PERBANDINGAN METODE ALGORITMA SUPERVISED NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI PENDERITA STUNTING DI KABUPATEN DELI SERDANG

Lidia Pebrianti, Elmanani Simamora, Sudianto Manullang, Insan Taufiq, Chairunisah
Ilmu Komputer, Universitas Negeri Medan
Jalan Willem Iskandar, Pasar V Medan Estate, Percut Sei Tuan, Deli Serdang, Sumatera Utara, Indonesia
lidiapebrianti503@gmail.com

ABSTRAK

Fenomena gizi buruk di Indonesia dipengaruhi oleh kualitas kesehatan sumber daya manusia (SDM) yang rendah, yang umumnya disebabkan oleh konsumsi pangan yang tidak seimbang. Kondisi ini terlihat pada anak-anak yang mengalami gangguan fisik dan psikis akibat kurangnya asupan gizi yang dibutuhkan. Dalam penelitian ini, dilakukan perbandingan antara dua algoritma, yaitu Naive Bayes dan Support Vector Machine (SVM), untuk mengklasifikasikan penderita stunting di Kabupaten Deli Serdang. Penelitian menggunakan dataset dari Dinas Kesehatan Deli Serdang, dengan jumlah data sebanyak 422 sampel. Data tersebut dibagi menjadi beberapa split, yaitu 60:40, 70:30, 80:20, dan 90:10, untuk melihat performa masing-masing algoritma pada proporsi pembagian data yang berbeda. Hasil penelitian menunjukkan bahwa akurasi klasifikasi menggunakan Support Vector Machine (SVM) lebih tinggi daripada dengan hasil Naive Bayes. Dari pengujian variansi, didapati bahwa Support Vector Machine (SVM) memiliki variansi sebesar 2,87%, sementara Naive Bayes menunjukkan variansi sebesar 28,35%. Hal ini menandakan bahwa Support Vector Machine (SVM) lebih konsisten kinerjanya di berbagai pembagian dataset. Dengan kata lain, Support Vector Machine (SVM) tidak hanya lebih akurat tetapi juga lebih stabil dalam performanya dibandingkan dengan Naive Bayes, menegaskan keunggulannya sebagai metode yang lebih andal untuk kasus klasifikasi stunting ini

Kata kunci : *Naive Bayes, Support Vector Machine (SVM), Stunting, Gizi Buruk*

1. PENDAHULUAN

Stunting merupakan gangguan pertumbuhan pada masa tubuh balita. Melonjaknya angka stunting di Indonesia sendiri termasuk kedalam barisan negara ketiga dengan kategori stunting tertinggi di Asia Tenggara dengan angka rata-rata yang mencapai 36,4% kasus [1].

Stunting dapat terjadi dimana pertumbuhan bayi tidak dapat bertumbuh secara optimal karena kurangnya asupan gizi pada tubuh anak dalam 1000 hari pertama sejak bayi dilahirkan. Stunting dapat mengakibatkan beberapa dampak yang cukup serius, baik dalam jangka pendek maupun dalam jangka panjang. Jangka pendek stunting dapat mengakibatkan efek gagal berkembang dan tumbuh, terhambatnya perkembangan kognitif dan motorik sehingga berpengaruh pada perkembangan otak, serta penyakit metabolisme anak dan ukuran tubuh yang kurang ideal. Disisi lain, stunting memiliki dampak jangka panjang yang mencakup penurunan kapasitas intelektual, gangguan permanen pada struktur tubuh dan fungsi sel saraf dan otak, seta menurunnya pemahaman pembelajaran pada usia sekolah, yang semuanya dapat berdampak pada pertumbuhan anak [2].

Berdasarkan permasalahan yang terjadi, penelitian ini membuat suatu pemodelan dengan mengandalkan machine learning metode naïve bayes dan Support Vector Machine (SVM) [3].

Dengan kemajuannya teknologi informasi yang berkembang sangat pesat di segala bidang kehidupan, sangat berbanding lurus dengan data yang dihasilkan. Mulai dari bidang industri, kesehatan dan berbagai bidang lainnya. Dengan penerapan teknologi informasi di dunia kesehatan dan medis mampu menghasilkan data yang berlimpah [4], Beberapa penelitian tentang prediksi di bidang kesehatan dan prediksi status gizi buruk dengan menggunakan algoritma berbasis machine learning telah banyak dipublikasikan dalam literature.

Dalam penelitian [5], dalam menganalisis perbandingan metode naïve bayes dan metode Support Vector Machine (SVM) untuk mengelompokkan jumlah pembaca, dan menghasilkan nilai akurasi metode Support Vector Machine (SVM) adalah yang terbaik tanpa begging yaitu 63,39% dengan 9 kali percobaan.

Metode naïve bayes mendapatkan hasil 57,10% dengan 10 kali percobaan. Contoh lainnya dalam penelitian [6] yang berjudul 'Comparison of Naive Bayes Classifier and Support Vector Machine Algorithms for Classifying Student Mental Health Data Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Data Kesehatan Mental Mahasiswa' menuliskan bahwa hasil dari metode naïve bayes dan Support Vector Machine (SVM) menghasilkan nilai akurasi sebesar 94,37 untuk Support Vector Machine (SVM) dimana hasil ini menjadi nilai paling optimal daripada dengan nilai akurasi naïve bayes sebesar 86,87%.

Dalam penelitian [7] yang berjudul 'Implementasi Algoritma Metode Naive Bayes dan

Support Vector Machine Tentang Pembobolan dan Kebocoran Data di Twitter' menghasilkan bahwa nilai akurasi pada naïve bayes lebih unggul 97%. sedangkan metode Support Vector Machine hanya menghasilkan nilai akurasi 93%.

2. TINJAUAN PUSTAKA

2.1. Klasifikasi

Klasifikasi merupakan sebuah metode dengan cara mengelompokkan objek berdasarkan karakteristik yang dimiliki pada objek klasifikasi. Klasifikasi memiliki tujuan untuk dapat mengorganisasikan suatu objek dengan sistem tertentu, sehingga mudah objek ditemukan.

Menurut [8] klasifikasi dilakukan dengan tujuan untuk membentuk urutan atau susunan yang agar memudahkan saat dilakukan identifikasi sehingga menghasilkan urutan yang berguna. Dalam penelitian [9], algoritma *Naïve Bayes* digunakan dalam Memprediksi Penyakit Stroke, dimana pada penelitian ini algoritma *Adaboost* digunakan untuk memaksimalkan nilai akurasi agar mendapatkan hasil akurasi yang optimal. Metode *Naïve Bayes* digunakan untuk mengukur tingkat akurasi yang optimal [10].

Dalam penelitian yang lain, mechine learning dengan metode *Support Vector Machine* di implementasikan untuk mengklasifikasi stunting pada anak dibawah umur lima tahun (BALITA) di Kota Padang [11].

2.2. Naïve Bayes

Naïve bayes merupakan sebuah metode algoritma klasifikasi pada penerapan machine learning yang supervised (memiliki label). Dimana pada klasifikasi Naïve bayes memiliki label kelas "Ya" dan "Tidak", atau "1" dan "0". Algoritma Naïve Bayes merupakan algoritma yang didasari dengan teorema bayes yaitu dengan mengasumsikan bahwa setiap atribut bersifat independen sendiri [12].

Kelebihan naïve bayes menururt [13] [14] dalam penelitiannya, keleihan naïve bayes memiliki cara yang sederhana dan mengasilkan nilai akurasi yang tinggi.

Dalam penelitian lainnya [15] menjelaskkan bahwa naïve bayes merupakan algirutna yang praktis, sederhana, dapat melakukan klasifikasi text dan memiliki akurasi yang tinggi.

Kelebihan lainnya menurut [16] kelebihan naïve bayes bias digunakan untuk mengklasifikasi data kuantitatif dan kualitatif, serta cara perhitungannya tergolong cepat dan efisien.

2.3. *Support Vector Machine*

Support Vector Machine (SVM) adalah salah satu machine learning yang umumnya digunakan untuk menyelesaikan masalah klasfikasi untuk berbagai bidang [17].

Support Vector Machine (SVM) merupakan algoritma machine learning yang digunakan untuk regresi dan klasifikasi. Cara kerja *Support Vector Machine* (SVM) berdasar pada Structural Risk Minimization (SRM) untuk mengolah data menjadi Hyperplane dengan mengklasifikasikan data menjadi dua kelas. Teori *Support Vector Machine* (SVM) diawali dengan pengelompokan kasus-kasus linier yang dapat dipisahkan dengan H*yperplane* dan dibagi menurut kelasnya. Proses *Support Vector Machine* (SVM) diawali dengan masalah klasifikasi dua kelas, sehingga membutuhkan set pelatihan positif dan negative, sehingga *Support Vector Machine* (SVM) akan berusaha mendapatkan Hyperplane (pembatas) yang optimal untuk memisahkan kedua kelas dan memaksimalkan margin kedua kelas tersebut.

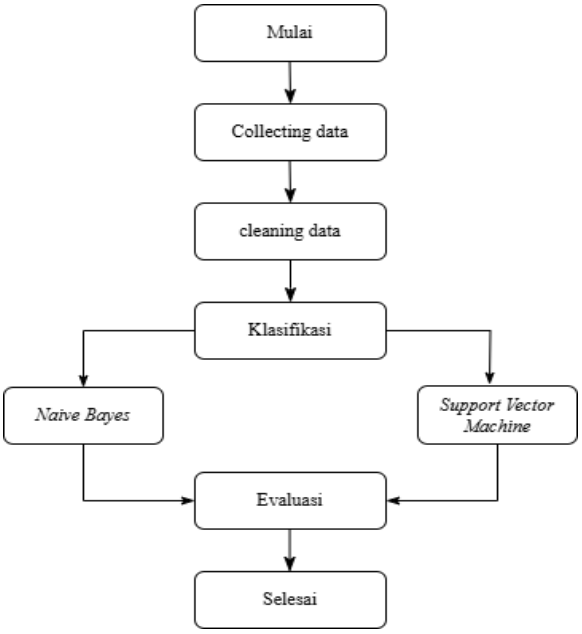
Kelebihan metode *Support Vector Machine* (SVM) menurut [18] adalah proses komputasi cepat, karena *Support Vector Machine* menggunakan teknik *Hyperplane* yang dapat memisahkan dua kelas yang berbeda menjadi satu data. Konsep *Support Vector Machine* (SVM) secara sederhana dapat dijelaskan untuk mencari *Hyperplane* yang paling optimal sebagai pemisah dua kelompok kelas. *Support Vector Machine* (SVM) akan berusaha menemukan fungsi dari pemisah *(Hyperplane)* dengan memaksimalkan margin antar kelas. Dengan proses ini, *Support Vector Machine* (SVM) dapat memastikam kemampuan dalam generalisasi yang tinggi untuk data-data yang akan datang [14].

Kelebihan lainnya menurut [17] pengklasifikasian dalam *Support Vector Machine* (SVM) menawarkan hasil akurasi yang tinggi dan mampu bekerja sangat baik dengan kapasitas dimensi tinggi dan menggunakan subset dari poin *training* sehingga hasilnya menggunakan memori yang sangat sedikit. Dengan dilakukannya perbandingan antara dua algoritma, yaitu Naive Bayes dan Support Vector Machine (SVM), untuk mengklasifikasikan penderita stunting di Kabupaten Deli Serdang. Maka akan terlihat hasil akurasi dari masing-masing metode.

3. METODE PENELITIAN

Penelitian ini berfokus pada metode kuantitatif yang menggunakan data primer dari Dinas Kesehatan Kabupaten Deli Serdang. Data yang akan dikumpulkan meliputi informasi stunting yang terdapat di Kabupaten Deli Serdang. Pengumpulan data dilakukan secara komprehensif untuk memastikan keakuratan dan representasi yang memadai dari kondisi stunting di wilayah tersebut.

3.1. Alur Penelitian



Gambar 1. Alur Penelitian

3.2. Tahap Alur Penelitian

Tahapan ini dimuai dengan collecting data. Pada penelitian digunakan data primer yang berjumlah 422 buah, dataset diambil dari Dinas Kesehatan Kabupaten Deli Serdang. Dataset yang terkumpul disusun dalam format excel, dataset yang diambil dari Dinas Kesehatan Kabupaten Deli Serdang meliputi nama, jenis kelamin, BB Lahir, TB Lahir, Tinggi Badan, BB/U, ZS BB/U, TB/U ZS TB/U, BB/TB, ZS BB/TB dan keterangan tentang stunting dan tidak stunting. Setelah data dikumpulkan, langkah selanjutnya adalah melakukan pre-processing data agar dataset menjadi lebih terstruktur dan siap untuk dianalisis.

Langkah selanjutnya adalah split dataset, Penelitian ini dilakukan pemisahan dataset atau split dataset sebagai bagian dari pemrosesan data. Split dataset adalah teknik pemisahan data menjadi dua bagian: data training dan data testing. Data training digunakan untuk melatih model dengan data yang ada, sementara data testing digunakan untuk menguji performa model pada data yang belum pernah dilihat sebelumnya. Tujuan dari pemisahan ini adalah untuk mengevaluasi kinerja model secara objektif. Penelitian ini menggunakan beberapa kombinasi split dataset untuk menguji model yang dikembangkan. Kombinasi-kombinasi ini dirangkum dalam Tabel 1., yang menampilkan berbagai rasio pemisahan yang digunakan.

Tabel 1. Split Dataset

Split dataset	Training	Testing
60;40	253	169
70;30	295	127
80;20	338	84
90;10	398	24

3.3. Klasifikasi *Naïve Bayes*

Langkah awal dalam klasifikasi *Naïve bayes* adalah melakukan input data/import data, selanjutnya membaca jenis data yang diimport. selanjutnya mengklasifikasikan *Naive Bayes* dengan membagi data, langkah selanjutnya membentuk fungsi klasifikasi dari *Naive Bayes* yang selanjutnya memasukkan data training pada fungsi klasifikasi *Naive Bayes*, dan dari hasil yang didapat maka dilanjutkan dengan menentukan probabilitas hasil prediksi

Selanjutnya membuat tabel confusion atau tabel prediksi. *Confusion Matrix* merupakan matrix yang berisi ketepatan prediksi.

3.4. Klasifikasi *Support Vector Machine*

langkah pertama yang lakukan adalah menginput data. Kemudian mencari data untuk mengetahui variabel independen/features dan nama target. Selanjutnya, dataset dibagi menjadi data training dan data testing. Kemudian membuat model untuk melakukan klasifikasi. Langkah berikutnya adalah melihat Confusion Matrix untuk melihat akurasi dari hasil pengklasifikasian

Evaluasi Model, dimana pada tahapan ini metode akan terus di uji coba sampai dengan batas maksimal.

4. HASIL DAN PEMBAHASAN

4.1. Pra-processing dataset

Proses pre-processing melibatkan serangkaian langkah seperti pembersihan data, transformasi variabel, encoding data label dan penghapusan outlier guna memastikan kualitas data yang dihasilkan. Data yang telah melalui tahap ini diharapkan memiliki kualitas yang optimal dan siap digunakan untuk analisis lebih lanjut dalam penelitian ini. Tujuan preprocessing ini untuk memastikan data yang dianalisis memiliki integritas yang tinggi dan relevansi yang optimal. Dataset pada penelitian ini pada fitur jenis kelamin, laki-laki diubah menjadi angka 1 dengan metode encoding label dan perempuan diubah menjadi angka 0 dengan metode encoding label. Selanjutnya pada fitur BB/U berat badan normal diubah menjadi 2, kurang diubah menjadi 1 dan sangat kurang diubah menjadi 0 dengan teknik label encoding. Begitu juga untuk fitur TB/U dan BB/TB masing-masing

4.2. Split data

Split Dataset 60;40 adalah 60% data training dan 40% data testing, 70;30 adalah 70% data training dan 30% data testing, 80;20 adalah 80% data training dan 20% data testing dan 90;10 adalah 90% data training dan 10% data testing. Semua split dataset akan digunakan untuk membangun model Support Vector Machine (SVM) dan Naive Bayes.

4.3. Pemodelan

4.3.1. Bangun Model Naïve Bayes

a. Inisialisasi Parameter
Parameter diinisialisasi dengan nilai awal:
- 'classes': Kelas unik dalam dataset.
- 'mean': Mean (rata-rata) dari fitur untuk setiap kelas.
- 'var': Varians dari fitur untuk setiap kelas.
- 'priors': Probabilitas prior untuk setiap kelas.

b. Fit Function
Pada bagian ini, dilakukan pelatihan untuk menghitung mean, varians, dan prior untuk setiap kelas.
Persamaan:
- Mean $(\mu_{c,j})$:

$$\mu_{c,j} = \frac{1}{N_c}\sum_{i=1}^{N_c} X_{i,j}$$

- Varians (σ_{cj}^2) :

$$\sigma_{c,j}^2 = \frac{1}{N_\varepsilon}\sum_{i=1}^{N_\varepsilon} \left(X_{i,j} - \mu_c j\right)^2$$

- Priors $(P(c))$:

$$P(c) = \frac{N_c}{N}$$

di mana:
- N_c adalah jumlah sampel dalam kelas c.
- $P(c)$adalah jumlah total sampel.
- $X_{i,j}$ adalah nilai fitur ke- j dari sampel ke- i dalam kelas c.

4.4. Pembangunan model *Support Vrctor Machine*

Tahapan Membangun Model *Support Vector Machine* (SVM) adalah sebagai berikut :

a. Inisialisasi Parameter
Parameter diinisialisasi dengan nilai awal:
- learning_rate (η) : laju pembelajaran
- '1ambda_param (λ) : parameter regulasi
- 'n_iters': jumlah iterasi
- 'w': vektor bobot (diinisialisasi dengan nol)
- 'b': bias (diinisialisasi dengan nol)

b. Fungsi Loss
Loss function yang digunakan adalah kombinasi dari hinge loss dan regularization term. Hinge loss memastikan margin maksimal, sedangkan regularization term mencegah overfitting:

$$L(w,b) = \frac{1}{2}\parallel w \parallel^2 + C\sum_{i=1}^{n} \max\left(0, 1 - y_i(w \cdot x_i + b)\right)$$

Loss function akan dioptimasi menggunakan gradient descent dengan pembaruan parameter.

c. Pembaruan Parameter
Gradient descent digunakan untuk meminimalkan loss function. Update untuk bobot w dan bias b tergantung pada apakah kondisi margin dipenuhi atau tidak.
Jika kondisi margin dipenuhi $(y_i(w \cdot x_i + b) \geq 1)$:

$$w = w - \eta(2\lambda w)$$

Tidak ada pembaruan untuk bias b.
Jika kondisi margin tidak dipenuhi $(y_i(w \cdot x_i + b) < 1)$:

$$w = w - \eta(2\lambda w - y_i x_i)$$
$$b = b - \eta y_i$$

4.5. Evaluasi Model

Selanjutnya setelah modelling selesai, dilakukan uji kinerja model klasifikasi *Support Vector Machine* (SVM) *dan Naive Bayes* dengan menggunakan *Confusion Matrix* dan *classification report* untuk menguji akurasi kinerja model.

4.5.1. *Naïve Bayes*

Tabel 2. Classification Report *Naive Bayes* 60% *Training* 40% *Testing*

Kelas Data	*Precision*	*Recall*	*F-1 score*	*Support Data*
Stunting	97%	92%	94%	192
Tidak *Stunting*	78%	90%	84%	40
Akurasi	92%			

Tabel 2. menunjukkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang sangat buruk dengan akurasi keseluruhan sebesar 24%. Model ini memiliki *precision*, *recall*, dan *F-1 score* sebesar 0% untuk kategori stunting, yang berarti model ini sama sekali tidak mampu mengidentifikasi kasus stunting dengan benar. Untuk kategori tidak stunting, model memiliki *precision* sebesar 24% dan *recall* sebesar 100% dengan *F-1 score* sebesar 38%, yang menunjukkan bahwa model mengidentifikasi semua kasus tidak stunting dengan benar, tetapi juga sering salah mengidentifikasi kasus stunting sebagai tidak stunting. Data dukungan menunjukkan terdapat 129 kasus stunting dan 40 kasus tidak stunting dalam dataset. Dengan demikian, model ini sangat tidak efektif dalam mendeteksi kasus stunting dan tidak dapat diandalkan dalam memprediksi kedua kategori tersebut**.**

Tabel 3. Classification Report *Naive Bayes* 70% *Training* 30% *Testing*

Kelas Data	*Precision*	*Recall*	*F-1 score*	*Support Data*
Stunting	97%	69%	91%	95
Tidak *Stunting*	51%	94%	66%	32
Akurasi	76%			

Tabel 3. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang cukup dengan akurasi keseluruhan sebesar 76%. Model ini memiliki *precision* sebesar 97% dan *recall* sebesar 69% untuk

kategori stunting, serta *F-1 score* sebesar 81%, yang menunjukkan bahwa model sangat tepat dalam mengidentifikasi stunting namun kurang sensitif dalam mendeteksi semua kasus stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 51% dan *recall* sebesar 94% dengan *F-1 score* sebesar 66%, Data dukungan menunjukkan terdapat 95 kasus stunting dan 32 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mengidentifikasi kasus yang benar-benar stunting tetapi perlu peningkatan dalam mendeteksi semua kasus stunting dan dalam mengidentifikasi kasus tidak stunting dengan tepat.

Tabel 4.Classification Report *Naive Bayes* 80% *Training* 20% *Testing*

Kelas Data	***Precision***	***Recall***	***F-1 score***	***Support Data***
Stunting	95%	71%	82%	59
Tidak *Stunting*	59%	91%	72%	26
Akurasi	78%			

Tabel 4. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang cukup dengan akurasi keseluruhan sebesar 78%. Model ini memiliki *precision* sebesar 95% dan *recall* sebesar 71% untuk kategori stunting, serta *F-1 score* sebesar 82%, yang menunjukkan bahwa model sangat tepat dalam mengidentifikasi stunting namun kurang sensitif dalam mendeteksi semua kasus stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 59% dan *recall* sebesar 92% dengan *F-1 score* sebesar 72%,. Data dukungan menunjukkan terdapat 59 kasus stunting dan 26 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mengidentifikasi kasus yang benar-benar stunting tetapi perlu peningkatan dalam mendeteksi semua kasus stunting dan dalam mengidentifikasi kasus tidak stunting dengan tepat.

Tabel 5. Classification Report *Naive Bayes* 90% *Training* 10% *Testing*

Kelas Data	***Precision***	***Recall***	***F-1 score***	***Support Data***
Stunting	96%	84%	90%	31
Tidak *Stunting*	59%	91%	72%	18
Akurasi	86%			

Tabel 5. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang cukup dengan akurasi keseluruhan sebesar 86%. Model ini memiliki *precision* sebesar 96% dan *recall* sebesar 84% untuk kategori stunting, serta *F-1 score* sebesar 90%, yang menunjukkan bahwa model sangat tepat dalam mengidentifikasi stunting dan juga cukup sensitif dalam mendeteksi sebagian besar kasus stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 69% dan *recall* sebesar 92% dengan *F-1 score* sebesar 79%, data dukungan menunjukkan terdapat 31 kasus stunting dan 12 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mengidentifikasi kasus yang benar-benar stunting tetapi perlu peningkatan dalam mengidentifikasi kasus tidak stunting dengan tepat.

4.5.2. *Support Vector Machine*

Tabel 6. Classification Report *SVM* 60% *Training* 40% *Testing*

Kelas Data	***Precision***	***Recall***	***F-1 score***	***Support Data***
Stunting	97%	92%	94%	129
Tidak *Stunting*	78%	90%	84%	40
Akurasi	92%			

Tabel 6. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang cukup baik dengan akurasi keseluruhan sebesar 92%. Model ini memiliki *precision* sebesar 97% dan *recall* sebesar 92% untuk kategori stunting, serta *F-1 score* sebesar 94%, yang artinya menunjukkan hasil yang seimbang antara sensitivitas dan ketepatannya dalam mengidentifikasi stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 78% dan *recall* sebesar 90% dengan *F-1 score* sebesar 84%. Data dukungan menunjukkan terdapat 129 kasus stunting dan 40 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mendeteksi kasus stunting dibandingkan tidak stunting, namun secara keseluruhan tetap andal dalam memprediksi kedua kategori tersebut.

Tabel 7. Classification Report *SVM* 70% *Training* 30% *Testing*

Kelas Data	***Precision***	***Recall***	***F-1 score***	***Support Data***
Stunting	98%	94%	96%	95
Tidak *Stunting*	83%	94%	88%	32
Akurasi	94%			

Tabel 7. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang sangat baik dengan akurasi keseluruhan sebesar 94%. Model ini memiliki *precision* sebesar 98% dan *recall* sebesar 94% untuk kategori stunting, serta *F-1 score* sebesar 96%, yang menunjukkan keseimbangan yang sangat baik antara ketepatan dan sensitivitas dalam mengidentifikasi stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 83% dan *recall* sebesar 94% dengan *F-1 score* sebesar 88%. Data dukungan menunjukkan terdapat 95 kasus stunting dan 32 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mendeteksi kasus

stunting dibandingkan tidak stunting, namun secara keseluruhan tetap andal dalam memprediksi kedua kategori tersebut.

Tabel 8. Classification Report *SVM* 80% *Training* 20% *Testing*

Kelas Data	*Precision*	*Recall*	*F-1 score*	*Support Data*
Stunting	96%	92%	94%	59
Tidak *Stunting*	83%	92%	87%	26
Akurasi	92%			

Tabel 8. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang baik dengan akurasi keseluruhan sebesar 92%. Model ini memiliki *precision* sebesar 96% dan *recall* sebesar 92% untuk kategori stunting, serta *F-1 score* sebesar 94%, yang artinya menujukkan hasil yang seimbangan antara sensitivias dan ketepatannya mengidentifikasi stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 83% dan *recall* sebesar 92% dengan *F-1 score* sebesar 87%. Data dukungan menunjukkan terdapat 59 kasus stunting dan 26 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mendeteksi kasus stunting dibandingkan tidak stunting, namun secara keseluruhan tetap andal dalam memprediksi kedua kategori tersebut.

Tabel 9. Classification Report *SVM* 90% *Training* 10% *Testing*

Kelas Data	*Precision*	*Recall*	*F-1 score*	*Support Data*
Stunting	96%	87%	92%	31
Tidak *Stunting*	73%	92%	81%	12
Akurasi	88%			

Tabel 9. menampilkan hasil evaluasi model klasifikasi dalam mendeteksi kasus stunting menunjukkan performa yang baik dengan akurasi keseluruhan sebesar 88%. Model ini memiliki *precision* sebesar 96% dan *recall* sebesar 87% untuk kategori stunting, serta *F-1 score* sebesar 92%, yang artinya menunjukkan hasil yang seimbangan antara sensitivitas dan ketepatannya mengidentifikasi stunting. Untuk kategori tidak stunting, model memiliki *precision* sebesar 73% dan *recall* sebesar 92% dengan *F-1 score* sebesar 81%. Data dukungan menunjukkan terdapat 31 kasus stunting dan 12 kasus tidak stunting dalam dataset. Dengan demikian, model ini lebih efektif dalam mendeteksi kasus stunting dibandingkan tidak stunting, namun secara keseluruhan tetap andal dalam memprediksi kedua kategori tersebut.

4.6. Perbandingan Evaluasi Model

Tabel 10. Perbandingan evaluasi model

Split dataset	*Akurasi Naïve Bayes*	*Akurasi SVM*
60;40	24%	92%
70;30	76%	94%
80;20	78%	92%
90;10	86%	88%

Penelitian ini menunjukkan bahwa *Support Vector Machine* (SVM) memiliki kinerja yang lebih baik dibandingkan dengan *Naive Bayes* dalam berbagai split dataset (60:40, 70:30, 80:20, 90:10). SVM menunjukkan akurasi yang tinggi dan konsisten, berkisar antara 88% hingga 94%, sedangkan *Naive Bayes* menunjukkan peningkatan akurasi yang signifikan dari 24% pada split 60:40 hingga 86% pada split 90:10. Rata-rata akurasi SVM adalah 91.5% dengan variasi akurasi sebesar 2.87%, menunjukkan kinerja yang lebih stabil. Sebaliknya, *Naive Bayes* memiliki rata-rata akurasi sebesar 66% dengan variasi akurasi sebesar 28.35%, menunjukkan kinerja yang kurang konsisten. Dengan demikian, SVM terbukti lebih unggul secara keseluruhan dalam hal akurasi dan konsistensi, menjadikannya pilihan yang lebih baik untuk tugas klasifikasi dalam penelitian ini

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, metode *Support Vector Machine* (SVM) menunjukkan kinerja yang lebih unggul dan konsisten dibandingkan dengan metode Naïve Bayes dalam mengklasifikasikan penderita stunting di Kabupaten Deli Serdang. *Support Vector Machine* (SVM) memiliki rata-rata akurasi sebesar 91.5%, dengan akurasi yang berkisar antara 88% hingga 94% pada berbagai split dataset (60:40, 70:30, 80:20, dan 90:10). Sebaliknya, Naïve Bayes menunjukkan peningkatan akurasi yang signifikan dari 24% pada split 60:40 menjadi 86% pada split 90:10, dengan rata-rata akurasi sebesar 66%. Variasi akurasi *Support Vector Machine* (SVM) yang lebih rendah (2.87%) dibandingkan dengan Naïve Bayes (28.35%) mengindikasikan bahwa *Support Vector Machine* (SVM) memiliki performa yang lebih stabil. Secara keseluruhan, *Support Vector Machine* (SVM) tidak hanya lebih akurat tetapi juga lebih stabil dibandingkan dengan Naïve Bayes

DAFTAR PUSTAKA

[1] G. Z. Rahmah, R. Kurniasari, F. I. Kesehatan, and U. S. Karawang, "The Influence Of Nutrition Education Media forms On Increasing Mother's Knowledge To Prevent Stunting In Children," *J. Gizi Kesehat.*, vol. 15, no. 1, pp. 131–139, 2023.

[2] M. R. Nugroho, R. N. Sasongko, and M. Kristiawan, "Faktor-faktor yang Mempengaruhi Kejadian Stunting pada Anak Usia Dini di Indonesia," *J. Obs. J. Pendidik. Anak Usia*

Dini, vol. 5, no. 2, pp. 2269–2276, 2021.

[3] M. Zaimy, S. Defit, and G. W. Nurcahyo, "Monte Carlo Prediksi Tingkat Prevalensi Stunting Kabupaten Lima Puluh Kota Menggunakan Metode Monte Carlo," *J. Inf. dan Teknol.*, vol. 3, pp. 245–250, 2021.

[4] Vega Herliansyah, Roswan Latuconsina, and Ashri Dinimaharawati, "Prediksi Stunting Pada Balita Dengan Menggunakan Algoritma Klasifikasi Naïve Bayes," *e-Proceeding Eng.*, vol. 8, no. 5, p. 6642, 2021.

[5] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *J. Ris. Komput.*, vol. 8, no. 6, pp. 219–225, 2021.

[6] H. Dwi Putra, L. Khairani, D. Hastari, P. Studi Sistem Informasi, F. Sains dan Teknologi, and C. Author, "Comparison of Naive Bayes Classifier and Support Vector Machine Algorithms for Classifying Student Mental Health Data Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine," *SENTIMAS Semin. Nas. Penelit. dan Pengabdi. Masy.*, pp. 120–125, 2023.

[7] 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.

[8] F. N. Josua, "Klasifikasi Penderita Stunting dengan Metode Support Vector Machine (Studi kasus: Lima Puskesmas di Kota Bandar Lampung)," *Univ. Lampung*, vol. 1, p. 1, 2021.

[9] A. Byna and M. Basit, "PENERAPAN METODE ADABOOST UNTUK MENGOPTIMASI PREDIKSI PENYAKIT STROKE DENGAN ALGORITMA NAÏVE BAYES," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 3, pp. 407–411, Nov. 2020.

[10] E. R. Arumi, S. A. Subrata, and A. Rahmawati, "JURNAL RESTI Implementation of Naïve bayes Method for Predictor Prevalence Level for Malnutrition Toddlers in Magelang City," *J. Rekayasa Sist. dan Teknol. Inf.*, vol. 5, no. 158, pp. 201–207, 2023.

[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," *J. Math. Its Apl.*, vol. 16, no. 3, pp. 771–778, 2022.

[12] D. M. Muafa and I. Lizda, "Pengembangan Aplikasi Berbasis Web dengan Rshiny untuk Data Klasifikasi Menggunakan Metode Naive Bayes," *Nuviversitas Islam Indones.*, 2020.

[13] A. P. Catur, A. A. Amalia, and S. Agus, "Klasifikasi Akun Buzzer Pada Twitter Menggunakan Algoritma Naive Bayes," *J. Ilm. Tek. Inform. dan Komun.*, vol. 3, no. 1, pp. 1–12, 2023.

[14] A. O. Pusphita, W. Yuciana, and I. Dwi, "Penerapan Metode KlasifikasiSupport Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*, vol. 3, no. 2339–2541, pp. 811–820, 2014.

[15] A. R. Khoirul, Berlilana, and A. Primandani, "Perbandingan Metode Support Vector Machine dan Decission Tree untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *J. Inf. Syst. Manag.*, vol. 2, no. 2, 2021.

[16] N. Maulidah, R. Supriyadi, D. Y. Utami, F. N. Hasan, A. Fauzi, and A. Christian, "Prediksi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine dan Naive Bayes," *Indones. J. Softw. Eng.*, vol. 7, no. 1, pp. 63–68, 2021.

[17] A. Weni, T. F. Muh., and R. Bayu, "Implementasi Metode Support Vector Machine (SVM) Untuk Klasifikasi Rumah Layak Huni (Studi Kasus : Desa Kidal Kecamatan Tumpang Kabupaten Malang)," *J. Pengemb. Teknol. Inf. dan ILmu Komput.*, vol. 2, no. 10, pp. 3366–3372, 2018.