

Penerapan Algoritma *Support Vector Machine* untuk Klasifikasi Stunting pada Balita di Kabupaten Enrekang

¹Andi Widya Mufila Gaffar, ²Andi Muhammad Halis, ³Purnawansyah, ⁴Sitti Rahmah Jabir
^{1, 2, 3, 4}Universitas Muslim Indonesia

¹widya.mufila@umi.ac.id, ²andimuhammadhalis24@gmail.com, ³purnawansyah@umi.ac.id,
⁴rahmahjabir@umi.ac.id

ABSTRAK

Stunting adalah salah satu bentuk gizi yang kurang yang ditandai dengan tinggi badan berdasarkan umur dan diukur dengan standar deviasi referensi WHO yang dapat dilihat pada Indikator TB/U (tinggi badan dengan usia). Salah satu kabupaten dengan jumlah penduduk stunting terbanyak di provinsi Sulawesi Selatan adalah Kabupaten Enrekang. Metode yang digunakan pada penelitian ini yaitu klasifikasi stunting menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel *polynomial* yang bertujuan untuk mengelompokkan data anak balita di bawah 60 bulan apakah mengalami stunting atau tidak (normal). Klasifikasi Stunting pada balita memiliki signifikansi penting karena menjadi landasan untuk merancang program pencegahan Stunting. Untuk menilai performa dan cara kerja model *Support Vector Machine* pada data anak balita di Kabupaten Enrekang, digunakan metode pengujian *cross validation*. Selain itu, hasil prediksi model dibandingkan dengan fakta aktual menggunakan *confusion matrix*. Pada pengujian dengan 10 K-Fold *Cross Validation* menggunakan *Support Vector Machine* menunjukkan hasil dengan nilai tertinggi berada pada *fold* ke-4 dengan tingkat *accuracy* 99.13% *precision* 99.13% *recall* 99.13% *f1-score* 99.13%. sedangkan nilai terendah berada pada *fold* ke-0 dengan tingkat *accuracy* 95.63% *precision* 95.74% *recall* 95.63% *f1-score* 95.51%. Untuk rata-rata dari pengujian *fold* menunjukkan hasil *accuracy* 96.98% *precision* 96.99% *recall* 96.98% *f1-score* 96.94%. Sedangkan untuk hasil dari *Confusion Matrix* dengan nilai *accuracy* sebanyak 98% secara total.

Kata Kunci: Klasifikasi, stunting, SVM, enrekang, kernel polynomial

PENDAHULUAN

Indonesia mengalami masalah gizi yang cukup berat, ditandai dengan banyaknya kasus gizi buruk pada anak balita. Stunting adalah salah satu bentuk gizi yang kurang yang ditandai dengan tinggi badan berdasarkan umur dan diukur dengan standar deviasi referensi WHO. Indikator TB/U (tinggi badan dengan usia) memberikan indikasi masalah gizi yang bersifat kronis sebagai akibat dari keadaan yang berlangsung lama, seperti kemiskinan, perilaku hidup yang tidak sehat dan pola asuh/pemberian makan yang kurang baik dari saat anak dilahirkan, yang menyebabkan anak menjadi pendek (Drajana & Bode, 2022).

Banyak anak yang mengalami Stunting yang dapat memberikan dampak buruk pada masyarakat, masalah ini berlangsung cukup lama. Stunting dipengaruhi oleh banyak faktor diantaranya pemberian ASI (Air Susu Ibu) eksklusif. Masalah kesehatan masyarakat dianggap berat bila prevalensi Stunting sebesar 30-39 persen dan sangat berat bila prevalensi Stunting ≥ 40 persen (Byna & Anisa, 2018).

Pada tahun 2019, Sulawesi Selatan termasuk dalam daftar 4 provinsi dengan jumlah penduduk Stunting terbanyak di Indonesia, dan Kabupaten Enrekang menjadi daerah dengan jumlah penduduk Stunting terbanyak di provinsi Sulawesi Selatan. Berdasarkan data BPS-

Kemenkes dan SSGBI (Studi Status Gizi Indonesia) 2019, angka Stunting di Kabupaten Enrekang merupakan kasus Stunting tertinggi sebesar 43,7%. Berdasarkan data SSGI (Studi Status Gizi Indonesia) tahun 2021, angka anak Stunting di Kabupaten Enrekang sebanyak 31,9 miliar, menempati urutan ke-8 di Provinsi Sulawesi Selatan. Berdasarkan data Dinas Kesehatan (Dinkes) Enrekang, data anak Stunting tahun 2021 di kabupaten Enrekang mencapai 22,8% atau 3.277 dari total 15.275 balita di kabupaten Enrekang. Data tersebut didasarkan pada hasil pemantauan status gizi tahun 2021 yang dilakukan oleh seluruh Puskesmas di Kabupaten Enrekang (Riyadh, Batara, & Nurlinda, 2023).

Klasifikasi Stunting pada balita memiliki signifikansi penting karena menjadi landasan untuk merancang program pencegahan Stunting. Salah satu pendekatan yang dapat diterapkan dalam klasifikasi Stunting adalah menggunakan algoritma *Support Vector Machine*. SVM merupakan contoh algoritma yang berguna dalam pembelajaran mesin dan memiliki penerapan luas di berbagai sektor, termasuk dalam konteks kesehatan. Berdasarkan penjelasan di atas, maka peneliti melakukan penelitian yang akan melakukan Prediksi Penyakit Stunting pada Balita menggunakan *Support Vector Machine* di Kabupaten Enrekang untuk mendapatkan *accuracy* yang sesuai dan tepat. Adapun tujuan penelitian ini untuk untuk mengklasifikasikan stunting pada balita di Kabupaten Enrekang menggunakan algoritma *Support Vector Machine* dan mengetahui nilai *accuracy*, *precision*, *recall* dan *f1-score* dari algoritma *Support Vector Machine* dalam mengklasifikasi tingkat stunting pada balita di Kabupaten Enrekang.

TINJAUAN PUSTAKA

Balita

Balita merupakan kelompok penduduk yang memiliki risiko tinggi terhadap masalah gizi. Balita yang diharapkan akan menjadi sumber daya manusia yang berkualitas di masa depan, membutuhkan perhatian khusus karena usia di bawah lima tahun dianggap sebagai "usia emas" dalam pembentukan sumber daya manusia. Dalam hal ini, status gizi yang baik sangat penting untuk menentukan kesuksesan upaya peningkatan sumber daya manusia, karena mempengaruhi pertumbuhan fisik dan kecerdasan anak. Namun, anak balita adalah salah satu kelompok masyarakat yang rentan terhadap masalah gizi. Anak balita membutuhkan asupan makanan dan gizi yang cukup dan memadai karena mengalami pertumbuhan dan perkembangan yang pesat. Kurang gizi pada masa balita dapat menyebabkan gangguan pertumbuhan dan gangguan perkembangan mental (Usman, Rosdiana, & Misnawati, 2021).

Gizi

Gizi memiliki peran penting dalam proses tumbuh kembang balita dan memiliki hubungan yang erat dengan kesehatan dan kecerdasan. Pemberian gizi yang tidak memadai, khususnya pada anak-anak, dapat mengurangi potensi sumber daya untuk pembangunan masyarakat (Alhamid, 2021). Status gizi adalah keadaan tubuh sebagai akibat konsumsi makanan dan penggunaan zat-zat gizi. Dibedakan antara status gizi buruk, kurang, baik, dan lebih. Status gizi mencerminkan kecukupan zat gizi pada tubuh. Apabila terjadi kekurangan asupan makanan, anak dapat mengalami gizi kurang atau bahkan gizi buruk, sedangkan di sisi lain, sebagian anak yang makan secara berlebihan dan mengalami kelebihan gizi dapat mengalami obesitas (Simanjuntak & Siregar, 2020).

Stunting

Stunting merupakan kondisi di mana pertumbuhan fisik anak terganggu dengan ciri menurunnya tinggi dan berat badan secara signifikan. Ini didefinisikan sebagai kegagalan dalam pertumbuhan pada anak balita (di bawah usia 60 bulan) yang disebabkan oleh malnutrisi kronis yang menyebabkan postur tubuh pendek jika dibandingkan dengan teman sebayanya. Malnutrisi ini dapat terjadi sejak dalam kandungan dan pada masa awal setelah bayi lahir (HIDAYANI, 2020). Menurut standar World Health Organization (WHO) Child Growth Standart, Stunting diukur dengan perbandingan antara panjang badan atau tinggi badan dengan usia (PB/U atau TB/U)

dengan batas kurang dari -2 SD (Standar Deviasi). Menurut peraturan Kementerian Kesehatan Republik Indonesia No 2 tahun 2020, kategori status gizi anak berdasarkan panjang badan atau tinggi badan menurut umur (PB/U atau TB/U) anak usia 0 – 60 bulan terdiri dari sangat pendek (Severely Stunted) dengan ambang batas (Z Score <-3 SD), pendek (Stunted) dengan ambang batas (Z Score -3 SD sd <-2 SD), normal dengan ambang batas (Z Score -2 SD sd +3 SD), dan tinggi dengan ambang batas (Z Score >+3 SD). Anak dengan PB/U atau TB/U dibawah minus dua standar deviasi (<-2 SD) adalah anak dengan perawakan pendek. Pada anak dengan PB/U atau TB/U terletak di atas tiga standar deviasi (>+3 SD) artinya anak berperawakan tinggi dan perlu dirujuk ke faskes yang lebih tinggi (KESEHATAN, 2020).

Support Vector Machine

Support Vector Machine (SVM) merupakan salah satu pengklasifikasi diskriminatif dengan hyperplane pemisah. Metode ini melakukan pelatihan data berlabel (supervised learning), sementara itu, algoritma menghasilkan hyperplane optimal yang mengkategorikan contoh baru. Hyperplane adalah garis yang memisahkan sebuah plane menjadi dua bagian pada setiap kelas yang terletak di kedua sisi. Menurut Santoso, Support vector machine (SVM) merupakan salah satu teknik terbaik untuk menyelesaikan proses prediksi, dalam kasus klasifikasi atau regresi (Seno & Wibowo, 2019).

METODE PENELITIAN

Dataset

Proses pertama dengan melakukan input data stunting sebanyak 3437 data. *Dataset* menggunakan format excel agar memudahkan data terbaca dalam *Google Collab* yang akan diolah menggunakan pemrograman *Python*.

Preprocessing

Preprocessing data dimulai dengan mengidentifikasi missing value pada dataset, menggunakan ringkasan (dataset) fungsi (Gaffar & Sitanggang, 2019). Pada tahap ini, preprocessing yang dilakukan terdapat 3 tahap, yaitu pembersihan data (data cleaning) yaitu proses menghilangkan noise (pengganggu) pada data yang dianggap kurang penting. Transformasi data atau label Encoder yaitu proses perubahan data ke dalam format tertentu agar memudahkan data saat di proses bertujuan untuk merubah semua data bertipe kategorik menjadi numerik. Normalisasi data yaitu mengubah skala data sehingga semua fitur atau atribut memiliki rentang yang serupa. Metode normalisasi yang digunakan yaitu MinMax Scalling yang bertujuan untuk membuat skala data seragam, sehingga fitur-fitur dengan skala besar tidak memiliki pengaruh yang lebih besar daripada fitur-fitur dengan skala kecil dalam algoritma pembelajaran mesin.

Split Data

Setelah dataset melakukan tahap preprocessing, dataset akan dipisahkan menjadi dua kategori yaitu data latih dan data uji, dalam penelitian ini split data dengan perbandingan 90:10, 80:20, 70:30, 60:40.

Training Support Vector Machine

Pada tahap ini melatih model Support Vector Machine menggunakan data latih dengan tujuan untuk memastikan bahwa model tersebut dapat bekerja dengan baik pada data stunting.

Model Support Vector Machine

Pada tahap ini, dilakukan pengujian Support Vector Machine menggunakan kernel linear, kernel polynomial, kernel sigmoid, dan kernel rbf. Tujuannya mencari nilai accuracy paling tinggi diantara kernel tersebut.

Evaluasi

Dasar penelitian ini adalah proses bentuk klasifikasi, sehingga hasil dari proses klasifikasi yang telah dilakukan harus dievaluasi. Tahap evaluasi merupakan tahap terakhir serangkaian tahap dari perancangan sistem, dalam penelitian ini penulis menggunakan metode cross validation dan confusion matrix. K-fold cross validation dengan nilai K sebanyak 10. Tujuan dari pengujian ini adalah untuk mengukur kinerja model SVM dengan kernel polynomial secara lebih akurat dan konsisten dengan melihat nilai accuracy, precision, recall dan f1-score. Confusion matrix digunakan untuk mengetahui nilai accuracy dalam mengklasifikasikan data dengan benar.

HASIL DAN PEMBAHASAN

Preprocessing

Cleaning data atau pembersihan data dilakukan dengan menghilangkan data yang bernilai *Nan* atau yang tidak memiliki nilai pasti, jumlah data yang di hapus sebanyak 168 sehingga total data keseluruhan sebanyak 3269 data.

Tabel 1. Transformasi Atribut TB/U

Sebelum Transformasi	Sesudah Transformasi
Normal	0
Pendek	1
Sangat Pendek	2

Pada tabel 1 tersebut menunjukkan transformasi dari atribut TB/U yang sebelumnya tipe datanya kategorikan menjadi tipe data numerikal dengan menggunakan label *encoder*. Sehingga atribut TB/U menjadi dalam bentuk angka 0 sampai 2.

Tabel 2. Hasil Normalisasi data

index	BB_Lahir	TB_Lahir	Usia_Saat_Ukur	Berat	Tinggi	ZS_B B/U	ZS_T B/U	ZS_BB /TB	TB/U
0	0.0001 10	0.099 010	0.233333	0.26 8551	0.382 872	0.571 83	0.299 88	0.0047	2
1	0.0001 17	0.099 010	0.200000	0.19 0813	0.363 980	0.270 42	0.304 55	0.00183	1
2	0.0001 00	0.097 030	0.350000	0.31 0954	0.487 406	0.550 704	0.383 9	0.00402	1
3	0.0000 00	0.000 000	0.650000	0.32 5088	0.540 302	0.322 535	0.199 53	0.00338	2
4	0.0002 33	0.085 149	0.850000	0.08 4806	0.065 491	0.733 803	0.826 14	0.00401	0

Pada Tabel 2 dilakukan normalisasi data menggunakan *MinMax Scalling* pada semua variabel kecuali TB/U kedalam rentang angka 0 sampai 1 dengan tujuan agar nilai setiap variabel tidak memiliki rentang yang terlalu jauh dengan nilai variabel lain.

Pengujian Kernel Support Vector Machine

Pada Tabel 3 kernel yang memiliki tingkat *accuracy* yang tinggi adalah kernel *polynomial* dengan tingkat *accuracy* sebesar 98%. Maka kernel tersebut akan digunakan dalam melakukan pengujian *cross validation* dan *confusion matrix*.

Tabel 3. Hasil Pengujian Kernel *Support Vector Machine*

Kernel	Accuracy
Linear	93%
<i>Polynomial</i>	98%
Sigmoid	46%
<i>Radial Basic Funtion (RBF)</i>	97%

Pengujian Model Teknik Holdout

Pada Tabel 4 Hasil performa *Support Vector Machine* memiliki tingkat *accuracy* sebanyak 98%. Untuk parameter 0 dengan jumlah data normal 86 menghasilkan *precision* 99%, *recall* 97%, *f1-score* 98%. Parameter 1 dengan jumlah data pendek sebanyak 678 menghasilkan *precision* 96%, *recall* 100%, *f1-score* 98%. Sedangkan parameter 2 dengan data sangat pendek sebanyak 217 menghasilkan *precision* 99%, *recall* 88%, *f1-score* 94%.

Tabel 4. Hasil performa *Support Vector Machine*

Class	accuracy	precision	recall	F1-score	Support
0	98%	99%	97%	98%	86
1		96%	100%	98%	678
2		99%	88%	94%	217

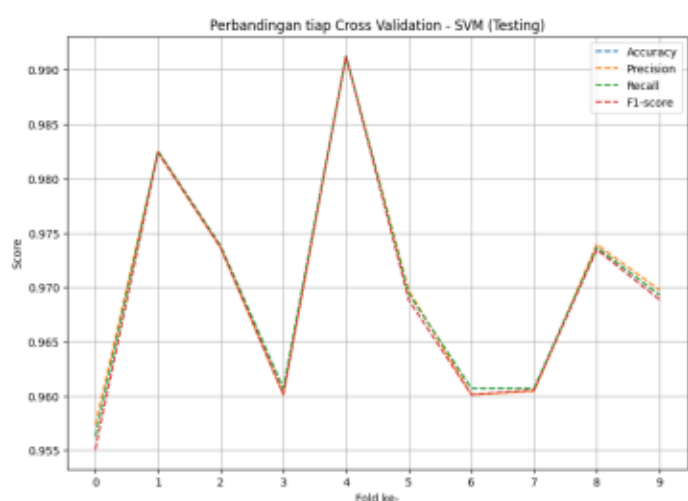
Hasil dari pengujian *K-Fold cross validation* metode *Support Vector Machine* menggunakan kernel *polynomial* menunjukkan hasil dengan nilai tertinggi berada pada *fold* ke-4 dengan tingkat *accuracy* 99.13% *precision* 99.13% *recall* 99.13% *f1-score* 99.13%. sedangkan nilai terendah berada pada *fold* ke-0 dengan tingkat *accuracy* 95.63% *precision* 95.74% *recall* 95.63% *f1-score* 95.51%. untuk rata-rata dari pengujian *fold* menunjukkan hasil *accuracy* 96.98% *precision* 96.99% *recall* 96.98% *f1-score* 96.94%.

Tabel 5. Hasil Pengujian *Confusion Matrix*

Fold Ke -	Accuracy	Precision	Recall	F1-Score
0	95.63%	95.74%	95.63%	95.51%
1	98.25%	98.25%	98.25%	98.24%
2	97.38%	97.37%	97.38%	97.36%
3	96.07%	96.01%	96.07%	96.01%
4	99.13%	99.13%	99.13%	99.13%
5	96.94%	96.97%	96.94%	96.88%
6	96.07%	96.01%	96.07%	96.01%

<i>Fold Ke -</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
7	96.07%	96.04%	96.07%	96.05%
8	97.37%	97.40%	97.37%	97.35%
9	96.93%	96.97%	96.93%	96.89%
Rata-rata	96.98%	96.99%	96.98%	96.94%

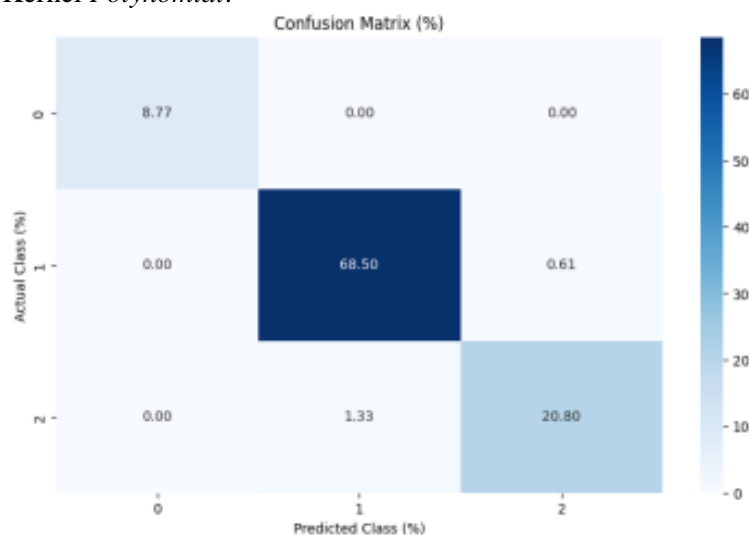
Berdasarkan gambar 3 diatas, grafik pengujian *K-Fold cross validation* sesuai dengan penjelasan tabel 10 dengan hasil *accuracy*, *precision*, *recall* dan *f1-score* terendah berada pada *fold ke-0* dan tertinggi berada pada *fold ke-4*.



Gambar

Gambar 1. Grafik Hasil pengujian *K-Fold cross validation*

Gambar 2. menunjukkan hasil dari proses evaluasi dari pengujian menggunakan *Support Vector Machine* dengan *Kernel Polynomial*.



Gambar 2. Hasil pengujian *confusion matrix*

Berdasarkan gambar 4 diatas, bahwa dari 981 data *testing*, terdapat 8.77% *True Positive* (TP(0.0)), 68.50% *True Positive* (TP(1.1)), dan 20.80% *True Positive* (TP(2.2)). Berdasarkan persamaan (8), maka hasil *accuracy* yang dihasilkan sebanyak 98%. Untuk *class 0* (normal) memiliki nilai *accuracy* 8.77% . Untuk 1 (pendek) *class* dan 2 (sangat pendek) memiliki nilai *accuracy* 91.24%

KESIMPULAN

Berdasarkan dari hasil penelitian dan evaluasi maka dapat diperoleh kesimpulan sebagai berikut yaitu hasil dari penelitian yang dilakukan memperlihatkan bahwa klasifikasi Stunting pada balita menggunakan algoritma *Support Vector Machine* mendapatkan tingkat *accuracy* tinggi. Pada pengujian *Support Vector Machine* dengan 10 *K-Fold Cross Validation* menunjukkan hasil dengan nilai tertinggi berada pada *fold* ke-4 dengan tingkat *accuracy* 99.13% *precision* 99.13% *recall* 99.13% *f1-score* 99.13%. sedangkan nilai terendah berada pada *fold* ke-0 dengan tingkat *accuracy* 95.63% *precision* 95.74% *recall* 95.63% *f1-score* 95.51%. untuk rata-rata dari pengujian *fold* menunjukkan hasil *accuracy* 96.98% *precision* 96.99% *recall* 96.98% *f1-score* 96.94%. Sedangkan untuk hasil dari *Confusion Matrix* dengan nilai *accuracy* sebanyak 98%. Untuk *class* 0 (normal) memiliki nilai *accuracy* 8.77%. Untuk 1 (pendek) *class* dan 2 (sangat pendek) memiliki nilai *accuracy* 89.3%. Dengan demikian dapat disimpulkan bahwa berdasarkan jumlah data testing sebanyak 981 data, balita stunting sebanyak 895 data dan balita normal sebanyak 86 data.

REFERENSI

- Drajana, C., & Bode, A. (2022). Prediksi Status Penderita Stunting Pada Balita Provinsi Gorontalo Menggunakan K-Nearest Neighbor Berbasis Seleksi Fitur Chi Square. *Jurnal Nasional Komputasi dan Teknologi Informasi*, 5(2).
- Byna, A., & Anisa, F. (2018). backward Elimination Untuk Meningkatkan Akurasi Kejadian Stunting Dengan Analisis Algoritma Support Vector Machine. *Dinamika Kesehatan*.
- Riyadh, N., Batara, A., & Nurlinda, A. (2023). Efektivitas Kebijakan dalam Pelaksanaan Program Penanggulangan Stunting di Kabupaten Enrekang. *Journal of Muslim Community Health*, 1-17.
- Usman, A., Rosdiana, & Misnawati, A. (2021). Faktor Resiko Kejadian Perasalinan Prematur di Rumah Sakit Umum Polewali Tahun 2021. *Jurnal Kesehatan*.
- Alhamid, S. (2021). Studi Mengenai Status Gizi Balita. *Jurnal Kebidanan*, 131-138.
- Simanjuntak, F., & Siregar, R. (2020). Hubungan Penerapan Pedoman Gizi Seimbang Dengan Status Gizi Balita Usia 36-60 Bulan di Puskesmas Karang Kitri Bekasi 2018. *Ayurveda Medistra*, 30-37.
- Hidayani, W. (2020). Peran Tenaga Kesehatan Dalam Menurunkan Kejadian Stunting' Tahun 2020 DI INDONESIA : Literature Review. *Prosiding Seminar Nasional Kesehatan* (pp. 45-53). STIKES Respati.
- Kesehatan, K. (2020). *Peraturan Menteri Kesehatan Republik Indonesia Tentang Standar Antropometri Anak*. Jakarta.
- Seno, D., & Wibowo, A. (2019). Analisis Sentimen Data Twitter Tentang Pasangan Capres-Cawapres Pemilu 2019 Dengan Metode Lexicon Based Dan Support Vector Machine. *Jurnal Ilmiah FIFO*, 144-155.
- Gaffar, A., & Sitanggang, I. (2019). Spatial model for predicting sugarcane crop productivity using support vector regression. *IOP Conference Series: Earth and Environmental Science*. Bogor: Institute of Physics Publishing.