementerian Kesehatan RI, 2022). Oleh karena itu, penelitian ini ingin membuat model prediksi menggunakan *Decision Tree Regression* untuk memprediksi prevalensi *stunting* di NTT. Penelitian ini menggunakan *Decision Tree Regression* karena ingin memprediksi nilai berupa angka prevalensi *stunting*. Model prediksi ini juga dapat mengidentifikasi faktor-faktor yang menyebabkan *stunting* di NTT. Dengan demikian, hasil penelitian ini selain menghasilkan model prediksi juga dapat memberikan pemahaman yang lebih dalam tentang faktor-faktor yang mempengaruhi *stunting* di NTT dan mendukung upaya untuk mengurangi prevalensinya di provinsi tersebut.

## METODE

Metode *Decision Tree Regression* populer karena kemampuannya untuk melakukan prediksi sekaligus mengidentifikasi hubungan antara atribut (Nike Aria Kurniawan, 2023). Dalam menangani fokus permasalahan, penelitian ini merancang kerangka kerja agar pelaksanaannya lebih terukur dan terarah. Adapun bentuk kerangka kerja ditunjukkan oleh Gambar 1.



**Gambar 1.** Kerangka Kerja Penelitian

### **A. Studi Literatur**

Tujuan dilakukan studi literatur untuk pengumpulan informasi dari berbagai sumber sebagai dukungan untuk pelaksanaan penelitian ini. Jurnal penelitian ilmiah yang dijadikan kerangka berpikir penelitian ini mencakup topik *stunting* pada balita, prediksi *stunting*, dan penerapan metode *Decision Tree Regression*.

### **B. Pengumpulan Data**

Data yang digunakan dalam penelitian ini adalah data prevalensi *stunting* di NTT yang diambil dari hasil survei kesehatan beberapa lembaga di tahun 2022. Peneliti menggunakan data dari tahun 2022 dikarena kurang lengkapnya data yang ada pada tahun 2023. Data bersumber dari Badan Pusat Statistik (BPS NTT, 2022) dan Kelompok Kerja (Pokja Penanganan *Stunting* NTT, 2022) pada tahun 2022. Berdasarkan data-data yang berhasil dikumpulkan, peneliti memperoleh 7 faktor yang menjadi penyebab *stunting* dan

memperoleh data untuk ketujuh faktor tersebut di masing-masing kabupaten/kota yang ada di Provinsi NTT.

Terdapat perbedaan rentang nilai pada data faktor-faktor tersebut, sehingga diperlukan normalisasi data. Penelitian ini menggunakan *Orange* untuk normalisasi data. *Orange Data Mining* adalah *software open source* untuk analisis atau penambangan data. *Orange* memiliki keunggulan dalam visualisasi atau yang sering disebut sebagai pemrograman visual. *Orange* menyajikan berbagai *widget* yang dapat ditempatkan di atas kanvas atau *drawingboard*, lalu dihubungkan dengan *widget* lainnya. Setiap *widget* memiliki fungsinya sendiri dan mampu menerima *input* serta menghasilkan *output* (Abdul Khalim et al., 2023)

### C. Seleksi Data

Langkah pertama dalam proses data *mining* adalah mengidentifikasi *dataset* yang akan digunakan untuk prediksi (Karlina & Nurdiawan 2023). Pada tahap seleksi data, dilakukan ekstraksi data karena tidak semua data akan digunakan. Oleh karena itu, hanya data yang relevan yang akan diambil dari *dataset* (Bahtiar, & Ali 2023). Dalam tahap seleksi data, peneliti memilih data faktor-faktor yang diperlukan dan menyesuaikannya dengan kebutuhan penelitian, sehingga terdapat 7 data yang akhirnya dipilih untuk proses ini, yaitu:

- Persalinan yang dibantu oleh nakes
- Berat badan lahir rendah
- Persentase anak di bawah usia 2 tahun yang diberi ASI
- Balita yang tidak mendapatkan imunisasi lengkap
- Balita yang mendapatkan imunisasi lengkap
- Indeks pembangunan manusia
- Masyarakat miskin

### D. Normalisasi Data

Dalam penelitian ini, *preprocessing* data dilakukan untuk mengubah data dengan menerapkan teknik normalisasi menggunakan bantuan *Orange*. Normalisasi data menggunakan *widget preprocessing* data dengan menormalkan fitur pada *Orange* kemudian

dipilih *Divide by number of values* dan *Normalize to interval [0,1]*. Harapan dari dilakukannya *preprocessing* data adalah tercapainya kondisi ideal dalam pengolahan data (Karo & Hendriyana, 2022).

### E. Decision Tree Regression

Metode *Decision Tree Regression* dimulai dengan menghitung nilai *entropy*, yang mengukur tingkat ketidakpastian atau impuritas dalam *dataset*. Persamaan untuk mencari nilai *entropy* ditunjukkan sebagai berikut ini:

$$entro(S) = \sum_{i=1}^m - p(w_i|S) \cdot \log_2(w_i|S)$$

Keterangan:

S: Himpunan kasus data yang sedang dianalisis

M: Jumlah total kelas yang berbeda dalam himpunan data S

$w_i|S$ : Probabilitas kemunculan kelas  $w_i$  dalam himpunan data S

Setelah mendapatkan nilai *entropy*, langkah selanjutnya adalah menghitung nilai *gain* yang merupakan ukuran seberapa besar informasi yang diperoleh dari pemisahan data berdasarkan atribut tertentu. Nilai *gain* yang lebih tinggi menunjukkan bahwa pemisahan berdasarkan atribut tersebut memberikan informasi yang lebih signifikan dan lebih baik dalam mengurangi ketidakpastian. Persamaan untuk menghitung *gain* ditunjukkan sebagai berikut:

$$Gain(S, J) = Entropy(S) - \sum_{i=j}^n p(v_i|S) * Entropy(S_i)$$

Keterangan:

S: Himpunan data yang sedang dianalisis

J: Fitur/atribut yang dipertimbangkan untuk pemisahan data

n: Banyak kelas di *node*

$v_i|S$ : Proporsi nilai  $v$  yang muncul pada kelas dalam *node*

$S_i$ : Entropi dari komposisi nilai  $v$  untuk kelas ke- $j$  dalam *node* data ke- $i$  *node*

Setelah melakukan perhitungan *entropy* dan *gain*, langkah selanjutnya adalah membentuk cabang untuk setiap nilai yang terdapat dalam akar tersebut. Proses ini kemudian dilanjutkan dengan membagi atribut di dalam setiap cabang dan pada akhirnya akan membentuk pohon keputusan.

## F. Prediksi

Setelah seluruh pohon keputusan telah dibangun, prediksi dilakukan dengan mengintegrasikan hasil prediksi dari masing-masing pohon. Dalam konteks klasifikasi, hasil prediksi dapat diperoleh dengan voting terbanyak. Sedangkan dalam regresi, hasil prediksi dapat diperoleh dengan keluarnya *mean* atau median dari prediksi yang dihasilkan pohon (Rofiani et al., 2024). Dengan demikian, penelitian ini akan menghasilkan nilai prediksi untuk prevalensi *stunting* yang terjadi di NTT.

## G. Pengujian Model

*Root Mean Square Error* (RMSE) digunakan untuk mengukur akurasi model prediksi dengan cara membandingkan data aktual dengan data hasil pemrosesan untuk menentukan tingkat kesalahan yang terjadi. RMSE mengukur seberapa besar perbedaan antara nilai yang diprediksi dan nilai yang sebenarnya dengan satuan yang sama dengan data asli. Semakin kecil nilai RMSE (mendekati 0) maka semakin akurat hasil prediksi tersebut (Kurniawan et al., 2023). Persamaan untuk menghitung RMSE ditunjukkan sebagai berikut:

$$RMSE = \left( \frac{\sum (y_i - \hat{y}_i)^2}{n} \right)^{1/2}$$

Keterangan:

RMSE: nilai *Root Mean Square Error*

$Y$ : nilai hasil observasi

$\hat{y}$ : nilai hasil prediksi

$I$ : index/urutan pada daftar data

$N$ : jumlah data

## HASIL DAN PEMBAHASAN

Pada pengumpulan data, terdapat 12 faktor-faktor yang menyebabkan prevalensi *stunting* di NTT. Setelah dilakukan seleksi data, hanya 7 faktor yang memenuhi persyaratan. Ketujuh faktor tersebut dipilih karena tidak memiliki nilai yang hilang (*missing value*). Kemudian ketujuh faktor tersebut diisi dengan data per wilayah kota/kabupaten yang ada di NTT, sehingga terdapat 22 baris data yang merepresentasikan data per kota/kabupaten. Data yang telah dikumpulkan kemudian diproses melalui normalisasi untuk menstandarkan rentang nilai. Pada normalisasi data, peneliti menggunakan *Orange* sehingga didapatkan hasil seperti yang ditunjukkan Tabel 1 berikut ini.

**Tabel 1.** Normalisasi Menggunakan *Orange*

No	V1	V2	V3	V4	V5	V6	V7	V8
1.	0.763257	0.334735	0.6720	0.269730	0.730270	0.292825	0.1859625	0.39155
2.	0.691413	0.267452	0.848615	0.773620	0.370852	0.226380	0.5590590	0.34085
3.	0.269586	0.700168	0.629538	0.614672	0.385328	0.320179	0.6834244	0.43099
4.	0.661649	0.093356	0.951385	0.748796	0.251204	0.216592	1.00000	0.59155
5.	0.687308	0.304878	0.239385	0.553538	0.446462	0.285202	0.3696798	0.64225
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
22.	0.862128	0.180404	0.530462	0.802519	0.197481	1.0000	0.2166146	0.45915

Kolom nomor mewakili data wilayah kabupaten/kota di NTT. Kolom V1 mewakili data faktor persalinan yang dibantu oleh tenaga kesehatan. Kolom V2 mewakili data faktor berat badan lahir rendah. Kolom V3 mewakili data faktor presentase anak dibawah 2 tahun yang diberi ASI. Kolom V4 mewakili data faktor balita yang mendapatkan imunisasi tidak lengkap. Kolom V5 mewakili data faktor balita yang mendapatkan imunisasi lengkap. Kolom V6 mewakili data faktor indeks pembangunan manusia. Kolom V7 mewakili data faktor masyarakat miskin. Kolom V8 mewakili data prevalensi *stunting* per kabupaten. Kolom V1 sampai dengan V7 sebagai fitur, dan kolom V8 sebagai prediksi prevalensi *stunting*.

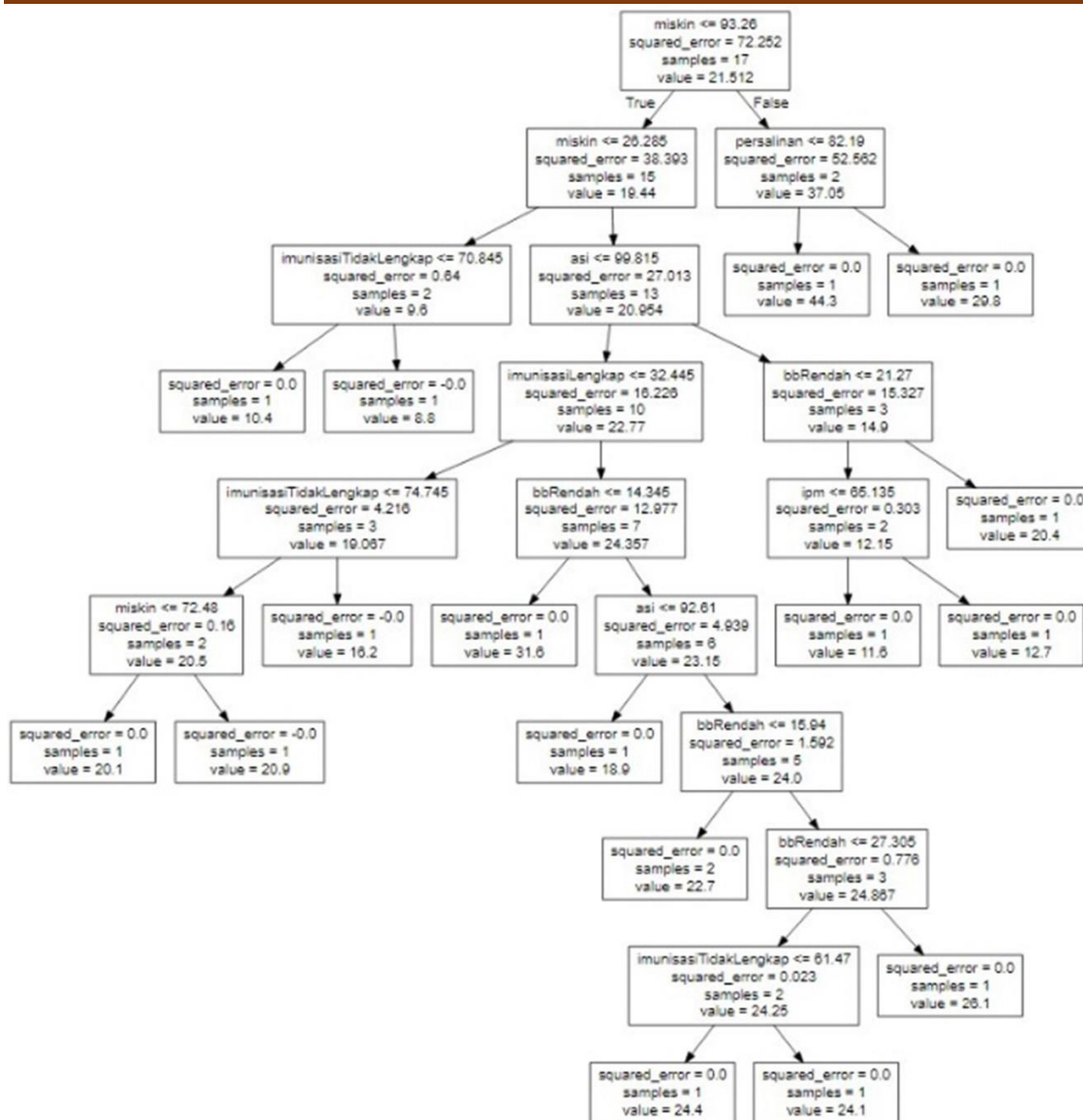
Data normalisasi di atas kemudian diolah menggunakan bahasa pemrograman Python. Pustaka yang digunakan meliputi matplotlib untuk visualisasi data, numpy untuk komputasi numerik, dan scikit-learn untuk mengimpor kelas *Decision Tree Regression*. Fungsi



`read_csv()` digunakan untuk membaca *file* CSV dan mengonversinya menjadi data *frame* yang merupakan struktur data tabel yang sering dipakai dalam analisis data (Siska & Putra, 2023). Selanjutnya, *dataset* dibagi menjadi data *training* dan data *testing* dengan ratio 80:20. Alasan pembagian ratio ini dikarenakan metode pembagian yang umum digunakan (Angga Sabda, 2023). Data *training* digunakan untuk melatih model, sementara data *testing* digunakan untuk menguji kinerja model yang telah dilatih. *Dataset* kemudian diproses menggunakan model *Decision Tree Regression* untuk memprediksi prevalensi *stunting*.

Setelah melakukan beberapa percobaan, hasil evaluasi menunjukkan bahwa model dengan metrik RMSE adalah 0,093. Hal ini membuktikan bahwa model *Decision Tree Regression* yang digunakan memiliki tingkat kesalahan prediksi yang relatif rendah, sehingga cukup efektif dalam memprediksi prevalensi *stunting* berdasarkan data yang digunakan. Setelah membangun dan mengeksekusi model, langkah selanjutnya adalah menganalisis struktur pohon dari model yang telah dibuat.

Untuk melakukan ini, peneliti menggunakan *WebGraphviz* (*WebGraphviz*). *WebGraphviz* dapat digunakan untuk visualisasi struktur pohon secara interaktif, sehingga memudahkan pemahaman terhadap bagaimana model *Decision Tree Regression* mengambil keputusan dan membuat prediksi. Visualisasi ini dapat digunakan untuk mengidentifikasi faktor-faktor dominan yang menyebabkan prevalensi *stunting*. Hal ini juga dapat digunakan untuk memahami alur logika yang digunakan oleh model *Decision Tree Regression* dalam proses prediksi. Hasil visualisasi pohon keputusan ditunjukkan oleh Gambar 2. Setiap cabang dalam pohon keputusan mewakili sebuah aturan berdasarkan fitur data, dan setiap daun (*leaf*) mewakili hasil atau keputusan akhir (Moerdyanto & Nuryana, 2023).



**Gambar 2.** Pohon keputusan prediksi prevalensi *stunting*

Berdasarkan pohon keputusan untuk prediksi prevalensi *stunting* seperti yang ditunjukkan dalam Gambar 2, ditemukan beberapa aturan atau *rules* yang terbentuk yaitu:

- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $\leq 26.28$  AND balita tidak imunisasi lengkap  $\leq 70.84$  THEN prevalensinya 9.6.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $\leq 26.28$  AND balita tidak imunisasi lengkap  $> 70.84$  THEN prevalensinya 8.8.

- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.28$  AND balita dibawah 2 tahun diberi asi  $\leq 90.81$  AND balita mendapat imunisasi lengkap  $\leq 32.44$  AND balita tidak imunisasi lengkap  $\leq 74.74$  AND masyarakat miskin  $\leq 72.48$  THEN prevalensinya 20.5.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.28$  AND balita dibawah 2 tahun diberi asi  $\leq 90.81$  AND balita mendapat imunisasi lengkap  $\leq 32.44$  AND balita tidak imunisasi lengkap  $\leq 74.74$  AND masyarakat miskin  $> 72.48$  THEN prevalensinya 20.9.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.28$  AND balita dibawah 2 tahun diberi asi  $\leq 90.81$  AND balita mendapat imunisasi lengkap  $\leq 32.44$  AND balita tidak imunisasi lengkap  $> 74.74$  THEN prevalensinya 31.6.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $\leq 90.815$  AND balita mendapat imunisasi lengkap  $> 32.445$  AND berat badan lahir rendah  $\leq 14.346$  and balita dibawah 2 tahun diberi asi  $\leq 92.61$  THEN prevalensinya 23.15.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $\leq 90.81$  AND balita mendapat imunisasi lengkap  $> 32.445$  AND berat badan lahir rendah  $\leq$  AND balita dibawah 2 tahun diberi asi  $> 92.61$  THEN prevalensinya 18.9.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $\leq 90.815$  AND balita mendapat imunisasi lengkap  $> 32.445$  AND berat badan lahir rendah  $> 14.346$  AND IPM  $\leq 68.135$  THEN prevalensinya 12.15.
- IF masyarakat miskin  $\leq 93.26$  and masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $\leq 90.815$  AND balita mendapat imunisasi lengkap  $> 32.445$  AND berat badan lahir rendah  $> 14.346$  AND IPM  $> 68.135$  THEN prevalensinya 20.4.
- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $> 90.815$  AND berat badan lahir rendah  $\leq 21.073$  THEN prevalensinya 20.4

- IF masyarakat miskin  $\leq 93.26$  AND masyarakat miskin  $> 26.285$  AND balita dibawah 2 tahun diberi asi  $> 90.815$  AND berat badan lahir rendah  $> 21.073$  THEN prevalensinya 12.7.
- IF masyarakat miskin  $> 93.26$  AND persalinan dibantu nakes  $\leq 82.19$  THEN prevalensinya 37.05.
- IF masyarakat miskin  $> 93.26$  AND persalinan dibantu nakes  $> 82.19$  THEN prevalensinya 20.9.

## KESIMPULAN DAN REKOMENDASI

Berdasarkan penjabaran pada bagian hasil dan pembahasan, dapat disimpulkan bahwa model *Decision Tree Regression* untuk prediksi prevalensi *stunting* di NTT berhasil dibangun. Melalui serangkaian uji coba, hasil evaluasi menunjukkan bahwa model *Decision Tree Regression* yang diuji dengan metrik RMSE mendapatkan hasil 0,093. Nilai ini membuktikan bahwa model *Decision Tree Regression* yang digunakan memiliki tingkat kesalahan prediksi yang relatif rendah, sehingga cukup efektif dalam memprediksi prevalensi *stunting* berdasarkan data yang tersedia.

Model prediksi yang terbentuk juga menghasilkan visualisasi pohon keputusan yang dapat digunakan untuk membuat *rules* dalam menentukan apakah prevalensi *stunting* tinggi atau rendah. Berdasarkan *rules* tersebut, dapat dianalisis faktor-faktor dominan penyebab prevalensi *stunting* diantaranya adalah masyarakat miskin, persalinan yang dibantu oleh tenaga kesehatan, berat badan lahir rendah, indeks pembangunan manusia, presentase anak dibawah 2 tahun yang diberi ASI, balita yang mendapatkan imunisasi lengkap, dan balita yang mendapatkan imunisasi tidak lengkap. Diharapkan pengetahuan akan faktor-faktor ini dapat mendukung upaya untuk mengurangi prevalensi *stunting* di NTT.

## REFERENSI

- Abdul Khalim, K., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naive Bayes. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 1). DOI: <https://doi.org/10.36040/jati.v7i1.6376>
- Angga Sabda, M. (2023). Implementasi Data Mining Dalam Memprediksi Penjualan Parfum Terlaris Menggunakan Metode K-Nearest Neighbor. In *Jurnal Sistem Komputer Dan Informatika (JSON)* Hal: 415–, 422(2), 415–422. <https://doi.org/10.30865/json.v5i2.7194>
- BPS NTT. (2022). <https://ntt.bps.go.id/subject/30/kesehatan.html#subjekViewTab3>
- Bulkisah, S., Astuti, R., & Bahtiar, A. (2024). *Implementasi Data Mining Algoritma Decision Tree Untuk Klasifikasi Status Gizi Balita*. 29, 1–12. DOI: <http://dx.doi.org/10.35760/ik.2024.v29i1.10346>
- Haris, M. S., Khudori, A. N., & Kusuma, W. T. (2022). Perbandingan Metode Supervised Machine Learning untuk Prediksi Prevalensi Stunting di Provisi Jawa Timur. In *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 9(7), 1571. <https://doi.org/10.25126/jtiik.2022976744>
- Karo, I. M., & Hendriyana (2022). *Klasifikasi Penderita Diabetes Menggunakan Algoritma Machine Learning Dan Z-Score*. In *Jurnal Teknologi Terpadu* Vol. 8, No. 2, pp. 94-99. DOI: <https://doi.org/10.54914/jtt.v8i2.564>
- Karlina, L., & Nurdiawan, O. (2023). Penerapan K- Medoids Dalam Klasifikasi Persebaran Lahan Kritis Di Jawa Barat Berdasarkan Kabupaten/Kota. In *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 527–532. <https://doi.org/10.36040/jati.v7i1.6348>
- Kementerian Kesehatan RI. (2022). *Buku Saku: Hasil Survei Status Gizi Indonesia (SSGI) 2022*. Kementerian Kesehatan Republik Indonesia, 1–99. <https://promkes.kemkes.go.id/materi-hasil-survei-status-gizi-indonesia-ssgi-2022>
- Kresnawati, W., Ambarika, R., & Saifulah, D. (2022). Pengetahuan dan Sikap Ibu Balita Sadar Gizi terhadap kejadian Stunting. In *Journal Of Health Science Community*, 3(1), 26–33. DOI: DOI: <https://doi.org/10.30994/jhsc.v3i1.167>

- Kurniawan, R., Halim, A., & Melisa, H. (2023). Prediksi Hasil Panen Pertanian Salak di Daerah Tapanuli Selatan Menggunakan Algoritma SVM (Support Vector Machine). In *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(2), 903–912. <https://doi.org/10.30865/klik.v4i2.1246>
- Labolo, A. Y., Mooduto, S., Bode, A., Colanus, I., & Drajana, R. (2022). Penerapan Algoritma Support Vector Machine dan K-Nearest Neighbor Menggunakan Feature Selection Backward Elimination Untuk Prediksi Status Penderita Stunting Pada Balita. In *Jurnal Tecnoscienza*, 6(2), April 2022. <https://doi.org/10.51158/tecnoscienza.v6i2.713>
- Mahanani Mulyaningrum, F., & Mulya Susanti, M. (2021). Faktor-Faktor Yang Mempengaruhi Stunting Pada Balita Di Kabupaten Grobogan. In *Jurnal Keperawatan dan Kesehatan Masyarakat STIKER Cendekia Utama Kudus*. Vol. 10, No.1, Maret 2021. DOI: <https://doi.org/10.31596/jcu.v10i1.704>
- Moerdyanto, O. P., & Nuryana, I. K. D. (2023). Prediksi Kelulusan Tepat Waktu Menggunakan Pendekatan Pohon Keputusan Algoritma Decision Tree. In *Journal of Informatics and Computer Science*, 05(1), 90–96. <https://ejournal.unesa.ac.id/index.php/jinacs/article/view/55329>
- M., Bahtiar, A., & Ali, I. (2023). Transformasi Strategi Penjualan Batik Cirebon Dengan Pendekatan Analisis Pengelompokan K-Means. In *KOPERTIP Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 7(1), 1–7. <https://doi.org/10.32485/kopertip.v7i1.309>
- Nike Aria Kurniawan, D. (2023). Implementasi Metode Decision Tree pada Sistem Prediksi Status Gizi Balita. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 7, Issue 2). DOI: <http://dx.doi.org/10.30645/j-sakti.v7i2.681>
- Pokja Penanganan Stunting NTT. (2022). <https://www.victorynews.id/kupang/pr-3313018484/simak-sebaran-stunting-2022-per-kabupatenkota-di-ntt>
- Rofiani, R., Oktaviani, L., Vernanda, D., & Hendriawan, T. (2024). Penerapan Metode Klasifikasi Decision Tree dalam Prediksi Kanker Paru-Paru Menggunakan Algoritma C4. 5. In *Jurnal Tekno Kompak*, 18(1), 126–139. DOI: <https://doi.org/10.33365/jtk.v18i1.3525>

Siska, A. F., & Putra, R. E. (2023). Klasifikasi Tingkat Kepuasan Wali Murid Terhadap Hasil Belajar Anak Menggunakan Algoritma C5.0. In *Journal of Informatics and Computer Science*, 4(4), 432–435.

Sofiyya Wahyurin, I., Nur Aqmarina, A., Rahmah, H. A., Hasanah, A. U., Nataly, C., & Silaen, B. (2019). In *Ilmu Gizi Indonesia*, 2(2), 141-146, Februari 2019.

*WebGraphviz is Graphviz*. <http://www.webgraphviz.com/>