

Analisis Perbandingan Klasifikasi Balita Beresiko Stunting Menggunakan Metode Support Vector Machine dan Decision Tree

Muhamad Amirudin¹, Alz Danny Wowor*²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Kristen Satya Wacana
Jl. Diponegoro No. 52-60, Salatiga

¹672021602@student.uksw.eu

²alzdanny.wowor@uksw.edu

Received on 27-10-2023, revised on 19-11-2023, accepted on 10-12-2023

ABSTRACT

Permasalahan status gizi penderita stunting dilihat dari balita berusia dibawah lima tahun dan perkembangannya sejak awal kehidupan. Fase-fase awal kehidupan merupakan hal yang terpenting dalam proses pertumbuhan dan perkembangan anak. Permasalahan stunting di Indonesia menjadi target khusus pemerintah dalam penurunan angka kasus yang terjadi, maka dari itu sebagai langkah wujud nyata akan hal itu, penelitian dilakukan guna mengetahui pengoptimalan perbandingan sebuah metode dalam klasifikasi balita beresiko stunting. Penelitian ini dilakukan dengan membandingkan kedua metode Support Vector Machine dan Decision Tree untuk mengklasifikasi bayi beresiko stunting untuk mendapatkan hasil akurasi yang tepat serta efisien untuk mendapatkan rekomendasi dalam penentuannya. Hasil dari Model SVM terbaik memiliki akurasi sekitar 83% pada dataset pengujian, sedangkan model Decision Tree memiliki akurasi sekitar 78%. Model SVM juga memiliki recall yang baik yaitu 100% pada target '1' atau stunting, yang berarti model ini efektif dalam menentukan balita beresiko stunting. Dari keunggulan Model, Model SVM dapat disimpulkan lebih cocok jika keseimbangan antara precision dan recall sangat penting dalam konteks aplikasi tertentu. Dalam penelitian ini, SVM adalah pilihan yang dapat diandalkan untuk menganalisis risiko stunting pada balita.

Keywords: : Support Vector Machine, Decision Tree, Stunting, Klasifikasi Stunting

This is an open access article under the [CC BY-SA](#) license.



Penulis Korespondensi:

Alz Danny Wowor

Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Kristen Satya Wacana Jl.

Diponegoro No. 52-60, Salatiga

Email: alzdanny.wowor@uksw.edu

I. PENDAHULUAN

Perkembangan dan pertumbuhan anak pada usia dini sangat berpengaruh dalam kemampuan anak di masa yang akan datang. Apabila terdapat penyimpangan pertumbuhan anak tidak diketahui sejak dini dengan baik, makan akan berpengaruh terhadap pertumbuhan dan perkembangan anak selanjutnya. Namun pada kenyataannya setiap anak memiliki perbedaan dalam perkembangan dan pertumbuhannya, gagalnya perkembangan maupun pertumbuhan tersebut adalah Stunting [1]. Menurut *World Health Organization* (WHO) Stunting merupakan masalah gangguan perkembangan dan pertumbuhan anak yang disebabkan oleh gizi buruk, tinggi badan yang tidak sesuai dengan median pertumbuhan yang ditetapkan oleh WHO. Adapun bayi dikatakan stunting apabila standar pertumbuhan yang dimiliki kurang dari -2

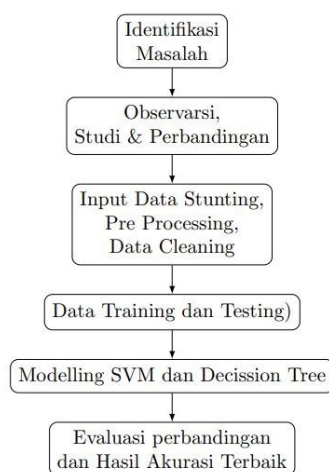
standar deviasi [2]. Permasalahan status gizi penderita stunting dilihat dari bayi berusia dibawah lima tahun dan perkembangannya sejak awal kehidupan. Fase-fase awal kehidupan merupakan hal yang terpenting dalam proses pertumbuhan dan perkembangan anak. Penyebab stunting memiliki banyak faktor seperti kurangnya asupan makanan bergizi, pola asuh orang tua, keturunan pendek, pendapatan ASI eksklusif, sanitasi, dan kemiskinan [3]. Kini kasus stunting menjadi prioritas penanganan pemerintah dalam bidang kesehatan balita periode tahun 2019-2024. Berdasarkan hasil survei status gizi indonesia (SSGI) Kementerian Kesehatan Republik Indonesia, angka prevalensi stunting per januari tahun 2023 kasus stunting di Indonesia turun dari 24,4 % di tahun 2021 menjadi 21,6 % di tahun 2022. Walaupun mengalami penurunan, 2 namun perlu percepatan untuk mencapai target program pemerintah sebesar 14% di tahun 2024.

Kemajuan teknologi informasi menghasilkan metode-metode dalam data science yang mampu memberikan kemudahan dalam mengambil suatu keputusan dalam berbagai bidang, salah satunya adalah metode *Support Vector Machine* (SVM). Metode SVM merupakan algoritma yang menggunakan ruang hipotesis berupa fungsi-fungsi linear berdasarkan teori optimalisasi. Metode ini dikategorikan sebagai teknik yang baru dibanding dengan yang lainnya. Sementara itu, "*Decision Tree*" (Pohon Keputusan) adalah sebuah metode pembelajaran yang membangun model prediksi dalam bentuk struktur pohon. Pohon tersebut terdiri dari simpul (*node*) yang mewakili keputusan berdasarkan pada nilai-nilai atribut, serta cabang-cabang yang mewakili hasil dari keputusan tersebut. *Decision Tree* juga telah digunakan secara luas dalam klasifikasi bayi berisiko stunting karena kemampuannya untuk memberikan pemahaman yang intuitif tentang faktor-faktor yang berkontribusi terhadap risiko stunting.

Penelitian ini bertujuan untuk melakukan analisis perbandingan antara metode "*Support Vector Machine*" dan "*Decision Tree*" dalam klasifikasi balita berisiko stunting berdasarkan pada data atribut dan kriteria yang relevan. Dengan membandingkan kedua metode ini, diharapkan dapat diperoleh pemahaman yang lebih baik tentang mana metode yang lebih cocok dan efektif untuk digunakan dalam pengenalan dini risiko stunting pada balita. Melalui penelitian ini, diharapkan bahwa hasilnya dapat memberikan kontribusi positif dalam upaya pencegahan dan penanganan stunting pada bayi, serta menjadi dasar untuk pengembangan metode klasifikasi yang lebih canggih dan akurat di masa depan.

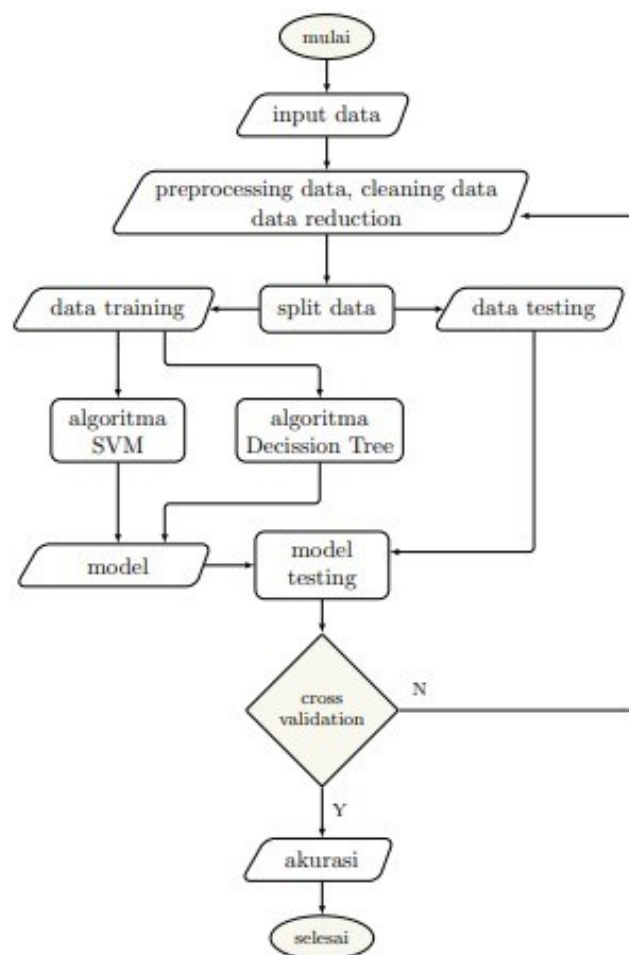
II. METODOLOGI PENELITIAN

Adapun tujuan penelitian ini untuk memprediksi dan mengklasifikasi bayi berisiko stunting dari dataset yang telah tersedia. Penelitian ini menggunakan dua metode untuk membandingkan hasil akurasi terbaik dalam proses klasifikasi bayi berisiko stunting. Ada beberapa langkah yang dilakukan dalam menyelesaikan penelitian ini, dimana langkah tersebut terdapat pada Gambar 1.



Gambar 1. Proses Penelitian

Pada penelitian memiliki beberapa tahap atau langkah, yang pertama adalah identifikasi masalah. Identifikasi masalah berupa melihat serta mengidentifikasi angka stunting yang jarang diprediksi menggunakan metode algoritma yang ada dalam lingkup teknologi informasi serta melakukan perancangan kerangka pada awal penelitian yang dilakukan. Tahapan kedua yaitu observasi, studi literatur dan perbandingan. Adapun observasi dilakukan melalui kajian pustaka dari penelitian terdahulu mengenai stunting, faktor penyebab, dan prediksi angka stunting dalam berbagai metode serta memilih metode yang paling efektif. Tahapan ketiga adalah input data, pre prosesi data dan data cleaning. Pada tahap ini peneliti telah mendapatkan beberapa data angka stunting yang memiliki beberapa atribut penyebab balita beresiko stunting, melalui tahap cleaning dan pre processing data ini menghasilkan data yang siap dianalisis dengan metode yang dipilih. Tahap berikutnya yaitu membagi data yang siap dianalisis menjadi dua bagian yaitu data training atau data latih dan data pengujian atau data testing. Setelah terbagi, data akan diolah dan dianalisis menggunakan metode SVM dan *Decission Tree* guna menghasilkan tahap terakhir, yaitu mengevaluasi hasil perbandingan antar metode untuk menentukan hasil akurasi terbaik dalam memprediksi balita beresiko stunting. Alur kerja pada penelitian ini secara umum dijelaskan pada 4. Penelitian ini diawali dengan pencarian data stunting untuk eksperimen memprediksi angka stunting. Adapun data tersebut berjumlah 10.000 data dengan 7 variabel atau atribut yang diukur. Tahap berikutnya yaitu *preprocessing* data yang berisi data cleaning dan data reduction untuk menjadikan *data set*. Setelah itu, *data set* dibagi menjadi dua bagian yakni data latih atau training dan data uji atau *data testing*. Dari kedua data tersebut akan ditentukan ratio. presentase atau perbandingan dalam penentuan data yang termasuk *data training* dan *data testing*.



Gambar 2. Skema Umum Proses Penelitian

Apabila *data training* dan *dataset* telah ditentukan rasio jumlahnya masing-masing, maka akan menuju proses analisis dengan modelling menggunakan metode SVM dan *Decission Tree*. Kedua hasil *confuse*

matrix yang didapatkan dari masing-masing model akan dibandingkan untuk mencari nilai hasil yang terbaik, hasil terbaik yang dimaksud adalah memiliki nilai akurasi presentase yang tinggi untuk memprediksi angka stunting berdasarkan eksperimen yang sudah dihasilkan.

III. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini merupakan data yang berasal dari kaggle.com dengan jumlah data sebanyak 10.000 data dengan melalui beberapa eksperimen dalam proses split data menjadi *data training* dan *data test* menggunakan metode SVM dan *Decission Tree*. Adapun tabel atribut data digunakan dalam Tabel 1.

TABEL 1. ATRIBUT DATA STUNTING

No	Atribut	Tipe	Deskripsi
1	Age	integer	umur
2	Birth Weight	float	berat lahir
3	Birth Length	float	panjang lahir
4	Body Weight	float	berat sekarang
5	Body Length	float	panjang sekarang
6	Breastfeeding	float	menyusui
7	Stunting	integer	stunting

Dapat dilihat pada tabel 1 diatas jumlah atribut yang digunakan sebanyak 7 atribut. Berikut Tabel 2 merupakan contoh data yang digunakan dalam penelitian ini :

TABEL 2. CONTOH DATA STUNTING

No	Age	BirthW	BirthL	BdyW	BdyL	BrsFd	Stunting
1	17	3.0	49	10.0	72.2	0	0
2	16	2.9	49	8.5	72.2	0	1
3	15	3.1	49	10.5	49.0	0	1
4	17	2.8	49	8.0	63.0	0	1
5	16	2.8	49	8.5	65.0	0	1
.
.
9998	16	2.8	49	7.7	49	0	0
9999	14	2.8	49	10	69	0	1
10000	10	3	49	7.7	80	0	1

Python adalah bahasa pemrograman yang biasa digunakan dalam proses analisis data, termasuk dalam implementasi dan penggunaan Support Vector Machines (SVM) dan *Decission Tree* untuk pemodelan dan analisis data. SVM dan *Decission Tree* adalah algoritma *machine learning* yang digunakan klasifikasi dan regresi. Eksperimen yang dilakukan dalam penelitian ini berupa analisis perbandingan dari kedua metode tersebut dengan mencari hasil akurasi terbaik dalam penentuan balita beresiko stunting dengan atribut yang tersedia. Adapun proses yang sama dalam analisa kedua metode tersebut sebagai berikut :

1. Proses *Split Data* yang menggunakan fungsi sama yaitu *train test split* dalam *sklearn.model selection* dengan membagi dua data yakni data latih atau *data test* dan data uji model atau *data training*. Hal ini memiliki fungsi penting dalam pengembangan model kedua metode tersebut, karena dapat membantu mengukur kinerja pada data yang belum pernah dilihat sebelumnya.

2. Penggunaan Model pada *python* Kode from *sklearn.svm import SVC* digunakan untuk mengimpor kelas SVM (*Support Vector Machine*) dari pustaka *scikit-learn (sklearn)*. Setelah mengimpor SVC, kita dapat membuat instance dari model SVM untuk melakukan pelatihan dan prediksi pada data klasifikasi. Begitu pula pada metode *DecisionTreeClassifier*: Ini adalah kelas dari *scikit-learn* yang digunakan untuk menginisialisasi model *Decision Tree* untuk tugas klasifikasi. Pada proses model di *Decision Tree* terdapat fungsi *Random State* dan pengisian kriteria "gini" serta *max depth=n* yang berarti untuk mengatur generator random acak untuk menghasilkan pemisahan data dalam pemrosesan hasil eksperimen yang berbeda, selain itu penggunaan kriteria berfungsi sebagai pengatur seberapa baik tingkat kedalaman *entropy* yang baik ketika memisahkan kelas target.

3. Hasil Akurasi Hasil akurasi dari eksperimen kedua metode tersebut merupakan laporan evaluasi kinerja performa klasifikasi seperti *precision* sebagai pengukur sejauh mana prediksi label atau target '0' dan '1' dalam bentuk presentase, *Recall* untuk mengukur sejauh mana model yang berhasil menemukan *instance* sama, *F1 Score* yang berupa rata-rata dari nilai *precision* dan *Recall*, *support* berupa jumlah *instance* pada masing-masing target dan *accuracy* yang merupakan hasil akhir sebagai pengukuran sejauh mana model tersebut dalam proses klasifikasi.

Skenario dalam proses eksperimen adalah melihat kemampuan metode dalam melakukan prediksi dengan penggunaan *test size* yaitu *test set* dan data *train test* yang berbeda. Digunakan lima eksperimen untuk setiap metode, secara berturut-turut *test size (x)* untuk proses eksperimen adalah $x \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. Banyak data yang digunakan untuk *train test (tt)* diberikan pada Persamaan 3

$$t_i(x) = x \times 100\% \quad (3)$$

Sebaliknya banyak ukuran data yang digunakan untuk setiap *test set (ts)* berdasarkan Persamaan 4. Sehingga akibat dari kedua persamaan ini akan diperoleh $tt + ts = 1$, dari relasi ini benar secara persentase untuk sebuah himpunan data.

$$t_s(x) = 100\% - t_i \quad (4)$$

Berikut merupakan hasil eksperimen dari metode SVM yang diberikan dalam tabel, dengan memperhatikan nilai target 0 atau 1 berdasarkan acuan nilai *precision*, *recall*, *f1-score* dan *support*. Nilai akurasi dari setiap metode merupakan indikator yang penting dalam riset ini, karena digunakan untuk menentukan tingkat resiko stunting.

TABEL 3. HASIL EKSPERIMEN SVM (TEST SIZE = 0.1)

Target	Precision	Rec all	F1-Score	Support
0	0.00	0.00	0.00	207
1	0.79	0.91	0.88	793
Accuracy: 79%				

Hasil eksperimen pertama untuk metode SVM diberikan pada Tabel 3, dimana ratio split data 0.1 yang berarti *test size* = 0.1. Sehingga dari data akan diambil untuk menjadi *test set (ts = 10%)* dan sementara 90% akan menjadi *train set (tt = 90%)*. Akurasi diperoleh adalah 0.79, tentunya menginformasikan bahwa metode SVM memiliki tingkat ketepatan 79 %, dalam melakukan klasifikasi bayi beresiko stunting.

Nilai *precision* yang dihasilkan untuk target 0 dan 1 masing-masing adalah 0.00% dan 0.79%, hal ini berarti nilai *precision* didapatkan dari rasio berupa jumlah prediksi positif yang benar (TP) dibagi dengan total prediksi positif (positif benar + positif salah). Dalam eksperimen ini, untuk target 0 bernilai 0.00% yang berarti tingkat prediksi positif yang dihasilkan 0% atau sama dengan sangat rendah untuk mengukur prediksi balita yang tidak beresiko stunting. Sedangkan pada target 1 bernilai 0.79% yang berarti tingkat prediksi positif yang benar bernilai 79% .

Recall yang dihasilkan dari eksperimen pertama masing-masing adalah 0.00% dan 0.91%, dimana *recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi, *recall* adalah rasio dari jumlah prediksi positif yang benar (*True Positive*) dibagi oleh total data aktual yang seharusnya positif (*True Positive + False Negative*). Dalam eksperimen pertama ini, untuk *recall* target 0 adalah 0.49, yang berarti model berhasil mengidentifikasi 49% dari keseluruhan data aktual yang seharusnya positif, sementara untuk *recall* target 1 adalah 0.91, yang berarti model berhasil mengidentifikasi 91% dari data aktual yang seharusnya positif.

F1-Score yang dihasilkan dari eksperimen pertama masing-masing adalah 0.00% dan 0.88%, *F1-score* adalah rata-rata hasil dari *precision* dan *recall*. Hal ini digunakan untuk mengukur keseimbangan antara *precision* dan *recall*. Dalam eksperimen pertama ini, untuk target 0, *F1-score* adalah 0%, sementara untuk kelas 1, *F1-score* adalah 89%. *F1score* memberikan gambaran keseluruhan tentang kualitas prediksi model, menggabungkan informasi tentang akurasi *precision* dan kemampuan model untuk menemukan semua kasus positif *recall*. *Support* merupakan jumlah data aktual yang termasuk dalam setiap kelas. Dalam kasus ini, ada 207 data untuk target 0 dan 793 data untuk target 1.

TABEL 4. HASIL EKSPERIMEN SVM (TEST SIZE = 0.2)

Target	Precision	Recall	F1-Score	Support
0	0.60	0.49	0.54	422
1	0.87	0.91	0.89	1578
Accuracy: 82%				

Hasil eksperimen kedua untuk metode SVM membuktikan ada kenaikan pada Tabel 4, dimana *ratio split data* 0.2 yang berarti *test size* = 0.2 bahwa 20 % dari data akan diambil untuk menjadi *test set*, sementara 80% akan menjadi *train set*. Dimana, nilai *precision* yang dihasilkan untuk target 0 dan 1 masing-masing adalah 0.60% dan 0.87% untuk memprediksi positif yang benar. *Recall* yang menghasilkan keberhasilan model dalam menemukan sebuah informasi kembali bernilai masing-masing adalah 0.49% dan 0.91%, dan *F1-Score* yang berisi rata-rata *precision* dan *recall* masing-masing adalah 0.54% dan 0.89% serta *Support* jumlah data aktual terdapat 422 data untuk target 0 dan 1578 untuk target 1. Akurasi keseluruhan diperoleh masih sama dengan eksperimen pertama yaitu sebesar 0.79, tentunya menginformasikan bahwa metode SVM memiliki tingkat ketepatan 79%, dalam melakukan klasifikasi bayi beresiko stunting.

TABEL 5. HASIL EKSPERIMEN SVM (TEST SIZE = 0.3)

Target	Precision	Recall	F1-Score	Support
0	0.62	0.48	0.54	629
1	0.87	0.92	0.89	2371
Accuracy: 83%				

Akurasi keseluruhan pada tabel 5 adalah 0.83, berarti eksperimen ini memiliki nilai akurasi baik dari eksperimen sebelumnya. Dimana model tingkat akurasi nya adalah 83 % pada *dataset* pengujian. Hasil eksperimen ketiga ini menggunakan *ratio split data* 0.3 yang berarti *test size*=0.3 bahwa 30% dari data akan diambil untuk menjadi *test set*, sementara 70% akan menjadi *train set*. Hal menginformasikan bahwa metode SVM memiliki tingkat ketepatan 83 %, dalam melakukan klasifikasi bayi beresiko stunting. Adapun

nilai akurasi tersebut didapatkan dari nilai *precision* yang dihasilkan untuk target 0 dan 1 masing-masing adalah 0.62% dan 0.87% untuk memprediksi positif yang benar. *Recall* yang menghasilkan keberhasilan model dalam menemukan sebuah informasi kembali bernilai masing-masing adalah 0.48% dan 0.92%, dan *F1-Score* yang berisi rata-rata *precision* dan *recall* masing-masing adalah 0.54% dan 0.89% serta *Support* jumlah data aktual terdapat 629 data untuk target 0 dan 2371 untuk target 1.

TABEL 6. HASIL EKSPERIMEN SVM (TEST SIZE = 0.4)

Target	Precision	Recall	F1-Score	Support
0	0.00	0.00	0.00	816
1	0.80	1.00	0.89	3184

Accuracy: 80%

Hasil eksperimen keempat mengalami penurunan untuk metode SVM yang diberikan pada Tabel 6, apabila *ratio split data* 0.4 yang berarti *test size* = 0.4. Sehingga dari data akan diambil untuk menjadi *test set* (ts = 40%) dan sementara 60% akan menjadi *train set* (tt = 60%). Dimana nilai dari *precision* yang dihasilkan untuk target 0 dan 1

masing-masing adalah 0.00% dan 0.80% untuk memprediksi positif yang benar. *Recall* yang menghasilkan keberhasilan model dalam menemukan sebuah informasi kembali bernilai masing-masing adalah 0.00% dan 1,0%, dan *F1-Score* yang berisi rata-rata *precision* dan *recall* masing-masing adalah 0% dan 0.89% serta *Support* jumlah data aktual terdapat 816 data untuk target 0 dan 3184 untuk target 1. Akurasi keseluruhan diperoleh masih sama dengan eksperimen pertama yaitu sebesar 0.80, tentunya menginformasikan bahwa metode SVM memiliki tingkat ketepatan 80 %, dalam melakukan klasifikasi bayi beresiko stunting.

TABEL 7. HASIL EKSPERIMEN SVM (TEST SIZE = 0.5)

Target	Precision	Recall	F1-Score	Support
0	0.00	0.00	0.00	1005
1	0.80	1.00	0.89	3995

Accuracy: 80%

Hasil eksperimen terakhir pada metode SVM yang diberikan pada Tabel 7 dengan *ratio split data* 0.5 yang berarti *test size* = 0.5 bahwa 40 % dari data akan diambil untuk menjadi *test set*, sementara 60% akan menjadi *train set* mendapat nilai akurasi keseluruhan sebesar 80% dalam melakukan klasifikasi bayi beresiko stunting. Dimana nilai *precision* yang dihasilkan untuk target 0 dan 1 masing-masing adalah 0.00% dan 0.80% untuk memprediksi positif yang benar. *Recall* yang menghasilkan keberhasilan model dalam menemukan sebuah informasi kembali bernilai masing-masing adalah 0.00% dan 1.00%, dan *F1-Score* yang berisi rata-rata *precision* dan *recall* masing-masing adalah 0.00% dan 0.89% serta *Support* jumlah data aktual terdapat 1005 data untuk target 0 dan 1578 untuk target 1. Berikut merupakan hasil eksperimen dari metode *Decision Tree*.

TABEL 8. HASIL EKSPERIMEN DECISION TREE (TEST SIZE = 0.1)

Target	Precision	Recall	F1-Score	Support
0	0.46	0.52	0.49	207
1	0.87	0.84	0.85	793

**Accuracy:
77%**

Hasil eksperimen pertama dengan metode DT memiliki akurasi yang lebih rendah dibandingkan dengan 5 eksperimen metode SVM, dimana dalam pemrosesannya menambahkan generator acak sebesar

40 dan *ratio split data* 0.1 yang berarti *test size*=0.1 bahwa 10 % dari data akan diambil untuk menjadi *test set* , sementara 90% akan menjadi *train set* mendapat nilai akurasi keseluruhan sebesar 77% dalam melakukan klasifikasi bayi beresiko stunting. Nilai *precision* yang dihasilkan untuk target 0 dan 1 masing-masing adalah 0.46% dan 0.87% untuk memprediksi positif yang benar. *Recall* yang menghasilkan keberhasilan model dalam menemukan sebuah informasi kembali bernilai masing-masing adalah 0.52% dan 0.84%, dan *F1-Score* yang berisi rata-rata *precision* dan *recall* masing-masing adalah 0.49% dan 0.85% serta *Support* jumlah data aktual terdapat 207 data untuk target 0 dan 793 untuk target 1.

TABEL 9. HASIL EKSPERIMEN DECISSION TREE (TESTSIZE = 0.2)

Target	Precision	Recall	F1-Score	Support
0	0.46	0.52	0.49	407
1	0.87	0.85	0.86	1593

Accuracy: 78%

Hasil eksperimen kedua pada tabel 9 adalah 0.78, berarti hasil kinerja akurasi model secara keseluruhan dalam memprediksi balita beresiko stunting sebesar 78 % pada dataset pengujian. Hasil ini memiliki sedikit peningkatan sebesar 1% dari eksperimen dimana yang menjadi pembeda ialah menggunakan *ratio split data* 0.2 yang berarti *test size* = 0.2 bahwa 80 % dari data akan diambil untuk menjadi *test set*, sementara 20% akan menjadi *train set* serta *random state* generator acak sebesar 40. Secara hasil dari nilai *precision*, *recall*, *F1-Score* dan *support* memiliki nilai yang mirip dengan eksperimen pertama.

TABEL 10. HASIL EKSPERIMEN DECISSION TREE (TEST SIZE = 0.3)

Target	Precision	Recall	F1-Score	Support
0	0.48	0.53	0.51	625
1	0.87	0.85	0.86	2375

Accuracy: 78%

Eksperimen ketiga memiliki nilai akurasi yang sama dengan eksperimen kedua, Akurasi pada tabel 10 adalah 0.78, berarti hasil kinerja akurasi model sebesar 78 % pada dataset pengujian. Hasil eksperimen ini menggunakan *ratio split data* 0.3 yang berarti *test size* = 0.3 bahwa 70 % dari data akan diambil untuk menjadi *test set*, sementara 30% akan menjadi *train set* serta *random state* generator acak sebesar 40. Secara hasil dari nilai *precision*, *recall*, *F1-Score* dan *support* memiliki nilai yang mirip dengan eksperimen pertama dan kedua.

TABEL 11. HASIL EKSPERIMEN DECISSION TREE (TEST SIZE = 0.4)

Target	Precision	Recall	F1-Score	Support
0	0.48	0.53	0.50	837
1	0.87	0.85	0.86	3163

Accuracy: 78%

Akurasi pada tabel 11 adalah 0.78, berarti hasil kinerja akurasi model sebesar 78 % pada dataset pengujian dalam mengklasifikasi balita beresiko stunting. Hasil tersebut menunjukkan nilai akurasi secara keseluruhan memiliki nilai yang sama. Hasil eksperimen ini menggunakan *ratio split data* 0.4 yang berarti *test size* = 0.4 bahwa 60 % dari data akan diambil untuk menjadi *test set*, sementara 40% akan menjadi *train set* serta *random state* generator acak sebesar 40. Secara hasil dari nilai *precision*, *recall*, *F1-Score* dan *support* memiliki nilai yang mirip dengan eksperimen pertama, kedua dan ketiga.

TABEL 12. HASIL EKSPERIMEN DECISION TREE (TEST SIZE = 0.5)

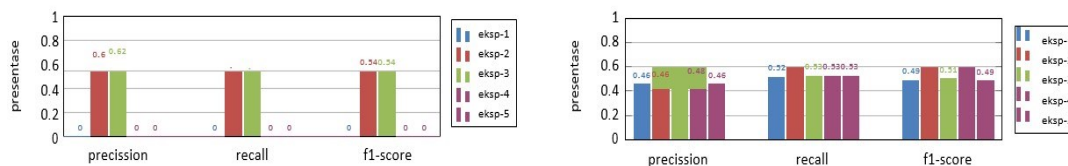
Target	Precision	Recall	F1-Score	Support
0	0.46	0.53	0.49	1019
1	0.87	0.84	0.86	3981

Accuracy: 78%

Berdasarkan empat eksperimen yang telah dilakukan, nilai akurasi keseluruhan pada tabel 12 adalah 0.78, berarti hasil kinerja akurasi model sebesar 78 % pada dataset pengujian dalam mengklasifikasi balita beresiko stunting . Hasil eksperimen ini menggunakan *ratio split data* 0.5 yang berarti *test size* = 0.5 bahwa 50 % dari data akan diambil untuk menjadi *test set*, sementara 50% akan menjadi *train set* serta *random state* generator acak sebesar 40. Nilai *precision*, *recall*, *F1-Score* dan *support* juga menunjukkan kemiripan dari empat eksperimen yang telah dilakukan.

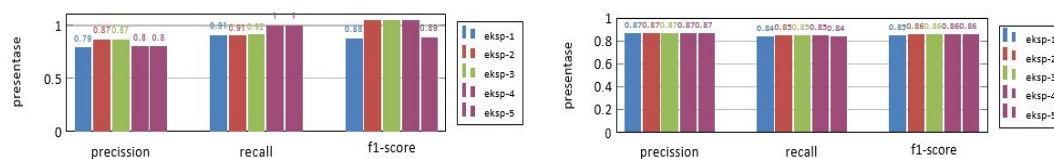
1) Perbandingan Confuse Matrix

Perbandingan hasil untuk lima eksperimen (eksp) untuk setiap metode diberikan pada Gambar, Berikut merupakan hasil 5 eksperimen presentase *precision*, *recall* dan *f1 score* metode SVM dengan label 0 atau tidak beresiko stunting dan label 1 atau beresiko stunting.



Gambar 3. Metode SVM (target 0) dan Metode DT (target 0)

Dari gambar 3 dapat diartikan bahwa nilai rata-rata *precision*, *recall* dan *f1 score* berturut-turut pada metode *Support Vector Machine* dengan label 0 atau tidak beresiko stunting yakni 0.24%, 0.194% dan 0.216% serta untuk metode *Decision Tree* berturut-turut yakni 0.46%, 0.52% dan 0.43. Dari hasil tersebut bisa ditarik kesimpulan bawasannya nilai kedua metode tersebut masih rendah dalam mengklasifikasi balita beresiko stunting dari segi indikator *confusion matrix* dan juga variabel-variabel atau atribut yang mempengaruhi balita beresiko stunting. Maka dari itu penelitian ini berfokus pada nilai target '1' yakni mengklasifikasi balita beresiko stunting



Gambar 4. Metode SVM (target 1) dan Metode DT (target 1)

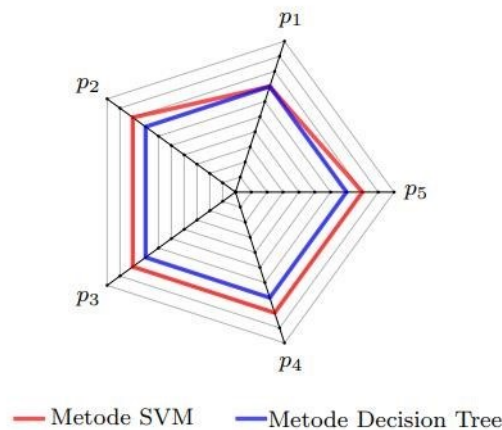
Dari gambar 4 dapat diartikan bahwa nilai rata-rata *precision*, *recall* dan *f1 score* berturut-turut pada metode *Support Vector Machine* dan *Decision Tree* dengan label 1 atau beresiko stunting memiliki nilai yang tinggi antar keduanya, tetapi berdasarkan dengan hasil jumlah 10 eksperimen yang dibuktikan dengan penggunaan data train dan data test, generator acak yang berbeda-beda disimpulkan bawasannya metode SVM lebih unggul dari nilai akurasi keseluruhan dan efektif dari nilai-nilai indikator *confusion matrix* yang ada.

Dalam perbandingan antara metode SVM dan *Decision Tree* untuk klasifikasi label 0 (tidak beresiko stunting) dan label 1 (beresiko stunting), nilai rata-rata *precision* lebih baik untuk label 1, artinya lebih

baik dalam mengidentifikasi yang beresiko stunting. SVM memiliki *recall* lebih tinggi untuk label 1, serta *F1 score* memberikan gambaran keseluruhan kinerja. Model dengan *F1 score* yang lebih tinggi maka efektif dalam situasi di mana kesalahan prediksi keduanya memiliki konsekuensi yang signifikan untuk memprediksi balita beresiko stunting

2) Perbandingan Hasil Akurasi

Gambar 9 merupakan perbandingan akurasi dari dari kedua metode. Berdasarkan hasil yang telah diberikan pada bagian sebelumnya, penelitian ini melakukan lima proses eksperimen untuk dapat melihat seberapa baik kedua metode dalam melakukan klasifikasi penentuan balita beresiko stunting.



Gambar 5. Perbandingan Hasil

Hasil perbandingan ditunjukkan dengan menggunakan diagram radar, dimana setiap p_i untuk $i = 1, 2, 3, 4, 5$ adalah banyak percobaan atau eksperimen yang dilakukan. Nilai akurasi digunakan secara spesifik untuk membandingkan kedua metode, dan nilai untuk setiap p_i adalah persentase dari nilai akurasi, dengan interval dari 0 sampai 10 rusuk menunjukkan 0% sampai 100%. Hasil yang nampak pada grafik, metode SVM lebih baik dari dengan metode DT. Percobaan dilakukan lima kali, karena nilai akurasi untuk metode DT tidak mengalami fluktuatif sejak percobaan ke-2 sampai percobaan ke-5 yaitu 78%, dan metode SVM percobaan ke-4 dan ke-5 mempunyai nilai yang sama juga yaitu 80%.

Pada percobaan pertama nilai akurasi kedua metode mempunyai selisih 2% dimana metode SVM lebih baik. Sedangkan pada percobaan kedua, metode SVM lebih unggul 4%. Selanjutnya untuk perbedaan nilai akurasi dari percobaan ke-3, ke-4, dan ke-5 secara berturut-turut adalah 5%, 2%, dan 2%, semuanya SVM tetap menjadi lebih baik. Hasil ini menunjukkan bahwa metode SVM lebih unggul bila dibandingkan dengan metode DT, terutama dalam kasus penentuan balita beresiko stunting

IV. KESIMPULAN

Model SVM terbaik memiliki akurasi sekitar 83% pada dataset pengujian, sedangkan model Decision Tree memiliki akurasi sekitar 78%. Dari segi akurasi, keduanya model menunjukkan hasil yang baik dalam mengklasifikasikan balita beresiko stunting. Kedua model SVM dan *Decision Tree* memiliki precision yang baik dalam mengidentifikasi balita beresiko stunting, dengan nilai precision masing-masing terbaik sekitar 87% dan 87%. Model SVM juga memiliki *recall* yang baik yaitu 100% pada target '1' atau stunting, yang berarti model ini efektif dalam menentukan balita beresiko stunting. Model *Decision Tree* juga memiliki *recall* yang baik sekitar 89%, menunjukkan kemampuan yang baik dalam mendeteksi balita beresiko stunting. *F1-Score* untuk kedua model juga sekitar 89% dan 86%, yang mengindikasikan keseimbangan yang baik antara precision dan recall.

Dari keunggulan Model, Model SVM dapat disimpulkan lebih cocok jika keseimbangan antara *precision* dan *recall* sangat penting dalam konteks aplikasi tertentu. Kedua model SVM dan Decision Tree dapat digunakan untuk mengklasifikasikan balita berisiko stunting dengan baik. Pemilihan model tergantung pada kebutuhan aplikasi dan faktor-faktor lain seperti interpretasi model dan kompleksitas. Dalam penelitian ini, SVM adalah pilihan yang dapat diandalkan untuk menganalisis risiko stunting pada balita.

V. DAFTAR PUSTAKA

- [1] Sujiono, Yuliani Nurani, Mencerdaskan Perilaku Anak Usia Dini, Gramedia. Jakarta, 2007
- [2] A.D.N Yadika, K.N Berawi, and S.H Nasution, Pengaruh Stunting Terhadap Perkembangan Kognitif dan Prestasi Belajar, *J.Majot* vol 8, 2019
- [3] Ernawati A, Gambaran Penyebab Balita Stunting di Desa Lokus Stunting Kabupaten Pati, *J Litbang Media Inf Penelitian Pengembangan dan IPTEK* , 2020
- [4] Khikmanto Supribadi, Nurul Khakhim, Taufik Hery Purwanto, *Analisis Metode Support Vector Machine (Svm) untuk Klasifikasi Penggunaan Lahan Berbasis Penutup Lahan pada Citra Alos Avnir-2*. Majalah Geografi. Universitas Gadjah Mada , 2014.
- [5] Christianini Nello, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*.Cambridge University Press. Royal Holloway. University of London. 2013
- [6] Ugur Fidan, Esma dan Ismail, *Classification of Dermatological Data with Self Organizing Maps and Support Vector Machine*.Afyon Kocatepe University. 2019
- [7] Hastie, Trevor. T, Robert and F Jerome, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*.Springer Science Business Media Library. Stanford University. 2009
- [8] Shawe-Taylor, J dan Cristianini, N, *Kernel Methods for Pattern Analysis*.Cambridge University Press. 2004
- [9] Eska, *Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C45*.JURTEKSI (Jurnal Teknol. dan Sistem Informasi vol. 2, pp. 9–13. 2016
- [10] Basuki, Achmad.,Iwan Syarif, *Decission Tree*.PENS-ITS.Page 2. Politeknik Elektronika Negeri Surabaya. 2003
- [11] Ula, Mutamminul, Ananda F, Mauliza, Muhammad Abdullah, Application Of Machine Learning In Determining The Classification Of Children’s Nutritionwith Decision Tree. *Jurnal Teknik Informatika*. Vol 3. 2022
- [12] Herliansyah, Vega, *Prediksi Stunting Pada Balita Dengan Menggunakan Algoritma Klasifikasi Naive Bayes*.Telkom University. 2021
- [13] Yunus, Abdul, Sarlis M, Andi B dan Ivo , *Penerapan Algoritma Spport Vector Machine dan KNearest Neighbor Menggunakan Feature Selection Backward Elimination Untuk Prediksi Status Penderita Stunting Pada Balita*. Technoscienza. Vol.6 No.2 . Universitas Kahuripan Kediri.2022
- [14] Rally, Ivo dan Andi Bode, *Prediksi Status Penderita Stunting Pada Balita Provinsi Gorontalo Menggunakan K-Nearest Neighbor Berbasis Seleksi Fitur Chi Square*. *Jurnal Nasional Komputasi dan Teknologi Informasi* Vol.5 No 2. 2022
- [15] Yoga, Saraswati, Maya Silvi dan Syahril E, *Optimization of Support Vector Machine Algorithm Using Stunting Data Classification*. *Jurnal Prisma Sains*. Vol 11. IKIP Mataram.2022