

Optimalisasi Algoritma Support Vector Machine (SVM) Dalam Klasifikasi Kejadian Data Stunting

Sonya Dian Wahyuni*, R. Hadapiningradja Kusumodestoni

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika, Universitas Islam Nahdlatul Ulama Jepara, Kota Jepara, Indonesia

Email: ¹sonyadianwahyuni@gmail.com, ²kusumodestoni@gmail.com

Corresponding Author: sonyadianwahyuni@gmail.com

Abstrak- Stunting adalah sebuah kondisi terhambatnya pertumbuhan linier akibat kekurangan gizi kronis, menjadi perhatian serius di Indonesia, terutama pada anak balita. Identifikasi stunting menggunakan parameter z-score tinggi badan terhadap umur (TB/U), dengan nilai kurang dari -2 SD sebagai indikator diagnostik. Faktor-faktor seperti kondisi sosial ekonomi, pola makan, infeksi, dan lingkungan mempengaruhi stunting, menjadikan upaya pencegahan dan intervensi dini sangat penting. Dampak jangka pendek stunting termasuk keterlambatan perkembangan motorik dan gangguan kognitif, sementara dampak jangka panjang meliputi risiko penyakit kronis di masa dewasa. Meskipun prevalensi stunting global menurun, Indonesia masih memiliki angka yang tinggi, menempati peringkat kedua di Asia Tenggara. Kolaborasi lintas sektor dan penerapan teknologi, seperti algoritma Support Vector Machine (SVM), dapat membantu mengidentifikasi faktor risiko stunting. Dalam penelitian ini, kernel SVM linier, polinomial, RBF, dan sigmoid dievaluasi. Kernel RBF SVM terbukti paling efektif, mencapai akurasi lebih dari 90% dengan AUC 0.926. Kolaborasi dan penggunaan teknologi memberikan harapan untuk mengatasi stunting di Indonesia, membutuhkan kerja sama antara pemerintah, lembaga swadaya masyarakat, dan komunitas ilmiah.

Kata Kunci: Stunting;SVM; Kernel Linier;Kernel Polinomial;Kernel RBF;Kernel Sigmoid

Abstract- Stunting is a condition of inhibited linear growth due to chronic malnutrition, which is a serious concern in Indonesia, especially among toddlers. Stunting identification utilizes the parameter of height-for-age z-score (HAZ), with a value of less than -2 SD as a diagnostic indicator. Factors such as socioeconomic conditions, dietary patterns, infections, and the environment influence stunting, making early prevention and intervention efforts crucial. Short-term effects of stunting include delayed motor development and cognitive impairment, while long-term impacts include the risk of chronic diseases in adulthood. Although the global prevalence of stunting has declined, Indonesia still has a high rate, ranking second in Southeast Asia. Cross-sector collaboration and the application of technology, such as Support Vector Machine (SVM) algorithms, can help identify stunting risk factors. In this study, SVM kernels including linear, polynomial, RBF, and sigmoid were evaluated. RBF SVM kernel proved to be the most effective, achieving an accuracy of over 90% with an AUC of 0.926. Collaboration and the use of technology provide hope for addressing stunting in Indonesia, requiring cooperation between the government, NGOs, and the scientific community.

Keywords: Stunting;SVM;Linear Kernel;Polynomial Kernel;RBF Kernel;Sigmoid Kernel

1. PENDAHULUAN

Stunting adalah suatu kondisi terhambatnya pertumbuhan linier yang disebabkan oleh kekurangan gizi kronis, infeksi berulang, atau faktor lainnya. Identifikasi stunting dapat dilakukan dengan mengukur nilai z-score tinggi badan terhadap umur (TB/U), dengan nilai kurang dari -2 SD sebagai parameter diagnostik. [1] Stunting pada anak balita di Indonesia terkait gizi buruk memerlukan perhatian khusus. Intervensi gizi pada bayi stunting hingga usia 2 tahun penting untuk mengatasi dampak jangka panjang dan meningkatkan kesehatan generasi mendatang. Faktor-faktor seperti kondisi sosial ekonomi, asupan makanan, infeksi dan lingkungan mempengaruhi kondisi stunting, sehingga upaya pencegahan dan intervensi dini perlu diperkuat. [2] [3] Stunting mempunyai dampak jangka pendek terhadap keterlambatan perkembangan motorik, gangguan fungsi kognitif, dan peningkatan kerentanan terhadap infeksi pada anak. Risiko jangka panjang antara lain penurunan tinggi badan, penurunan fungsi kognitif, dan peningkatan risiko penyakit kronis seperti diabetes dan penyakit kardiovaskular di masa dewasa. [4]

Berdasarkan data Organisasi Kesehatan Dunia (WHO), terdapat 22% atau 149,2 juta prevalensi stunting di seluruh dunia. Angka tersebut mengalami penurunan dibandingkan tahun 2005 sebanyak 192,3 juta jiwa, tahun 2010 sebanyak 180,5 juta jiwa, dan pada tahun 2015 terus menurun menjadi 163,4 juta jiwa. Pada tahun 2020, kejadian stunting di Asia Selatan memiliki proporsi terbesar yaitu 58,7%, disusul Asia Tenggara (8,2%), Asia Timur (4,8%), dan Asia Barat (3,5%), dengan proporsi terkecil.

Indonesia menduduki peringkat kedua tertinggi di antara negara-negara di Asia Tenggara dengan angka stunting sebesar (31,8%), setelah Timor Leste (48,8%). [5] Berdasarkan laporan hasil Survei Status Gizi Indonesia (SSGI) yang dilakukan Kementerian Kesehatan, prevalensi stunting di Indonesia mengalami fluktuasi dari tahun ke tahun. Pada tahun 2022 menunjukkan sebanyak (21,6%) anak balita akan mengalami stunting. Angka kejadian stunting mengalami penurunan sebesar (2,8%) dari tahun 2021 sebesar (21,6%). [6] Namun masih diperlukan upaya untuk menurunkannya sebesar (3,8%) per tahun agar target (14%) pada tahun 2024 dapat tercapai. [7]

Penanganan masalah stunting memerlukan keterlibatan dan tanggung jawab semua pihak, termasuk ilmu pengetahuan dan teknologi. [8] [9] Penggunaan algoritma Machine Learning untuk mengidentifikasi dan memprediksi faktor risiko utama stunting untuk mengenali potensi risiko malnutrisi. [10] Penggunaan pendekatan data mining dapat memprediksi stunting pada anak. Salah satu metode yang digunakan adalah Support Vector Machine (SVM), sebuah pendekatan yang berkembang dalam ilmu komputer sejak tahun 1990an dan semakin populer saat ini. [11] Telah dilakukan penelitian

mengenai metode prediksi dan penerapan algoritma SVM dengan memanfaatkan sejumlah lembar data yang berbeda-beda. Penelitian tersebut meliputi [12] [13] [14] [15]

KK Kishore dan JV Suman, melakukan penelitian prediksi menggunakan lembar data survei kesehatan dan gizi pada anak di bawah 2 tahun. Data yang digunakan diperoleh dari UNICEF dan mencakup informasi tinggi dan berat badan anak. Dalam penelitian ini, berbagai algoritma seperti SVM, KNN, Logistic Regression, Naive Bayes, dan Bilayered Neural Network. Hasil penelitian menunjukkan bahwa SVM dan Two Layer Neural Network menunjukkan kinerja yang lebih baik dalam mengenali malnutrisi terutama menggunakan Z score seperti LAZ dan WAZ dibandingkan dengan algoritma lainnya. [13] Diagnosis stunting diselidiki oleh [15]. Terdapat 1731 data yang diolah. Hasil akurasi tertinggi pada model SVM dengan kernel RBF sebesar 78,10%. SVM-RBF secara konsisten mendominasi akurasi dibandingkan 10 model lainnya dengan parameter default.

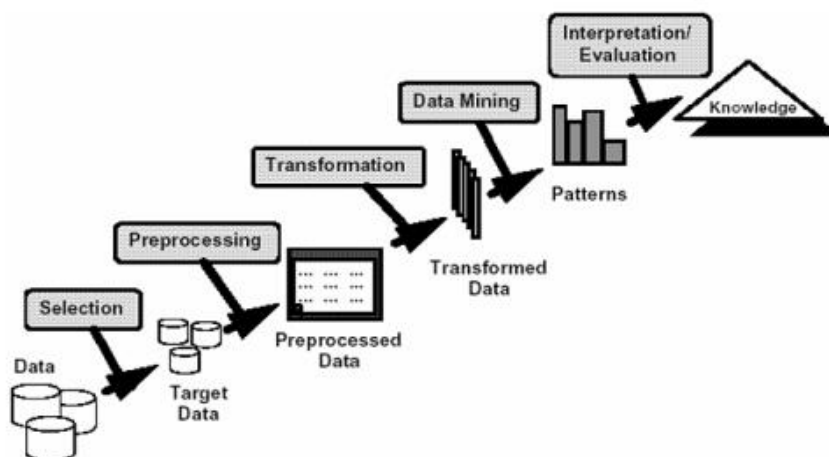
Penggunaan metode data mining mode clustering dapat dilakukan melalui berbagai aplikasi seperti Python, Rapid Miner, R, Orange, dan lain-lain. [16][17] [18] Penelitian menggunakan rapid miner diteliti oleh rapid miner. itu digunakan untuk melaksanakan klasifikasi penerima bantuan fasilitas. hasil clustering sehingga cluster ke 3 yang masuk dalam cluster prioritas mendapat bantuan. [19] Penelitian selanjutnya yaitu menggunakan Python, python dipilih untuk mengembangkan library kernelmethods karena dukungan pustaka ilmu data yang kuat seperti NumPy dan scikit-learn. Evaluasi menunjukkan bahwa kernelmethods memiliki cakupan pengujian 95% dan kinerja yang sebanding dengan scikit-learn, menyederhanakan implementasi metode kernel dalam pembelajaran mesin.[20]Kajian optimasi algoritma Support Vector Machine (SVM) untuk klasifikasi data stunting dengan Particle Swarm Optimization (PSO) meningkatkan kinerja SVM, dengan kernel Radial Basic Function terbukti paling efektif (Akurasi 78%, Presisi 89%, 66% Ingat, dan 72% Skor F1). [21]SVM sangat penting untuk mengidentifikasi dan memprediksi faktor risiko malnutrisi pada anak, sehingga membantu upaya mengatasi stunting di Indonesia. Kolaborasi antar pemangku kepentingan, termasuk komunitas ilmiah, sangat penting untuk mengatasi kompleksitas stunting. Teknik SVM dan optimasi dapat meningkatkan intervensi pemerintah dalam klasifikasi malnutrisi. Penelitian ini menggunakan lembar data akademik dan perilaku siswa, hasil SVM-POLY dengan nilai C sebesar 0,1, nilai γ sebesar 0,005, dan nilai derajat sebesar 1, model SVM-POLY ini meningkat secara signifikan dari akurasi 90,5% menjadi 95,2%, dapat memberikan kinerja terbaik dalam mengklasifikasikan kinerja siswa selama kegiatan pembelajaran online.[22]

Penelitian ini bertujuan untuk mengatasi permasalahan stunting dan gizi buruk pada anak di Indonesia dengan menggunakan algoritma pembelajaran, khususnya Support Vector Machines (SVM), untuk mengklasifikasikan data status gizi dan memprediksi faktor risiko stunting, wasting, dan underweight pada anak. Python dipilih sebagai bahasa pemrograman utama karena fleksibilitasnya dan dukungan terhadap berbagai fungsi kernel seperti linier, polinomial, sigmoid, dan RBF. Penggunaan Python memungkinkan penanganan yang efektif terhadap berbagai jenis data dan kendala pengambilan keputusan. Diharapkan, dengan memanfaatkan variasi fungsi kernel ini, SVM dapat meningkatkan kinerjanya dalam mengklasifikasikan data nonlinier terkait status gizi anak, serta berkontribusi pada upaya meningkatkan kesehatan dan kesejahteraan anak di Indonesia yang terkena dampak stunting dan gizi buruk.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian memanfaatkan Knowledge Discovery in Database (KDD). Kompleksitas proses KDD ditekankan, tidak hanya mencakup analisis data tetapi juga persiapan data, interpretasi pengetahuan, dan penerapan informasi yang baru diperoleh. [23]



Gambar 1. Metodologi Penelitian

Proses data mining merupakan tahap tunggal dalam perkembangan berulang prosedur KDD, seperti yang digambarkan pada Gambar 1. Perkembangan ini dimulai dengan identifikasi kumpulan data dan pembuatan model, kemudian dilanjutkan melalui tahapan pengumpulan data, pemrosesan, transformasi, data mining, dan interpretasi.[24]

Dataset yang digunakan adalah data stunting dari laman <https://www.kaggle.com/datasets/muhtarom/stunting>. Datasheet berisi 6500 data. Beberapa atribut yang dimanfaatkan dalam penelitian ini meliputi:

- a. Sex : Jenis Kelamin
- b. Age : Umur
- c. Birth Weight : Berat Badan Lahir
- d. Birth Leght : Panjang Badan Lahir
- e. Body Weight : Berat Badan
- f. Body Length : Panjang Badan
- g. Asi Eksklusif
- h. Outcome Stunting

2.2 Support Vector Machine (SVM)

Metode Support Vector Machine (SVM) merupakan salah satu teknik yang efektif dalam melakukan prediksi baik dalam pengklasifikasian maupun regresi. Thorsten Joachim adalah salah satu yang pertama kali mengusulkan penggunaan algoritma SVM untuk klasifikasi teks dengan menggunakan bobot indeks term sebagai fitur. Penggunaan SVM dalam pembelajaran telah diperkenalkan sejak tahun 1992 oleh Boser, Guyon, dan Vapnik. SVM mampu menangani permasalahan baik yang bersifat linier maupun non-linier.[25]

Dalam algoritma SVM, terdapat kernel di mana terdapat SVM linear dan SVM nonlinier. SVM hyperplane linear beroperasi hanya pada data yang dapat dipisahkan dengan cara linier dengan kernel linier, SVM mampu membangun hyperplane linier yang efisien memisahkan dua kelas data dalam ruang fitur berdimensi tinggi, memungkinkan klasifikasi yang akurat pada data yang dapat dipisahkan secara linier, sedangkan SVM nonlinier digunakan untuk data yang berdistribusi pada kelas yang tidak linier dengan menerapkan pendekatan kernel pada fitur awal set. Contoh kernel non-linier yang sering digunakan dalam SVM meliputi kernel polinomial, kernel berbasis radial (RBF), dan kernel sigmoid.[26] [27] Fungsi kernel memungkinkan pemetaan data ke ruang berdimensi lebih tinggi secara efisien, memfasilitasi pemisahan linier dan non-linier. Setiap jenis fungsi kernel memiliki parameter spesifik yang memerlukan penyetelan optimal untuk mencapai hasil optimal. Fungsi Kernel sebagai berikut:[28]

2.2.1 Kernel Linier

Fungsi kernel linier dalam mesin vektor dukungan (SVM) merupakan jenis fungsi kernel yang memetakan data ke dalam dimensi fitur yang lebih tinggi. Fungsi ini sederhana dan tidak melakukan transformasi nonlinier pada data, hanya memetakan data ke dalam dimensi fitur yang sama. Kernel linier sesuai digunakan saat data dapat dipisahkan secara linier dalam dimensi fitur asli. Meskipun simpel, kernel linier tetap efektif dalam banyak kasus klasifikasi data linier.

Parameter yang umum terkait dengan kernel linier adalah parameter bias (c), yang mengatur posisi batas keputusan SVM. Dalam kernel linier, nilai c dapat memengaruhi margin dan letak hyperplane yang memisahkan kelas data. Dengan menggunakan kernel linier, SVM mampu menangani data yang dapat dipisahkan secara linier dalam dimensi fitur asli tanpa perlu melakukan transformasi nonlinier yang rumit.

2.2.2 Kernel Polinomial

Gangguan data atau representasi fitur yang buruk kadang membuatnya sulit untuk memisahkan data secara linier. Untuk mengatasi hal ini, data dapat dipetakan ke dalam ruang dimensi yang lebih tinggi menggunakan fungsi kernel polinomial. Fungsi kernel polinomial digunakan baik untuk hard margin maupun soft margin dalam menangani pola nonlinier. Parameter bebas seperti c digunakan dalam polinomial untuk menyeimbangkan efek antara suku tingkat tinggi dan rendah. Selain itu, terdapat parameter optimasi lainnya seperti d , γ , c , dan $\bar{\gamma}$ yang perlu diperhatikan. Polinomial dianggap sebagai kernel non-stasioner yang cocok digunakan ketika seluruh sampel pelatihan sudah dinormalisasi. Pengaturan parameter kemiringan γ juga penting dalam penggunaan kernel ini. Dalam konteks ini, c merepresentasikan suku konstanta dan d adalah derajat polinomial yang digunakan.

2.2.3 Kernel RBF

Fungsi basis Radial, juga dikenal sebagai kernel RBF atau kernel Gaussian, menggunakan parameter γ untuk mengatur sebaran kernel. Penyetelan γ harus dilakukan dengan hati-hati karena berpengaruh besar terhadap kinerja kernel ini. Jika γ diestimasi terlalu tinggi, perilaku eksponensialnya bisa menjadi linier dan menyebabkan kehilangan kekuatan nonliniernya dalam proyeksi dimensi yang lebih tinggi. Kernel RBF memiliki kekurangan dalam regularisasi dan keputusan batasannya cenderung sangat sensitif terhadap gangguan dalam data pelatihan jika tidak diperhatikan. Oleh karena itu, pemilihan nilai parameter γ harus dilakukan dengan cermat karena berdampak pada kinerja SVM.

2.2.4 Kernel Sigmoid

Fungsi kernel sigmoid merupakan alat dalam SVM yang memungkinkan transformasi data ke dalam bentuk nonlinier menggunakan kurva sigmoid. Ini sangat berguna ketika data tidak dapat dipisahkan secara linier dalam dimensi fitur asli. Dengan mengatur parameter gamma dan konstanta c secara optimal, SVM dapat menghasilkan hasil klasifikasi yang optimal untuk data yang kompleks dan tidak linier. Namun, pengaturan parameter ini memerlukan perhatian khusus karena dapat mempengaruhi kinerja keseluruhan SVM dengan fungsi kernel sigmoid. Oleh karena itu, pemilihan parameter yang tepat sangat penting untuk memastikan SVM dapat menangani pola nonlinier dengan efektif dan memberikan hasil klasifikasi yang akurat.

3. HASIL DAN PEMBAHASAN

3.1 Penentuan Library yang digunakan

Pemilihan library yang akan digunakan dalam proses pengembangan model data mining. Pemilihan library ini penting karena akan memengaruhi kemudahan dalam manipulasi data, implementasi algoritma machine learning, dan evaluasi hasil model. Pada tahap ini, kita akan menggunakan tiga library utama, yaitu Pandas, NumPy, dan Scikit-learn.

- Pandas: Untuk membaca dan memanipulasi data tabular dari berbagai format file.
- NumPy: Untuk komputasi numerik, termasuk transformasi data dan operasi matematika pada array.
- Scikit-learn: Menyediakan algoritma ML dan alat evaluasi model.

Di sini, kita akan membangun model SVM. Penggunaan ketiga pustaka ini memanfaatkan kelebihan masing-masing dan integrasinya yang baik, memudahkan dalam pengembangan model data mining secara efisien.

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn import svm
from sklearn.metrics import accuracy_score, classification_report
```

Gambar 2. Penentuan Library

3.2 Load dan Standarisasi Dataset

Untuk memulai, dataset diunduh dari Kaggle dan dibaca menggunakan pustaka Pandas. Beberapa baris pertama dari dataset ditampilkan untuk pemeriksaan awal.

```
import pandas as pd
stunting_dataset = pd.read_csv('DataSetStunting.csv')

stunting_dataset.head()

stunting_dataset.shape

X = standardized_data
Y = stunting_dataset['Outcome']

print(X)
print(Y)
```

	Sex	Age	Birth Weight	Birth Length	Body Weight	Body Length	ASI Eksklusif	Outcome
0	1	56	2.9	50	11.0	90.0	1	0
1	1	20	3.3	49	11.1	80.5	0	0
2	0	4	2.8	48	6.5	63.0	0	0
3	1	14	2.0	49	7.0	71.0	1	0
4	0	32	3.2	49	11.0	88.7	1	0

Gambar 3. Load dan Standarisasi Dataset

Standarisasi data penting untuk memastikan konsistensi dan performa model yang lebih baik. Untuk melakukannya, digunakan StandardScaler dari pustaka scikit-learn untuk menormalkan data. Proses standarisasi ini dilakukan khususnya pada atribut-atribut numerik dari dataset.

3.3 Pemisahan Data Training dan Data Testing

Langkah ini adalah membagi dataset menjadi dua bagian terpisah: data training dan data testing. Tindakan ini krusial untuk mengevaluasi kinerja model dengan tepat. Dengan memisahkan dataset, kita dapat melatih model menggunakan data training dan menguji performanya menggunakan data testing yang terpisah. endekatan ini memastikan bahwa model tidak hanya dapat mempelajari pola yang ada dalam data pelatihan tetapi juga dapat menggeneralisasikan pengetahuannya dengan baik ke data baru yang belum pernah dilihat sebelumnya.

```
[19] X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size= 0.2, stratify=Y, random_state=2)
✓ 0.0s Python

[20] print(X.shape, X_train.shape, X_test.shape)
✓ 0.0s Python

... (6500, 7) (5200, 7) (1300, 7)
```

Gambar 4. Pemisahan Data Training dan Data Testing

3.4 Tahapan Pemodelan Data Mining Algoritma SVM

Fase pemodelan merupakan tahapan yang krusial dalam proses data mining, yang memanfaatkan metode SVM (Support Vector Machine), dengan Rapidminer sebagai perangkat utamanya. Tujuan utama dari pengujian model adalah untuk melakukan optimalisasi terhadap Algoritma Support Vector Machine (SVM) dalam melakukan klasifikasi terhadap Kejadian Data Stunting, dengan fokus untuk mendapatkan tingkat akurasi yang optimal.

3.4.1 Kernel Linier

Dalam penggunaan kernel linier, iterasi terhadap nilai parameter C dilakukan dengan menggunakan rentang nilai tertentu, yaitu 10 pangkat -1,0 dan 1. Setelah melakukan iterasi, evaluasi performa model menggunakan teknik validasi silang dilakukan untuk menentukan nilai parameter C yang optimal. Evaluasi dilakukan menggunakan confusion matrix untuk mengukur akurasi klasifikasi, memastikan bahwa nilai parameter C yang dioptimalkan menghasilkan peningkatan signifikan dalam akurasi klasifikasi dibandingkan dengan pengaturan default.

```
classifier = svm.SVC(kernel='linear', C=1, gamma='auto')
[23] ✓ 0.0s Python

classifier.fit(X_train, Y_train)
[24] ✓ 0.0s Python

...
* SVC
SVC(C=1, gamma='auto', kernel='linear')

krn=['linier']
rng_C=np.arange(0.125,0,10)
[25] ✓ 0.0s Python
```

Gambar 5. Optimasi Kernel Linier

3.4.2 Kernel Polinomial

Pada penggunaan kernel polinomial, terdapat tiga parameter yang dioptimalkan, yaitu C, gamma, dan degree. Rentang nilai untuk setiap parameter disesuaikan, misalnya dari parameter C yaitu 0,1, 0, dan 10, parameter gamma yaitu 10 pangkat -3 hingga suku pangkat -1 dan parameter degree yaitu 2 sampai 3. Melalui proses iterasi, kombinasi nilai parameter yang menghasilkan kinerja terbaik dicari.

```
classifier = svm.SVC(kernel='poly')
[24] ✓ 0.0s Python

classifier.fit(X_train, Y_train)
[25] ✓ 0.0s Python

...
* SVC
SVC(kernel='poly')

krn=['poly']
rng_C=np.arange(0.1,0,10)
rng_gamma=np.arange(0.125,0.25,0.5)
rng_degree=np.arange(2,3)
[26] ✓ 0.0s Python
```

Gambar 6. Optimasi Kernel Polinomial

3.4.3 Kernel RBF

Dalam penggunaan kernel RBF, dua parameter utama yang disesuaikan adalah C dan gamma. Rentang nilai untuk masing-masing parameter dapat ditentukan. Rentang nilai untuk setiap parameter disesuaikan, misalnya dari parameter C yaitu 0,1, 0, dan 10 dan parameter gamma yaitu 10^{-3} hingga 10^{-1} . Melalui proses iterasi, nilai parameter yang menghasilkan kinerja terbaik dapat diidentifikasi, memungkinkan peningkatan akurasi model.

```

classifier = svm.SVC(kernel='rbf')

[23] ✓ 0.0s Python

classifier.fit(X_train, Y_train)

[24] ✓ 0.4s Python

...

svm.SVC()

...

kernel['rbf']
rng_C=np.arange(0.1,0,10)
rng_gamma=np.arange(0.125,0.25,0.5)

[25] ✓ 0.0s Python

```

Gambar 7. Optimasi Kernel RBF

3.4.4 Kernel Sigmoid

Dalam penggunaan kernel Sigmoid, dua parameter utama yang disesuaikan adalah C dan gamma. Rentang nilai untuk masing-masing parameter dapat ditentukan. Rentang nilai untuk setiap parameter disesuaikan, misalnya dari parameter C yaitu 0,1, 0, dan 10 dan parameter gamma yaitu 10^{-3} hingga 10^{-1} . Melalui proses iterasi, nilai parameter yang menghasilkan kinerja terbaik dapat diidentifikasi, memungkinkan peningkatan akurasi model.

```

classifier = svm.SVC(kernel='sigmoid')

[21] ✓ 0.0s Python

> classifier.fit(X_train, Y_train)

[22] ✓ 0.5s Python

...

svm.SVC(kernel='sigmoid')

...

kernel['sigmoid']
rng_C=np.arange(0.1,0,10)
rng_gamma=np.arange(0.125,0.25,0.5)

[23] ✓ 0.0s Python

```

Gambar 8. Optimasi Kernel Sigmoid

3.5 Model Evaluasi

Tahap evaluasi model adalah memverifikasi nilai kegunaan dari model yang telah dibuat pada tahap sebelumnya. Dalam pengujian model untuk mengevaluasi dan mendapatkan nilai akurasi.

3.5.1 Hasil Pengujian Algoritma SVM

Pengujian model yang telah dilakukan menggunakan 4 (empat) tipe kernel yang dimiliki SVM, yaitu:

a. Kernel Linier

Uji coba model SVM dengan Kernel Linier yang memanfaatkan variabel Cost telah dilakukan, dan berikut adalah laporan hasil dari sejumlah percobaan yang telah dilakukan:

```

> print(classification_report(X_train_prediction, Y_train))
print("akurasi SVM : {:.2f}%".format(training_data_accuracy*100))
print('Akurasi data training adalah = ', training_data_accuracy)

[27] ✓ 0.0s

...

precision    recall  f1-score   support

      0       0.79      0.84      0.82     2412
      1       0.85      0.81      0.83     2788

 accuracy      0.82      0.83      0.82     5200
 macro avg      0.82      0.83      0.82     5200
 weighted avg    0.83      0.82      0.82     5200

akurasi SVM : 82.42%
Akurasi data training adalah = 0.8242307692307692

```

Gambar 9. Akurasi Kernel Linier

Hasil optimal dari eksperimen SVM dengan kernel Polinomial yang disebutkan di atas diperoleh tingkat akurasi mencapai 82,42% dengan Area Under Curve (AUC) sebesar 0,824.

b. Kernel Polinomial

Uji coba model SVM dengan Kernel Polinomial yang memanfaatkan variabel Cost, Gamma, dan Degree telah dilakukan, dan berikut adalah laporan hasil dari sejumlah percobaan yang telah dilakukan:

```
print(classification_report(X_train_prediction, Y_train))
print("akurasi SVM : {:.2f}%".format (training_data_accuary*100))
print('Akurasi data training adalah = ', training_data_accuary)
```

	precision	recall	f1-score	support
0	0.80	0.90	0.85	2280
1	0.91	0.83	0.87	2920
accuracy			0.86	5200
macro avg	0.86	0.86	0.86	5200
weighted avg	0.86	0.86	0.86	5200

akurasi SVM : 85.77%
Akurasi data training adalah = 0.8576923076923076

Gambar 10. Akurasi Kernel Polinomial

Hasil optimal dari eksperimen SVM dengan kernel Polinomial yang disebutkan di atas diperoleh tingkat akurasi mencapai 85,77% dengan Area Under Curve (AUC) sebesar 0,857.

c. Kernel RBF

Uji coba model SVM dengan Kernel RBF yang memanfaatkan variabel Cost dan Gamma telah dilakukan, dan berikut adalah laporan hasil dari sejumlah percobaan yang telah dilakukan:

```
print(classification_report(X_train_prediction, Y_train))
print("akurasi SVM : {:.2f}%".format (training_data_accuary*100))
print('Akurasi data training adalah = ', training_data_accuary)
```

	precision	recall	f1-score	support
0	0.90	0.94	0.92	2440
1	0.95	0.91	0.93	2760
accuracy			0.93	5200
macro avg	0.93	0.93	0.93	5200
weighted avg	0.93	0.93	0.93	5200

akurasi SVM : 92.65%
Akurasi data training adalah = 0.9265384615384615

Gambar 11. Akurasi Kernel RBF

Hasil optimal dari eksperimen SVM dengan kernel RBF yang disebutkan di atas diperoleh tingkat akurasi mencapai 92,65% dengan Area Under Curve (AUC) sebesar 0,926.

d. Kernel Sigmoid

Uji coba model SVM dengan Kernel Sigmoid yang memanfaatkan variabel Cost dan Gamma telah dilakukan, dan berikut adalah laporan hasil dari sejumlah percobaan yang telah dilakukan:

```
print(classification_report(X_train_prediction, Y_train))
print("akurasi SVM : {:.2f}%".format (training_data_accuary*100))
print('Akurasi data training adalah = ', training_data_accuary)
```

	precision	recall	f1-score	support
0	0.63	0.62	0.62	2611
1	0.62	0.64	0.63	2589
accuracy			0.63	5200
macro avg	0.63	0.63	0.63	5200
weighted avg	0.63	0.63	0.63	5200

akurasi SVM : 62.71%
Akurasi data training adalah = 0.6271153846153846

Gambar 12. Akurasi Kernel Sigmoid

Hasil optimal dari eksperimen SVM dengan kernel Sigmoid yang disebutkan di atas diperoleh tingkat akurasi mencapai 62,71% dengan Area Under Curve (AUC) sebesar 0,627.

4. KESIMPULAN

Stunting adalah kondisi terhambatnya pertumbuhan linier anak akibat gizi buruk, memerlukan perhatian khusus di Indonesia. Meskipun prevalensi stunting secara global telah menunjukkan penurunan, namun Indonesia masih memiliki angka yang tinggi. Kolaborasi lintas sektor dan penerapan teknologi, termasuk algoritma machine learning seperti Support Vector Machine (SVM), dapat membantu dalam mengidentifikasi dan memprediksi faktor risiko stunting. Penggunaan metode data mining dan teknologi memberikan peluang untuk meningkatkan pemahaman dan intervensi terhadap stunting. Dalam hasil penelitian ini menggunakan SVM dengan kernel RBF telah berhasil menghasilkan model yang efektif dalam mengklasifikasikan stunting dengan tingkat akurasi yang tinggi (92,65%) dan Area Under Curve (AUC) sebesar 0,926. Penggunaan teknologi ini membuka pintu untuk penelitian dan intervensi lebih lanjut terhadap stunting di Indonesia. Diperlukan kerjasama antara berbagai pihak, termasuk komunitas ilmiah, pemerintah, dan lembaga swadaya masyarakat, untuk mengatasi masalah kompleks stunting dan memastikan kesehatan anak-anak Indonesia di masa mendatang.

REFERENCES

- [1] L. Simamora, L. Simorangkir, and E. S. Rajagukguk, "Edukasi Pencegahan Stunting Dengan Metode Penyuluhan Di Puskesmas Sijamapolang Kabupaten Humbang Hasundutan Stunting Prevention Education Using Counseling Methods At The Sijamapolang Community Health Center , Humbang Hasundutan Regency," vol. 2, no. 1, 2024.
- [2] H. Hatijar, "Angka Kejadian Stunting Pada Bayi dan Balita Pendahuluan," pp. 12–17, 2023.
- [3] S. A. A. P. S. K. Haryani, "Pencegahan Stunting Melalui Pemberdayaan," *J. Pengabd. Kesehat.*, vol. 4, no. 1, p. 30, 2021.
- [4] D. J. Raiten and A. A. Bremer, "Exploring the nutritional ecology of stunting: New approaches to an old problem," *Nutrients*, vol. 12, no. 2, 2020, doi: 10.3390/nu12020371.
- [5] UNICEF/WHO/WORLD BANK, "Levels and trends in child malnutrition UNICEF / WHO / World Bank Group Joint Child Malnutrition Estimates Key findings of the 2021 edition," *World Heal. Organ.*, pp. 1–32, 2021, [Online]. Available: <https://www.who.int/publications/i/item/9789240025257>
- [6] S. Handayani and J. Kebidanan, "SAVE THE NATION ' S GENERATION FROM THE DANGERS OF STUNTING," vol. 3, pp. 87–92, 2024, doi: 10.36082/jmswh.v3i2.1082.
- [7] Sukmalalana;, D. Putra;, Hafiz;, and V. Cika, "Percepatan Penurunan Stunting Untuk Mewujudkan Sumber Daya Manusia Indonesia Yang Unggul," *Pus. Kaji. Akuntabilitas Keuang. Negara*, vol. 1, no. 1, pp. 1–15, 2022, [Online]. Available: <https://berkas.dpr.go.id/puskajiakn/analisis-ringkas-cepat/public-file/analisis-ringkas-cepat-public-52.pdf>
- [8] I. K. A. Wiraguna, E. Setyati, and E. Pramana, "Prediksi Anak Stunting Berdasarkan Kondisi Orang Tua Dengan Metode Support Vector Machine Dengan Study Kasus Di Kabupaten Tabanan-Bali," *Smatika J.*, vol. 12, no. 01, pp. 47–54, 2022, doi: 10.32664/smatika.v12i01.662.
- [9] W. Achmad, "Social Reality Stunting Prevention in Cianjur District," *J. EduHealth*, vol. 13, no. 02, pp. 467–477, 2022, [Online]. Available: <http://ejournal.seaninstitute.or.id/index.php/health/article/view/575%0Ahttp://ejournal.seaninstitute.or.id/index.php/health/article/download/575/463>
- [10] S. M. J. Rahman *et al.*, "Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach," *PLoS One*, vol. 16, no. 6 June 2021, pp. 1–11, 2021, doi: 10.1371/journal.pone.0253172.
- [11] I. Rahmi, M. Susanti, H. Yozza, and F. Wulandari, "Classification of Stunting in Children Under Five Years in Padang City Using Support Vector Machine," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 3, pp. 771–778, 2022, doi: 10.30598/barekengvol16iss3pp771-778.
- [12] R. Kusumaningrum, T. A. Indihatmoko, S. R. Juwita, A. F. Hanifah, K. Khadijah, and B. Surarso, "Benchmarking of multi-class algorithms for classifying documents related to stunting," *Appl. Sci.*, vol. 10, no. 23, pp. 1–13, 2020, doi: 10.3390/app10238621.
- [13] K. K. Kishore and J. V. Suman, "Prediction of malnutrition in newborn Infants using machine learning techniques," 2023.
- [14] A. N. N. Azmi, S. Khairunniza-Bejo, M. Jahari, F. M. Muharram, and I. Yule, "Identification of a suitable machine learning model for detection of asymptomatic Ganoderma boninense infection in oil palm seedlings using hyperspectral data," *Appl. Sci.*, vol. 11, no. 24, 2021, doi: 10.3390/app112411798.
- [15] S. Syahrial, R. Ilham, Z. F. Asikin, and S. S. I. Nurdin, "Stunting Classification in Children's Measurement Data Using Machine Learning Models," *J. La Multiapp*, vol. 3, no. 2, pp. 52–60, 2022, doi: 10.37899/journallamultiapp.v3i2.614.
- [16] M. D. Kurnia, "Klasifikasi Customer Relationship Management Perusahaan Telekomunikasi Seluler Dengan Metode Machine Learning," vol. 1, no. 4, pp. 63–76, 2023.
- [17] C. Aldama and M. Nasir, "Klasifikasi Penyakit Diabetes Menggunakan Metode Support Vector Machine Pada Rumah Sakit Umum Prabumulih," *J. Ilm. Betrik*, vol. 14, no. 02, pp. 376–383, 2023, [Online]. Available: <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/117>
- [18] M. Edris Effendi, I. Yuadi, and I. Puspitasari, "Prediksi Guru Kemungkinan Tetap Bekerja di Sekolah Al Uswah Surabaya Menggunakan Machine Learning," *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 129–137, 2023, doi: 10.37034/jidt.v5i2.361.
- [19] R. Hadapinigradja Kusumodestoni, "The 2 st Seminar Nasional dan Prosiding Scitech 2023 OCTAGONAL SHOE RACK DESIGN WITH DESIGN IDEAS," pp. 218–223, 2023.
- [20] P. R. Raamana, "Kernel methods library for pattern analysis and machine learning in python," vol. 1, pp. 1–7, 2020, [Online]. Available: <http://arxiv.org/abs/2005.13483>
- [21] S. Y. Andriyani, M. S. Lydia, and S. Efendi, "Optimization of Support Vector Machine Algorithm Using Stunting Data Classification," *Prism. Sains J. Pengkaj. Ilmu dan Pembelajaran Mat. dan IPA IKIP Mataram*, vol. 11, no. 1, p. 164, 2023, doi: 10.33394/j-ps.v11i1.6619.
- [22] M. H. Mohd Zaki, M. A. Mohd Aziz, S. Sulaiman, and N. Hambali, "Student Performance Classification using Support Vector Machine (SVM) with Polynomial Kernel on Online Student Activities," *J. Electr. Electron. Syst. Res.*, pp. 80–90, 2023, doi: 10.24191/jeesr.v23i1.009.
- [23] A. M. Abdulkadum, R. A. A. Shekan, and H. A. Hussain, "Application of Data Mining and Knowledge Discovery in Medical Databases," *Webology*, vol. 19, no. 1, pp. 4912–4924, 2022, doi: 10.14704/web/v19i1/web19329.
- [24] A. H. Nasyuha *et al.*, "Frequent pattern growth algorithm for maximizing display items," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 19, no. 2, pp. 390–396, 2021, doi: 10.12928/TELKOMNIKA.v19i2.16192.

- [25] A. Zy and Wahyu Hadikristanto, "Implementasi Algoritma Metode Naive Bayes dan Support Vector Machine Tentang Pembobolan dan Kebocoran Data di Twitter," *Bull. Inf. Technol.*, vol. 4, no. 1, pp. 49–56, 2023, doi: 10.47065/bit.v4i1.493.
- [26] I. Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin," *J. Responsif Ris. Sains dan Inform.*, vol. 4, no. 2, pp. 118–125, 2022, doi: 10.51977/jti.v4i2.762.
- [27] I. S. Al-Mejibli, J. K. Alwan, and D. H. Abd, "The effect of gamma value on support vector machine performance with different kernels," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 5, pp. 5497–5506, 2020, doi: 10.11591/IJECE.V10I5.PP5497-5506.
- [28] J. Shawe-Taylor and S. Sun, "Kernel Methods and Support Vector Machines," pp. 857–881, 2020, doi: 10.1016/b978-0-12-396502-8.00016-4.