

TESIS
**ANALISIS PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM) DAN RANDOM FOREST UNTUK KLASIFIKASI
STATUS STUNTING**



disusun oleh
NADA RIZKI FEBRIYANTI
22.55.1242
Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

TESIS
ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VEKTOR*
***MACHINE (SVM)* DAN *RANDOM FOREST* UNTUK KLASIFIKASI**
STUNTING

COMPARATIVE ANALYSIS OF SUPPORT VECTOR MACHINE (SVM)
AND RANDOM FOREST ALGORITHMS FOR STUNTING STATUS
CLASSIFICATION

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana
Program Studi Informatika



disusun oleh
NADA RIZKI FEBRIYANTI
22.55.1242
Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

HALAMAN PERSETUJUAN

ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VEKTOR MACHINE* (SVM) DAN *RANDOM FOREST* UNTUK KLASIFIKASI STUNTING

COMPARATIVE ANALYSIS OF SUPPORT VECTOR MACHINE (SVM) AND RANDOM FOREST ALGORITHMS FOR STUNTING STATUS CLASSIFICATION

yang disusun dan diajukan oleh

Nada Rizki Febrlyanti

22.55.1242

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 03 Februari 2026

Dosen Pembimbing,



Anggit Dwi Hartanto, M.Kom.

NIK. 190302163

HALAMAN PENGESAHAN

ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VEKTOR MACHINE* (SVM) DAN *RANDOM FOREST* UNTUK KLASIFIKASI STUNTING

COMPARATIVE ANALYSIS OF SUPPORT VECTOR MACHINE (SVM) AND RANDOM FOREST ALGORITHMS FOR STUNTING STATUS CLASSIFICATION

yang disusun dan diajukan oleh

Nada Rizki Febriyanti

22.55.1242

Telah dipertahankan di depan Dewan Penguji
pada tanggal 03 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096



Dr. Kumara Ari Yuana, S.T., M.T.
NIK. 190302575



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 03 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : **Nada Rizki Febriyanti**
NIM : **22.55.1242**

Menyatakan bahwa Tesis dengan judul berikut:

Analisis Perbandingan Algoritma *Support Vektor Machine (SVM)* dan *Random Forest* untuk Klasifikasi *Stunting*

Dosen Pembimbing : Prof. Dr. Kusriani, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 03 Februari 2026

Yang Menyatakan,



Nada Rizki Febriyanti

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam ke hadirat Tuhan Yang Maha Esa, karya ini saya persembahkan dengan tulus kepada:

1. **Kedua Orang Tua dan Keluarga Besar**, sumber kekuatan dan alasan utama saya untuk terus melangkah. Terima kasih yang tak terhingga atas setiap untaian doa, kasih sayang yang tulus, serta pengorbanan yang tidak mungkin bisa saya balas satu per satu. Pencapaian ini adalah wujud nyata dari dukungan dan kesabaran kalian yang luar biasa selama ini.
2. **Dosen Pembimbing dan Dosen Penguji**, yang telah meluangkan waktu dan memberikan arahan berharga. Terima kasih atas setiap diskusi, kritik yang membangun, serta ilmu yang telah dibagikan. Bimbingan Bapak dan Ibu bukan sekadar penyelesaian tugas akhir, tapi juga pelajaran berharga bagi perjalanan hidup saya ke depan.
3. **Teman-Teman Seperjuangan dan Rekan Kerja**, yang selalu hadir memberikan semangat di saat-saat tersulit. Terima kasih atas kebersamaan, bantuan, dan tawa yang membuat proses yang melelahkan ini terasa jauh lebih ringan. Kalian adalah bagian penting dari perjalanan ini.
4. **Almamater Terclnta**, tempat saya belajar banyak hal, tidak hanya soal akademik tapi juga tentang proses mendewasakan diri. Terima kasih atas segala pengalaman dan kesempatan yang telah diberikan.

Semoga karya ini bisa membawa manfaat bagi dunia ilmu pengetahuan dan menjadi pengingat bahwa setiap perjuangan yang dilakukan dengan sungguh-sungguh akan menemukan jalannya.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas limpahan rahmat, hidayah, serta karunia-Nya, sehingga tesis yang berjudul **“Anallsis Perbandingan Algoritma Support Vektor Machine (SVM) dan Random Forest untuk Klasifikasi Stunting”** ini dapat terselesaikan dengan baik. Penyusunan tesis ini merupakan salah satu syarat akademis dalam memperoleh gelar Magister Komputer (M.Kom.) pada Program Studi S2 PJJ Informatika, Universitas Amikom Yogyakarta.

Penulis menyadari bahwa keberhasilan dalam menyelesaikan tugas akhir ini tidak lepas dari bimbingan, dukungan, serta bantuan tulus dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan apresiasi dan terima kasih yang mendalam kepada:

1. **Prof. Dr. Kusriani, M.Kom.**, selaku Dosen Pembimbing Utama sekaligus Dosen Penguji Ketiga, yang dengan penuh kesabaran telah memberikan motivasi, bimbingan, arahan, dan masukan yang sangat berharga selama proses penyusunan tesis ini;
2. **Anggit Dwi Hartanto, M.Kom.**, selaku Dosen Pembimbing Kedua, terima kasih atas waktu dan dedikasi Bapak dalam memberikan arahan serta koreksi detail yang sangat membantu penulis dalam merapikan alur berpikir dan hasil penelitian ini.
3. **Hanif Al Fatta, S.Kom., M.Kom., Ph.D. dan Dr. Kumara Ari Yuana, S.T., M.T.** selaku Dosen Penguji, penulis haturkan terima kasih yang tulus atas diskusi yang membangun, kritik yang mencerahkan, serta dukungan semangat yang diberikan demi kesempurnaan akhir tesis ini;
4. **Prof. Dr. M. Suyanto, M.M.**, selaku Rektor Universitas Amikom Yogyakarta, yang telah memberikan kesempatan dan fasilitas kepada penulis untuk menyelesaikan studi dan penelitian ini;
5. **Seluruh Bapak dan Ibu Dosen di Program Studi S2 PJJ Informatika, Universitas Amikom Yogyakarta**, yang telah memberikan ilmu, pengalaman, dan dukungan akademik selama masa studi;

6. **Seluruh Admisi, Staf, dan Karyawan di Program Studi S2 PJJ Informatika**, Universitas Amikom Yogyakarta, yang telah memberikan bantuan dan pelayanan yang ramah serta sabar, sehingga sangat membantu kelancaran masa studi penulis;
7. **Keluarga Tercinta**, yang selalu memberikan doa, dukungan, dan semangat yang tiada henti, serta pengorbanan yang tak ternilai dalam perjalanan pendidikan penulis;
8. **Sahabat dan Rekan – Rekan Seperjuangan**, yang selalu memberikan semangat, motivasi, dan kebersamaan yang indah selama masa studi dan penyusunan tesis ini.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna. Oleh karena itu, penulis dengan tangan terbuka menerima segala bentuk saran dan kritik yang membangun demi perbaikan di masa depan. Akhir kata, penulis berharap semoga karya ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan dan menjadi referensi yang berguna bagi para pembaca.

Yogyakarta, 03 Februari 2026

Penulis

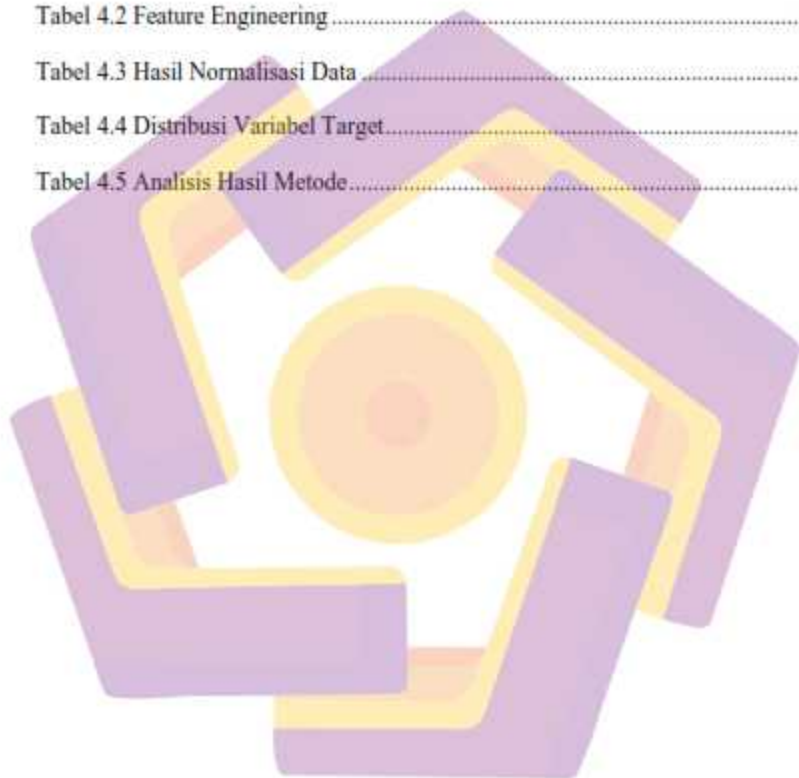
DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PERNYATAAN KEASLIAN TESIS.....	iv
HALAMAN PERSEMBAHAN.....	v
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
INTISARI.....	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	6
1.3. Batasan Masalah.....	7
1.4. Tujuan Penelitian.....	8
1.5. Manfaat Penelitian.....	9

BAB II TINJAUAN PUSTAKA.....	11
2.1. Tinjauan Pustaka.....	11
2.2. Keaslian Penelitian.....	18
2.3. Landasan Teori.....	24
BAB III METODE PENELITIAN.....	39
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	39
3.2. Metode Pengumpulan Data.....	40
3.3. Metode Analisis Data.....	44
3.4. Alur Penelitian.....	45
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	46
4.1. Hasil Penelitian.....	46
4.1 Deskripsi Data Awal.....	46
4.2 Persiapan Data.....	47
4.4 <i>Modelling</i>	47
4.5 Evaluasi.....	49
4.6 Validasi data Eksternal.....	60
4.6 Analisis Hasil.....	65
BAB V PENUTUP.....	67
5.1. Kesimpulan.....	67
5.2. Saran.....	68
DAFTAR PUSTAKA.....	69
LAMPIRAN.....	75

DAFTAR TABEL

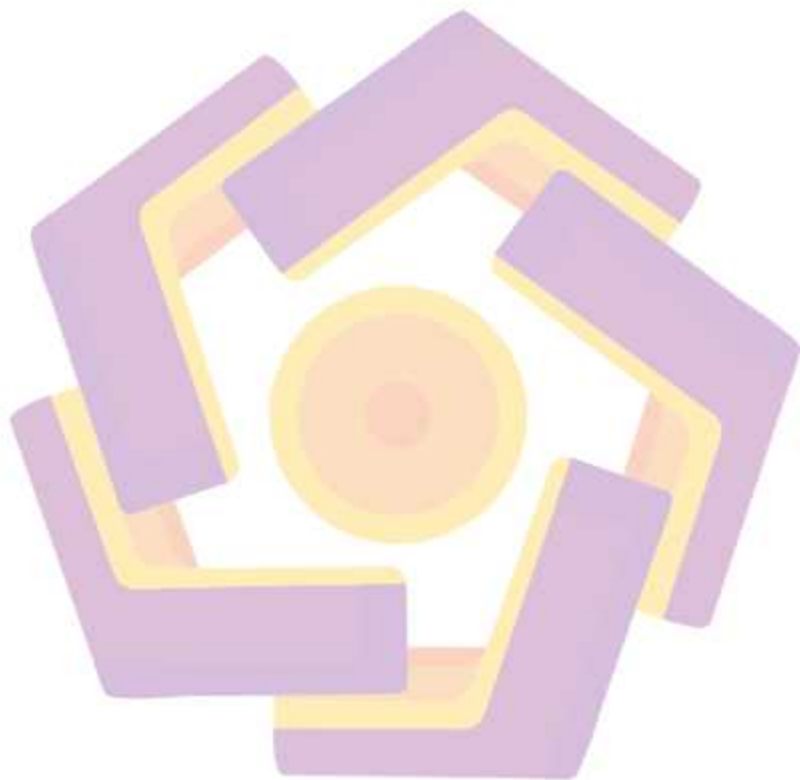
Tabel 2.1. Matriks literatur review.....	13
Tabel 4.1 Struktur Dataset Awal.....	47
Tabel 4.2 Feature Engineering.....	49
Tabel 4.3 Hasil Normalisasi Data.....	54
Tabel 4.4 Distribusi Variabel Target.....	55
Tabel 4.5 Analisis Hasil Metode.....	60



DAFTAR GAMBAR

Gambar 2.1 Struktur SVM Garis Hyperplane.....	28
Gambar 2.2 Flowchart Algoritma SVM.....	31
Gambar 2.3 Ilustrasi Metode Random Forest	33
Gambar 3. 1 Alur Penelitian.....	44
Gambar 4. 1 Hasil Missing Value.....	49
Gambar 4. 2 Hasil Prediksi Model 80:20.....	51
Gambar 4. 3 Hasil Prediksi Model 70:30.....	51
Gambar 4. 4 Confusion Matrix Support Vector Machine 80:20.....	52
Gambar 4. 5 Hasil Confusion Matrix Support Vector Machine 80:20.....	52
Gambar 4. 6 Confusion Matrix Support Vector Machine 80:20	53
Gambar 4. 7 Hasil Confusion Matrix Support Vector Machine 80:20	53
Gambar 4. 8 Confusion Matrix Support Vector Machine 70:30	54
Gambar 4. 10 Hasil Confusion Matrix Support Vector Machine 70:30	54
Gambar 4. 11 Confusion Matrix Random Forest 80:20	55
Gambar 4. 12 Hasil Confusion Matrix Random Forest 80:20	55
Gambar 4. 13 Confusion Matrix Random Forest 80:20	56
Gambar 4. 14 Hasil Confusion Matrix Random Forest 80:20	56
Gambar 4. 15 Confusion Matrix Random Forest 70:30	57
Gambar 4. 16 Hasil Confusion Matrix Random Forest 70:30	57
Gambar 4. 17 Confusion Matrix Random Forest 80:20	58
Gambar 4. 18 Hasil Confusion Matrix Random Forest 80:20	58

Gambar 4. 19 Kurva ROC Random Forest	59
Gambar 4. 20 Kurva ROC SVM	59
Gambar 4. 21 Validasi Silang 5-Fold.....	60



INTISARI

Masalah *stunting* merupakan tantangan gizi kronis yang signifikan di Indonesia, khususnya di Kota Banjarmasin. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Support Vector Machine* dan *Random Forest* dalam mengklasifikasikan status *stunting* pada balita. Data penelitian bersumber dari Dinas Kesehatan Kota Banjarmasin tahun 2024 sebanyak 2.231 rekam data, dengan tambahan data validasi eksternal dari Kabupaten Jeneponto tahun 2025. Tantangan berupa ketidakseimbangan kelas (*class imbalance*) ditangani dengan penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE). Hasil penelitian menunjukkan bahwa algoritma *Random Forest* secara konsisten mengungguli *Support Vector Machine* pada seluruh skenario pengujian. Pada konfigurasi data optimal 80:20 dengan penerapan SMOTE, *Random Forest* menghasilkan akurasi sebesar 93,67%, nilai AUC sebesar 0,980, dan *F1-score* sebesar 93,1%. Analisis *feature importance* mengungkapkan bahwa variabel tinggi badan dan usia merupakan kontributor paling determinan dalam proses klasifikasi, yang selaras dengan standar antropometri medis. Uji validasi eksternal menunjukkan kemampuan generalisasi model yang baik dengan akurasi 83,30% dan *recall* 0,930. Penelitian ini menyimpulkan bahwa *Random Forest* merupakan metode yang lebih efektif dan stabil untuk mendukung identifikasi status *stunting* balita secara akurat.

Kata Kunci: *Stunting*, Klasifikasi, *Support Vector Machine*, *Random Forest*, SMOTE.

ABSTRACT

Stunting is a significant chronic nutritional challenge in Indonesia, particularly in Banjarmasin City. This study aims to compare the performance of the Support Vector Machine and Random Forest algorithms in classifying stunting status among toddlers. The research data was sourced from the Banjarmasin City Health Office in 2024, consisting of 2,231 records, with additional external validation data from Jenepono Regency in 2025. The challenge of class imbalance was addressed using the Synthetic Minority Over-sampling Technique (SMOTE). The results demonstrate that the Random Forest algorithm consistently outperformed the Support Vector Machine across all testing scenarios. In the optimal 80:20 data configuration with SMOTE application, Random Forest achieved an accuracy of 93.67%, an AUC of 0.980, and an F1-score of 93.1%. Feature importance analysis revealed that body height and age are the most determinant contributors to the classification process, aligning with medical anthropometric standards. External validation testing showed strong model generalization capability with an accuracy of 83.30% and a recall of 0.930. This study concludes that Random Forest is a more effective and stable method for supporting the accurate identification of toddler stunting status.

Keywords: *Stunting, Classification, Support Vector Machine, Random Forest, SMOTE*



BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Stunting adalah persoalan gizi kronis yang masih menjadi tantangan serius di berbagai belahan dunia, khususnya di wilayah yang memiliki keterbatasan ekonomi dan akses terhadap sumber daya kesehatan. Kondisi ini tercermin dari pertumbuhan tinggi badan anak yang berada di bawah standar usianya sebagai akibat dari kekurangan asupan gizi dalam jangka panjang, terutama pada periode krusial seribu hari pertama kehidupan. Pada kelompok anak, hambatan pertumbuhan tersebut berpotensi menimbulkan konsekuensi jangka panjang yang serius terhadap kondisi kesehatan serta proses perkembangan fisik dan kognitif mereka. Masalah ini memberikan pengaruh terhadap perkembangan kognitif yang rendah dan tingkat kecerdasan yang menurun. Selain itu, stunting berdampak signifikan terhadap sumber daya manusia secara keseluruhan (Aryastami & Tarigan, 2017). Para ahli gizi menekankan bahwa periode kritis pertumbuhan anak berlangsung sejak masa kehamilan hingga anak mencapai usia dua tahun, yang secara luas dikenal sebagai periode seribu hari pertama kehidupan. Kegagalan pertumbuhan pada periode ini sulit untuk dikejar pada periode berikutnya dan dapat berpengaruh terhadap kualitas kesehatan, kapasitas intelektual, serta tingkat produktivitas individu ketika memasuki usia dewasa.

Indonesia masih menghadapi masalah serius terkait stunting, dengan angka kejadian sebesar 24,4% pada balita berdasarkan data pada tahun 2021

oleh Riset Kesehatan Dasar (Dhewi, 2024; Kurniawati & Ardiansyah, 2022). Kondisi ini menempatkan Indonesia di posisi kedua tertinggi di kawasan ASEAN. Dibandingkan dengan negara tetangga, situasi di Indonesia masih memprihatinkan, mengingat Vietnam hanya memiliki angka prevalensi 23%, Malaysia 17%, Thailand 16%, dan Singapura jauh lebih rendah yaitu hanya 4% (Rosida et al., 2023). Berdasarkan data Anwar et al. (2022) dan Anwar (2024), Kota Banjarmasin mengalami perubahan angka stunting yang tidak stabil dalam tiga tahun terakhir yakni 27,80% pada tahun 2021, turun menjadi 22,40% di tahun 2022, kemudian naik kembali menjadi 26,5% pada tahun 2023. Kondisi yang tidak stabil tersebut diduga berkaitan dengan pelaksanaan intervensi kesehatan yang belum berkesinambungan serta ketimpangan sosial ekonomi yang masih terjadi, di mana variabel seperti pendapatan rumah tangga dan tingkat pendidikan ibu terbukti memiliki keterkaitan yang signifikan dengan status gizi anak usia balita (Marniati et al., 2020).

Dampak buruk stunting dalam jangka panjang mencakup penurunan kekebalan tubuh, peningkatan risiko penyakit diabetes, *stroke*, kanker, penyakit jantung, hingga penurunan kualitas kerja (Agustian et al., 2023; Anggryni et al., 2021). Penelitian lain menambahkan bahwa dampak jangka panjang lainnya meliputi pertumbuhan yang pendek, peningkatan risiko obesitas, dan penyakit degeneratif di masa mendatang (Anmaru & Laksono, 2019). Kajian Kemenkes (2013) yang dikutip oleh Mentari & Hermansyah (2019) melaporkan bahwa berdasarkan usia balita, prevalensi stunting cenderung lebih tinggi pada anak berusia 24-59 bulan dibandingkan dengan

balita berusia 0-23 bulan. Fenomena ini menunjukkan efek kumulatif dari kekurangan gizi kronis semakin terlihat jelas pada usia yang lebih tua. Penelitian longitudinal menunjukkan bahwa pola pertumbuhan anak yang mengalami stunting seringkali sudah menunjukkan perlambatan sejak usia 6 bulan, namun baru terdeteksi secara signifikan pada usia 2 tahun ke atas (Kusuma et al., 2021).

Permasalahan stunting tidak semata-mata berpengaruh terhadap kondisi kesehatan individu, tetapi juga memberikan konsekuensi ekonomi yang signifikan pada skala nasional dan global. Stunting menyebabkan hilangnya 2-3% PDB (Produk Domestik Bruto) di setiap negara pada tiap tahunnya (Galasso & Wagstaff, 2019). Di Indonesia sendiri, dampak ekonomi yang ditimbulkan oleh stunting diperkirakan mencapai sekitar 2-3% dari PDB. Biaya ini timbul dari peningkatan anggaran kesehatan untuk menangani penyakit terkait stunting, penurunan produktivitas kerja, dan hilangnya potensi sumber daya manusia. Fakta ini menegaskan pentingnya identifikasi dini secara akurat untuk memetakan risiko status stunting sebagai dasar upaya pencegahan yang efektif. Dalam menghadapi tantangan kompleks ini, penggunaan analisis data dan teknologi kecerdasan buatan menjadi sarana pendukung pengambilan keputusan yang tepat sasaran melalui klasifikasi status gizi yang presisi.

Penelitian komprehensif di 42 negara dengan tingkat pendapatan rendah hingga menengah mengidentifikasi berbagai faktor utama yang berperan dalam terjadinya stunting bersifat *multidimensional* dan bervariasi

antar wilayah geografis (Haddad et al., 2020). Analisis tersebut mengungkapkan bahwa pendekatan “*one-size-fits-all*” kurang efektif dalam upaya penanggulangan stunting, sehingga diperlukan strategi penanganan yang lebih fleksibel serta disesuaikan dengan konteks lokal. Penelitian ini juga menyoroti peran teknologi klasifikasi dalam mengidentifikasi kelompok berisiko tinggi dan mengoptimalkan alokasi sumber daya intervensi. Laporan global WHO (2020) mengenai status nutrisi anak menekankan pentingnya identifikasi dini faktor risiko stunting melalui pendekatan berbasis data. Laporan tersebut menggarisbawahi bahwa pemanfaatan teknologi klasifikasi dapat secara signifikan meningkatkan efektivitas intervensi stunting, terutama di negara-negara berkembang dengan keterbatasan sumber daya kesehatan.

Pengembangan model klasifikasi untuk stunting memerlukan algoritma yang handal dan akurat. Berbagai pendekatan *machine learning* telah diterapkan dalam beberapa tahun terakhir, dengan *Support Vector Machine* dan *Random Forest* menjadi dua algoritma yang menunjukkan potensi besar. *Support Vector Machine* adalah salah satu algoritma dalam pembelajaran mesin yang dikenal memiliki tingkat efektivitas yang tinggi dalam memecahkan masalah klasifikasi dengan ruang fitur yang kompleks. Algoritma ini bekerja dengan menemukan *hyperplane* optimal yang memisahkan kelas-kelas data (misalnya, stunting dan non-stunting) dengan margin maksimal. Keunggulan *Support Vector Machine* terletak pada kemampuannya mengolah data dengan dimensi yang tinggi serta

menghasilkan performa generalisasi yang kuat, yang menjadikannya kandidat potensial untuk klasifikasi stunting yang melibatkan berbagai faktor risiko.

Sejumlah studi terdahulu telah membuktikan tingkat efektivitas penggunaan *Support Vector Machine* dalam mengklasifikasikan kasus stunting (Nisa et al., 2023; Anggraini et al., 2022). Penelitian tersebut mengidentifikasi beberapa faktor risiko kunci seperti berat badan lahir rendah, panjang badan lahir, dan status imunisasi sebagai fitur signifikan untuk klasifikasi stunting. Model *Support Vector Machine* yang dikembangkan mampu mengelompokkan kasus stunting dan non-stunting dengan tingkat akurasi yang baik.

Di sisi lain, *Random Forest* merupakan pendekatan *ensemble* yang bekerja dengan mengombinasikan sejumlah pohon keputusan guna menghasilkan klasifikasi yang lebih akurat dan konsisten. Keunggulan metode ini terletak pada kemampuannya dalam mengolah data yang tidak seimbang, mengatasi *noise* data, dan mengidentifikasi fitur-fitur penting (*feature importance*) dalam dataset. Algoritma ini juga relatif tahan terhadap *overfitting*, yang menjadi masalah umum dalam model klasifikasi yang kompleks. Karakteristik-karakteristik ini menjadikan *Random Forest* sebagai pilihan yang menarik untuk klasifikasi stunting yang sering melibatkan data dengan distribusi tidak seimbang dan variabel yang saling berkorelasi.

Berbagai penelitian telah menunjukkan efektivitas *Random Forest* dalam mengklasifikasikan kondisi kesehatan yang kompleks, termasuk stunting (Saputra & Wijaya, 2023; Lestari et al., 2022). Penelitian-penelitian

tersebut menggunakan pendekatan *machine learning ensemble* untuk menganalisis dan mengelompokkan risiko stunting berdasarkan faktor-faktor seperti status ekonomi keluarga, pola asuh, dan riwayat kesehatan ibu dan anak, dengan hasil akurasi yang baik. Melalui pendekatan *feature importance*, penelitian-penelitian tersebut berhasil mengidentifikasi faktor-faktor kunci yang paling membedakan antara kelompok stunting dan non-stunting.

Penelitian lain membandingkan berbagai algoritma *machine learning* dalam klasifikasi stunting di daerah perdesaan Indonesia (Mayasari et al., 2021). Temuan penelitian memperlihatkan bahwa *Random Forest* dan *Support Vector Machine* secara konsisten menunjukkan kinerja klasifikasi yang lebih unggul dibandingkan dengan algoritma lain. Penelitian tersebut mengidentifikasi bahwa *Random Forest* unggul dalam sensitivitas (kemampuan mendeteksi kasus positif stunting), sementara *Support Vector Machine* menunjukkan spesifisitas yang lebih baik (kemampuan mengidentifikasi kasus non-stunting dengan tepat). Temuan ini memberikan dasar yang kuat untuk analisis komparatif yang lebih mendalam antara kedua algoritma ini dalam konteks spesifik klasifikasi stunting.

1.2. Rumusan Masalah

Mengacu pada uraian latar belakang yang telah dipaparkan, dapat disimpulkan bahwa permasalahan yang dikaji dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana kinerja algoritma *Support Vector Machine* dalam mengklasifikasikan status *stunting* pada anak-anak?
2. Bagaimana kinerja algoritma *Random Forest* dalam mengklasifikasikan status *stunting* pada anak-anak?
3. Algoritma *machine learning* manakah yang memiliki tingkat efektivitas lebih baik antara *Support Vector Machine* dan *Random Forest* dalam melakukan klasifikasi *stunting*?

1.3. Batasan Masalah

Guna memfokuskan penelitian dan memberikan kejelasan ruang lingkup permasalahan yang akan dibahas, maka penelitian ini dibatasi pada beberapa hal sebagai berikut:

1. Penelitian ini hanya berfokus pada perbandingan performa algoritma *Support Vector Machine* dan *Random Forest* untuk klasifikasi *stunting*, tidak termasuk algoritma *machine learning* lainnya.
2. Data yang digunakan dalam penelitian ini dibatasi pada data status gizi anak balita di wilayah Kota Banjarmasin.
3. Penelitian ini berfokus pada kinerja model klasifikasi dan tidak menyertakan analisis mendalam mengenai hubungan sebab akibat (*causal analysis*) antara faktor risiko dengan kejadian *stunting*.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan sebelumnya, tujuan dari penelitian ini adalah sebagai berikut:

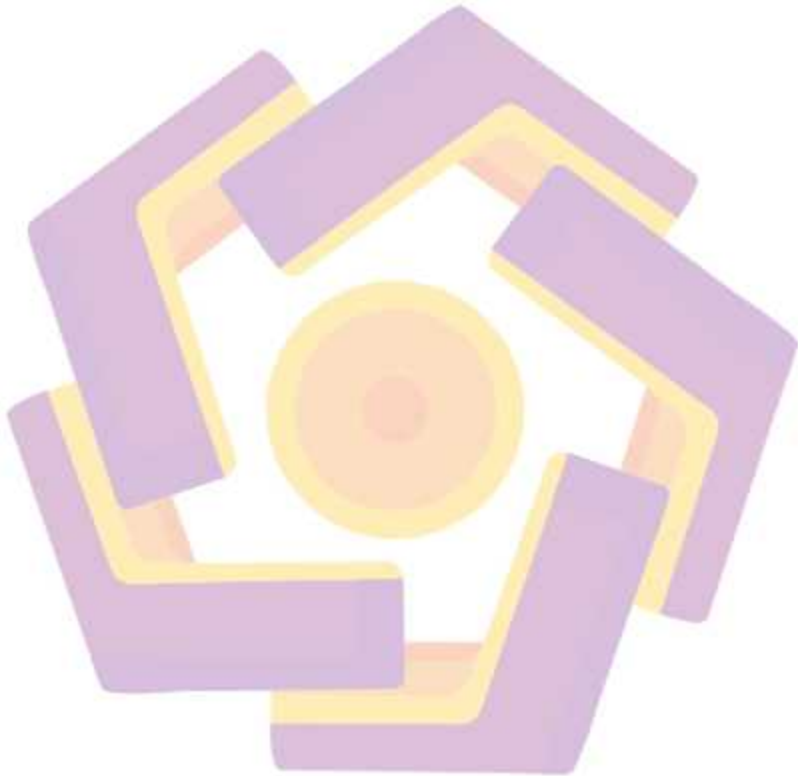
1. Mengukur kinerja algoritma *Support Vector Machine* dalam mengklasifikasikan status *stunting* pada anak-anak.
2. Mengukur kinerja algoritma *Random Forest* dalam mengklasifikasikan status *stunting* pada anak-anak.
3. Menganalisis dan menentukan algoritma *machine learning* yang memiliki tingkat efektivitas lebih baik antara *Support Vector Machine* dan *Random Forest* dalam melakukan klasifikasi *stunting*.

1.5. Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini dapat diuraikan sebagai berikut:

1. Memberikan kontribusi ilmiah dalam pengembangan metodologi *machine learning* untuk bidang kesehatan masyarakat, khususnya dalam konteks klasifikasi *stunting*.
2. Memperluas pemahaman tentang kekuatan dan keterbatasan algoritma SVM dan *Random Forest* dalam konteks data kesehatan yang memiliki karakteristik spesifik seperti ketidakseimbangan kelas dan interaksi kompleks antar variabel.

3. Menjadi sumber rujukan dan acuan bagi peneliti selanjutnya yang ingin mengembangkan atau membandingkan algoritma klasifikasi pada permasalahan kesehatan masyarakat yang serupa.



BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

Dasar tinjauan pustaka pada penelitian ini dari berbagai penelitian terdahulu yang membahas penerapan *machine learning* dalam proses klasifikasi maupun deteksi dini stunting. Kajian pustaka ini secara khusus menitikberatkan pada jenis algoritma pada jenis algoritma yang digunakan dalam penelitian-penelitian tersebut. Adapun tujuan utama penelitian ini untuk melakukan analisis perbandingan kinerja algoritma SVM dan Random Forest dalam mengklasifikasikan status stunting anak usia balita.

Saputra dan Wijaya (2023) melaksanakan sebuah penelitian yang berfokus pada pemanfaatan algoritma Random Forest untuk mengelompokkan tingkat risiko stunting pada balita. Penelitian tersebut diarahkan untuk mengenali anak-anak yang berpotensi mengalami stunting melalui analisis berbagai karakteristik sosial, demografis, serta kondisi lingkungan yang berkaitan dengan tumbuh kembang anak. Algoritma Random Forest digunakan karena kemampuannya dalam menangani data dengan banyak fitur dan ketahanannya terhadap *overfitting*. Temuan penelitian tersebut menunjukkan bahwa model yang dikembangkan memiliki kemampuan yang baik dalam mengelompokkan tingkat risiko stunting, dengan tingkat ketepatan klasifikasi mencapai 87,5%, dengan faktor pendapatan keluarga dan pendidikan ibu teridentifikasi sebagai fitur paling penting melalui analisis *feature importance*.

Penelitian berikutnya tentang klasifikasi status stunting dengan SVM dilakukan oleh Nisa dkk. (2023). Tujuan penelitian ini adalah untuk mengembangkan model deteksi dini stunting berdasarkan data antropometri dan riwayat kesehatan anak. Berbagai jenis kernel digunakan pada penelitian ini yang menggunakan algoritma SVM untuk menemukan pemisah optimal antara kelas stunting dan non-stunting. Hasil pengujian menunjukkan bahwa model SVM dengan kernel RBF mencapai akurasi tertinggi sebesar 89,2%, dengan performa

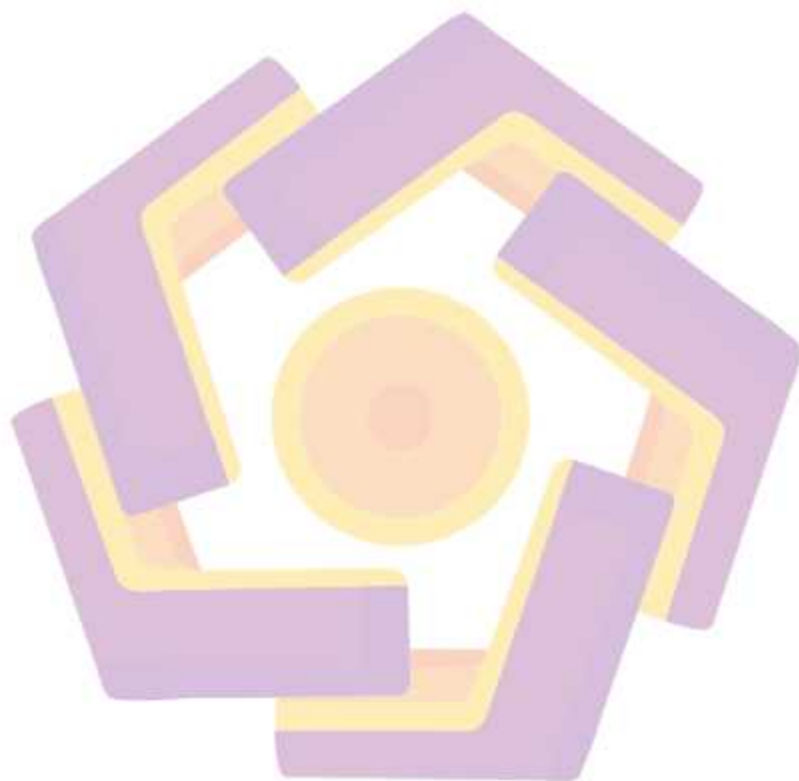
khususnya sangat baik dalam spesifisitas, yaitu mengurangi kesalahan dalam mengklasifikasikan anak sehat sebagai berisiko stunting.

Penelitian komparatif antara berbagai algoritma untuk klasifikasi stunting dilakukan oleh Mayasari dkk. (2021). Penelitian tersebut melakukan analisis perbandingan terhadap performa sejumlah algoritma, di antaranya Random Forest dan SVM, dalam mengidentifikasi kondisi stunting pada anak usia balita yang tinggal di wilayah pedesaan. Berdasarkan hasil pengujian, diperoleh temuan bahwa Random Forest dan SVM konsisten menempati peringkat dua terbaik. Random Forest mencapai recall (sensitivitas) tertinggi sebesar 85%, yang berarti sangat baik dalam menemukan kasus stunting yang sesungguhnya. Sementara itu, SVM mencapai precision tertinggi sebesar 88%, yang menunjukkan kemampuannya dalam meminimalkan kesalahan klasifikasi positif pada anak yang sebenarnya normal.

Berdasarkan hasil tinjauan pustaka yang telah diuraikan, dapat dipahami bahwa algoritma Random Forest dan SVM telah menunjukkan tingkat efektivitas yang baik dalam berbagai penelitian, khususnya dalam melakukan klasifikasi terhadap kondisi stunting. Namun, analisis perbandingan langsung yang mendalam antara kedua algoritma tersebut dengan menggunakan dataset yang komprehensif dan serangkaian metrik evaluasi yang lengkap masih dapat dikembangkan lebih lanjut. Dengan demikian, penelitian ini bertujuan untuk melaksanakan analisis perbandingan secara terstruktur antara algoritma Random Forest dan SVM dalam proses pengklasifikasian status stunting. Upaya ini dilakukan untuk mengidentifikasi algoritma yang memiliki kinerja paling optimal, sehingga dapat dimanfaatkan secara efektif dalam mendukung skrining serta pelaksanaan intervensi dini terhadap kasus stunting.

Selain itu, penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam memperkaya kajian mengenai penerapan metode machine learning pada bidang kesehatan masyarakat, khususnya dalam upaya deteksi dan klasifikasi stunting secara lebih akurat. Hasil penelitian ini juga diharapkan mampu menjadi

dasar pertimbangan dalam pemilihan algoritma yang paling sesuai untuk digunakan dalam sistem pendukung keputusan, terutama bagi tenaga kesehatan dan pemangku kebijakan. Dengan adanya model klasifikasi yang memiliki tingkat ketepatan yang tinggi, proses identifikasi kelompok berisiko dapat dilakukan secara lebih cepat dan tepat sasaran. Pada akhirnya, temuan penelitian ini diharapkan dapat mendukung perencanaan program pencegahan dan penanganan stunting secara lebih efektif, terarah, dan berbasis data



2.2. Keaslian Penelitian

Tabel 2.1. Matriks literatur review dan posisi penelitian

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1.	<i>Klasifikasi Balita Stunting Menggunakan Random Forest Classifier Di Kabupaten Blitar</i>	Peneliti: M. Ramadani Akbar Ariyadi, Sri Lestanti, Sabitul Kirom Tahun : 2023 Publikasi : <i>JATI (Jurnal Mahasiswa Teknik Informatika)</i>	Penelitian ini bertujuan untuk mencrapkan algoritma Random Forest sebagai metode klasifikasi dalam menentukan status stunting pada balita, serta mengevaluasi kinerja model melalui analisis Confusion Matrix.	Hasil pengujian menggunakan algoritma Random Forest pada data balita menunjukkan bahwa konfigurasi terbaik diperoleh dengan membagi data sebesar 80% untuk proses pelatihan dan 20% untuk proses pengujian. Pada skenario tersebut, model berhasil mencapai tingkat akurasi tertinggi sebesar 90,1%, dengan nilai presisi sebesar 71,4% dan recall sebesar 62,5%. Temuan ini menunjukkan bahwa algoritma Random	Saran untuk penelitian selanjutnya adalah mempertimbangkan teknik penyeimbangan data agar nilai <i>recall</i> meningkat, serta menggunakan dataset yang lebih luas agar model lebih representatif.	Ariyadi, dkk. melakukan klasifikasi status stunting menggunakan algoritma Random Forest, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi dengan melakukan komparasi metode Support Vector Machine dan Random Forest.

				Forest memiliki kemampuan yang cukup baik dalam mengklasifikasikan status stunting pada balita berdasarkan data yang digunakan.		
2.	Klasifikasi Penyakit Stunting Dengan Menggunakan Algoritma Support Vector Machine Dan Random Forest	Peneliti: Mahudin Banurea, Dinda Betaria Hutagaol, Oloan Sihombing Publikasi: Jurnal TEKINKOM Tahun: 2023	Penelitian ini bertujuan untuk melakukan analisis evaluatif dan perbandingan antara algoritma Support Vector Machine (SVM) dan Random Forest dalam proses klasifikasi stunting.	Hasil penelitian menunjukkan bahwa algoritma SVM menghasilkan tingkat akurasi sebesar 65,6% pada data pengujian dan 62,7% pada data pelatihan. Sementara itu, algoritma Random Forest menunjukkan performa yang lebih baik dengan tingkat akurasi sebesar 88,2% pada data pengujian dan 98,8% pada data pelatihan. Setelah dilakukan proses penyesuaian parameter (<i>hypertuning</i>) pada algoritma SVM, tingkat akurasinya meningkat secara signifikan hingga mencapai 81%.	Penelitian selanjutnya diharapkan dapat melakukan optimasi parameter yang lebih mendalam dan menambah jumlah dataset untuk menghindari indikasi <i>overfitting</i> pada model.	Banurea, dkk. melakukan analisis evaluatif antara SVM dan Random Forest, sedangkan dalam penelitian tesis ini akan melakukan komparasi kedua metode tersebut dengan fokus pada optimasi parameter yang berbeda

3.	<p>Optimization of Support Vector Machine Algorithm Using Stunting Data Classification</p>	<p>Peneliti: Saraswati Yoga, Andriyani, Maya Silvia Lydia, Syahril Efendi</p> <p>Tahun: 2023</p> <p>Publikasi: <i>Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram</i></p>	<p>Penelitian ini bertujuan untuk melakukan perbandingan kinerja beberapa jenis kernel pada algoritma Support Vector Machine (SVM), yaitu kernel linear, polynomial, sigmoid, dan radial basis function, setelah melalui tahap optimasi.</p>	<p>Hasil penelitian yang memanfaatkan data status gizi menunjukkan bahwa penerapan algoritma Support Vector Machine yang dioptimalkan menggunakan metode Particle Swarm Optimization memberikan tingkat kinerja yang bervariasi pada setiap jenis kernel yang diuji. Dari seluruh pengujian yang dilakukan, kernel radial basis function menunjukkan performa terbaik dengan nilai akurasi sebesar 78%, presisi sebesar 89%, recall sebesar 66%, dan skor F1 sebesar 72%.</p>	<p>Disarankan bagi penelitian mendatang untuk mempertimbangkan faktor-faktor lain seperti sosial ekonomi dan melakukan seleksi fitur guna meningkatkan performa model SVM yang dioptimasi.</p>	<p>Yoga, dkk. melakukan optimasi algoritma SVM menggunakan metode PSO, sedangkan dalam penelitian tesis ini akan melakukan komparasi antara metode Support Vector Machine dan Random Forest untuk klasifikasi stunting.</p>
4.	<p>Klarifikasi Status Penderita Gizi Stunting Pada Balita Menggunakan Metode Random Forest</p>	<p>Peneliti: Yunita Nur Aprilia, Dian Ahkam Sani, Nanda Martyan Angadimas</p> <p>Tahun : 2024</p>	<p>Penelitian ini bertujuan untuk mengembangkan model klasifikasi berbasis <i>machine learning</i> dengan memanfaatkan algoritma Random Forest guna menentukan status gizi stunting pada balita</p>	<p>Hasil penelitian menunjukkan bahwa sistem yang dikembangkan berhasil membentuk model klasifikasi yang efektif dalam menilai dan</p>	<p>Saran untuk penelitian kedepan adalah menggunakan dataset yang lebih besar dan representatif agar model Random Forest memiliki tingkat keandalan yang lebih</p>	<p>Aprilia, dkk. mengembangkan model klasifikasi stunting menggunakan metode Random Forest, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi</p>

		Publikasi: INTEGER: Journal of Information Technology	berdasarkan data nutrisi yang tersedia.	mengevaluasi kinerja algoritma Random Forest dalam menentukan status gizi stunting pada balita. Dengan menggunakan dataset yang terdiri dari 300 data, model tersebut mampu mencapai nilai akurasi rata-rata sebesar 81%, nilai presisi rata-rata sebesar 76%, nilai recall rata-rata sebesar 69%, serta skor F1 rata-rata sebesar 72%. Hasil ini menunjukkan bahwa algoritma Random Forest memiliki kemampuan yang cukup baik dan dapat diandalkan dalam proses klasifikasi status gizi stunting berdasarkan data nutrisi yang tersedia.	baik dalam menentukan status gizi.	dengan melakukan komparasi metode Support Vector Machine dan Random Forest.
5.	Optimalisasi Algoritma Support Vector Machine (SVM) Dalam	Peneliti: Sonya Dian Wahyuni, R. Hadapiningradja Kusumodestoni	Pengujian model ini bertujuan untuk meningkatkan kinerja algoritma SVM dalam mengklasifikasikan data	Hasil penelitian menunjukkan bahwa penerapan algoritma SVM dengan kernel Radial Basis	Penelitian selanjutnya dapat mencoba membandingkan performa SVM dengan algoritma berbasis	Wahyuni, dkk. melakukan optimalisasi pada algoritma SVM untuk data stunting, sedangkan dalam

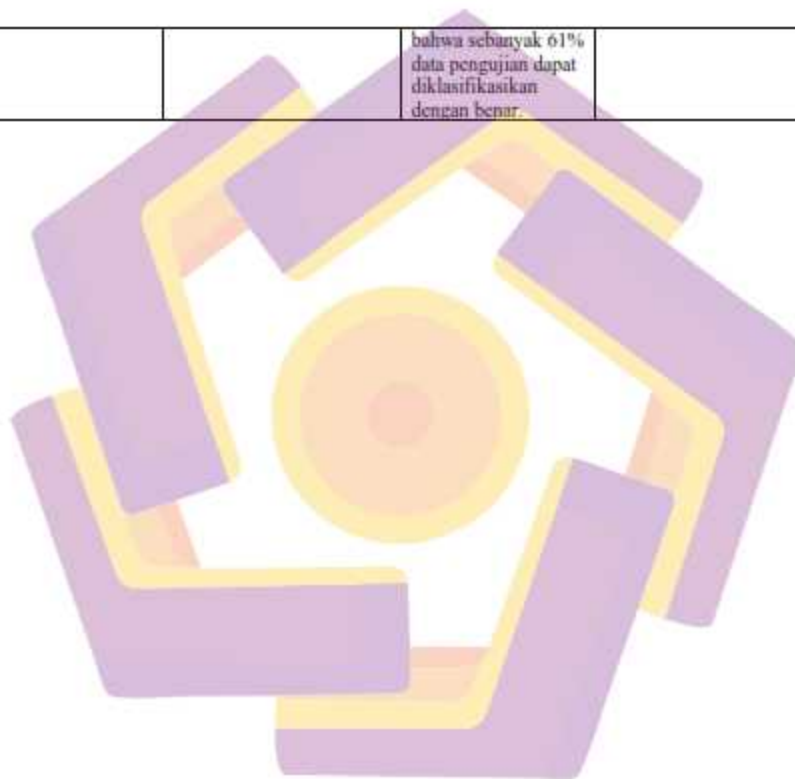
	Klasifikasi Kejadian Data Stunting	Publikasi: Bulletin of Information Technology (BIT) Tahun : 2024	kejadian stunting, dengan penekanan pada upaya memperoleh tingkat akurasi yang paling optimal.	Function (RBF) mampu menghasilkan model yang efektif dalam mengklasifikasikan kasus stunting. Model tersebut mencapai tingkat akurasi yang tinggi, yaitu sebesar 92,65%, serta nilai Area Under Curve (AUC) sebesar 0,926, yang menunjukkan kemampuan klasifikasi yang sangat baik.	pohon lainnya dan mempertimbangkan faktor lingkungan yang mempengaruhi kejadian stunting	penelitian tesis ini akan melakukan klasifikasi dengan melakukan komparasi metode Support Vector Machine dan Random Forest.
6.	Penerapan Algoritma Support Vector Machine untuk Klasifikasi Stunting pada Balita di Kabupaten Enrekang	Peneliti: Andi Widya Mufila Gaffar, Andi Muhammad Halis, Purnawansyah, Sitti Rahmah Jabir Tahun : 2024 Publikasi: <i>Jurnal Minfo Polgan,</i>	Penelitian ini bertujuan untuk menerapkan sekaligus mengevaluasi algoritma Support Vector Machine (SVM), khususnya yang menggunakan kernel polynomial, dalam mengklasifikasikan kondisi balita di Kabupaten Enrekang berdasarkan status stunting atau tidak.	Hasil penelitian menunjukkan bahwa penggunaan algoritma SVM dalam klasifikasi status stunting pada balita menghasilkan tingkat ketepatan yang sangat tinggi. Berdasarkan pengujian menggunakan metode 10-Fold Cross Validation, performa terbaik diperoleh pada fold	Disarankan agar penelitian berikutnya memperluas cakupan wilayah dan menggunakan variabel yang lebih beragam untuk memastikan konsistensi akurasi algoritma SVM.	Gaffar, dkk. menerapkan algoritma SVM untuk klasifikasi stunting di Kabupaten Enrekang, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi dengan melakukan komparasi metode Support Vector Machine dan Random Forest.

				<p>ke-4 dengan nilai accuracy, precision, recall, dan F1-score masing-masing sebesar 99,13%. Sementara itu, performa terendah tercatat pada fold ke-0, dengan nilai accuracy sebesar 95,63%, precision sebesar 95,74%, recall sebesar 95,63%, dan F1-score sebesar 95,51%. Secara keseluruhan, hasil tersebut menunjukkan bahwa algoritma SVM memiliki kemampuan yang sangat baik dalam mengklasifikasikan status stunting pada balita secara akurat dan konsisten.</p>		
7.	<p><i>a Implementation of the Support Vector Machine Algorithm for Classifying Stunting Status in Toddlers</i></p>	<p>Peneliti: Abdul Jalil, Ahmad Homaidi, Zachol Fatah Tahun: 2024</p>	<p>Penelitian ini bertujuan untuk mengimplementasikan algoritma Support Vector Machine (SVM) dengan kernel linear pada dataset balita guna</p>	<p>Hasil pengujian menunjukkan bahwa algoritma SVM dengan kernel linear mampu mencapai tingkat akurasi sebesar 82%, nilai</p>	<p>Penelitian selanjutnya diharapkan dapat mengeksplorasi penggunaan kernel non-linear pada SVM dan menambah faktor-faktor determinan</p>	<p>Jalil, dkk. mengimplementasikan algoritma SVM dengan kernel linear, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi</p>

		Publikasi: G-Tech: Jurnal Teknologi Terapan	mengklasifikasikan status stunting secara tepat	presisi sebesar 80%, dan recall sebesar 86% dalam proses klasifikasi kasus stunting pada balita. Temuan ini menunjukkan bahwa metode SVM memiliki kinerja yang baik dan efektif dalam mengidentifikasi status stunting.	stunting lainnya yang lebih spesifik.	dengan melakukan komparasi metode Support Vector Machine (dengan berbagai kernel) dan Random Forest.
8.	<i>Classification of Public Sentiment Related to Stunting in Indonesia Using BERT and SVM</i>	Peneliti: Yampi R Kacsmetan Tahun: 2025 Publikasi: JBASE- <i>Journal of Business and Audit Information Systems</i>	Penelitian ini bertujuan untuk menilai dan membandingkan kinerja model BERT (<i>Bidirectional Encoder Representations from Transformers</i>) dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral.	Hasil pengujian menggunakan algoritma SVM dengan berbagai jenis kernel menunjukkan adanya perbedaan tingkat kinerja dalam proses klasifikasi sentimen, yang mengindikasikan bahwa pemilihan kernel berpengaruh terhadap hasil klasifikasi yang diperoleh.	Saran untuk penelitian mendatang adalah memperkaya fitur data dan mempertimbangkan faktor sosial budaya dalam analisis sentimen terkait stunting di Indonesia.	Kacsmetan melakukan klasifikasi sentimen stunting menggunakan BERT dan SVM, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi status stunting pada balita dengan melakukan komparasi metode Support Vector Machine dan Random Forest.
9.	<i>Algoritma Random Forest, Decision Tree dan XGboost Untuk Klasifikasi Stunting Pada Balita</i>	Peneliti: Dhika Malita, Kartika Imam Santoso, Andri Triyono, Eko	Penelitian ini bertujuan untuk melakukan perbandingan antara tiga algoritma, yaitu Random Forest, Decision Tree, dan	Hasil penelitian menunjukkan bahwa penerapan algoritma <i>machine learning</i> efektif dalam	Disarankan bagi peneliti selanjutnya untuk melakukan validasi model pada dataset yang berbeda	Malita, dkk. membandingkan algoritma Random Forest, Decision Tree, dan XGBoost.

		<p>Supriyadi, Agus Susilo Nugrobo, Edi Widodo</p> <p>Tahun: 2025</p> <p>Publikasi: <i>Jurnal Transformatika</i></p>	<p>XGBoost, serta menentukan algoritma yang paling optimal dalam menganalisis data stunting pada bayi.</p>	<p>mengelompokkan status stunting pada balita dengan memanfaatkan data demografis dan pengukuran antropometri dasar sebagai variabel utama. Di antara algoritma yang dituji, Random Forest menunjukkan kinerja terbaik dengan tingkat akurasi sebesar 99,72%, diikuti oleh XGBoost dengan akurasi sebesar 99,58%, dan Decision Tree dengan akurasi sebesar 98,87%.</p>	<p>sedangkan dalam penelitian tesis ini akan melakukan klasifikasi dengan melakukan komparasi metode Support Vector Machine dan Random Forest.</p>	
10.	<p><i>Klasifikasi Status Gizi Pada Anak Balita Menggunakan Metode Random Forest (Studi Kasus: Puskesmas Kambaniru)</i></p>	<p>Peneliti: Jeaneh Ngguna, Arini Aha Pekuwali, Leonard Marthen Doni Ratu</p> <p>Tahun: 2025</p> <p>Publikasi: <i>Jurnal Inovatif</i></p>	<p>Penelitian ini bertujuan untuk mengembangkan dan menerapkan model klasifikasi berbasis <i>machine learning</i> dengan menggunakan algoritma Random Forest dalam mengidentifikasi status gizi balita, baik kategori gizi baik, gizi kurang, maupun gizi buruk.</p>	<p>Hasil penelitian menunjukkan bahwa model Random Forest yang digunakan dalam klasifikasi status gizi balita menghasilkan beberapa temuan penting. Model tersebut mampu mencapai tingkat akurasi sebesar 61%, yang menunjukkan</p>	<p>Penelitian selanjutnya perlu mempertimbangkan perbaikan kualitas dataset dan penambahan fitur-fitur klinis yang lebih representatif guna meningkatkan akurasi klasifikasi yang dihasilkan.</p>	<p>Ngguna, dkk. menerapkan model klasifikasi status gizi menggunakan algoritma Random Forest, sedangkan dalam penelitian tesis ini akan melakukan klasifikasi dengan melakukan komparasi metode Support Vector Machine dan Random Forest.</p>

				bahwa sebanyak 61% data pengujian dapat diklasifikasikan dengan benar.		
--	--	--	--	--	--	--



2.3. Landasan Teori

2.3.1. Stunting

Stunting adalah kondisi gagal tumbuh pada anak yang ditandai dengan tinggi badan yang tidak sesuai dengan usia, atau berada di bawah rata-rata dibandingkan dengan anak normal seusianya. Kondisi ini terjadi karena kekurangan gizi kronis dan infeksi yang berulang dalam waktu yang lama, terutama pada periode kritis 1000 Hari Pertama Kehidupan (HPK), yaitu sejak janin dalam kandungan hingga anak berusia dua tahun (Rahmawati & Agustin, 2020). Untuk mengidentifikasi stunting, pemeriksaan antropometri dapat dilakukan dengan cara yang sederhana yaitu dengan mengukur tinggi badan anak dan membandingkannya dengan nilai z-score menggunakan grafik pertumbuhan global yang dikembangkan oleh WHO (Candra, 2020).

Stunting tidak hanya mempengaruhi pertumbuhan fisik tetapi juga perkembangan otak. Dampak jangka panjangnya antara lain keterbelakangan mental, penurunan kemampuan kognitif dan penerimaan pembelajaran, penurunan produktivitas, dan peningkatan risiko penyakit metabolik seperti diabetes, hipertensi, dan obesitas di kemudian hari (Kurniawan et al., 2022). Kekurangan gizi pada masa pertumbuhan anak usia dini akan menghambat perkembangan fisik secara keseluruhan dan meningkatkan kerentanan terhadap berbagai penyakit.

Stunting merupakan masalah gizi serius yang masih menjadi isu penting di bidang kesehatan, terutama di negara miskin dan berkembang.

Kondisi ini merupakan bentuk gangguan tumbuh kembang akibat akumulasi dari kekurangan gizi yang berlangsung dalam jangka waktu yang cukup lama, mulai dari masa kehamilan hingga usia 24 bulan (Handika, 2020). Mengingat dampaknya yang signifikan terhadap kualitas hidup individu dan produktivitas masyarakat secara keseluruhan, maka penanganan masalah stunting memerlukan pendekatan yang komprehensif dan berkelanjutan dengan melibatkan berbagai sektor terkait. Merujuk pada *World Health Organization* (WHO) yang dikutip oleh Okselia (2024), standar pertumbuhan anak diukur berdasarkan indeks panjang badan terhadap umur (PB/U) atau tinggi badan terhadap umur (TB/U) dengan ambang batas (*z-score*) kurang dari -2 SD. Penentuan status stunting atau balita pendek dapat dilakukan melalui pengukuran panjang atau tinggi badan anak dibandingkan dengan standar, di mana hasil pengukuran berada dalam rentang normal. Untuk ciri-ciri stunting menurut Esha et al. (2023) sebagai berikut:

- a. Bertubuh pendek, anak-anak dengan stunting biasanya memiliki tinggi badan yang lebih rendah dari teman sebayanya. Ini adalah hasil dari pertumbuhan linier yang terhambat karena kekurangan nutrisi jangka panjang.
- b. Berat badan kurang dari standar
- c. Perkembangan fisik yang tertunda
- d. Gangguan fungsi kognitif
- e. Berkurangnya tingkat energi dan aktivitas
- f. Pubertas yang tertunda

- g. Penampilan fisik yang tampak lebih muda dari usia sebenarnya

2.3.1.1. Faktor-Faktor Penyebab Stunting

Merujuk pendapat Novita et al. (2020), terdapat beberapa faktor penyebab stunting diantaranya sebagai berikut:

- a. Gizi buruk dalam aspek komunikasi

Gizi buruk terkait dengan keterbatasan pemahaman atau penyampaian informasi yang tidak tepat tentang gizi. Informasi gizi yang tidak dikomunikasikan dengan baik membuat masyarakat tidak menyadari pentingnya asupan gizi yang cukup untuk perkembangan anak, sehingga meningkatkan angka stunting.

- b. Minimnya pengetahuan ibu tentang kesehatan dan gizi

Kurangnya pemahaman ibu tentang kesehatan dan gizi berkontribusi pada terjadinya stunting. Ketidapahaman tentang makanan bergizi dan kebutuhan nutrisi anak menyebabkan pemberian asupan yang tidak optimal, mempengaruhi pola makan anak, dan meningkatkan risiko defisiensi nutrisi. Peningkatan pengetahuan ibu melalui edukasi dapat menjadi solusi untuk mengurangi risiko stunting.

- c. Keterbatasan layanan kesehatan

Keterbatasan layanan kesehatan menghambat penyebaran informasi tentang perawatan dan gizi yang diperlukan untuk pertumbuhan anak. Komunikasi antara tenaga kesehatan dan masyarakat menjadi kunci dalam penyampaian informasi kesehatan yang tepat untuk pencegahan

stunting. Peningkatan akses dan kualitas layanan kesehatan dapat membantu mengurangi risiko stunting.

d. Keterbatasan akses makanan bergizi dari segi komunikasi

Keterbatasan akses makanan bergizi dapat disebabkan oleh kesenjangan informasi atau pemahaman yang rendah di masyarakat. Ketidaksetaraan dalam distribusi sumber daya, ekonomi, dan infrastruktur juga berkontribusi. Komunikasi efektif perlu berfokus pada peningkatan pemahaman masyarakat tentang haknya atas akses makanan bergizi. Edukasi dan advokasi penting untuk meningkatkan kesadaran masyarakat tentang pentingnya gizi dan upaya bersama mengatasi stunting.

e. Keterbatasan akses air bersih dan sanitasi

Ketidakmampuan menyampaikan informasi tentang pentingnya sanitasi dan pengelolaan air bersih dapat meningkatkan risiko stunting. Komunikasi yang tidak memadai juga mempengaruhi partisipasi masyarakat dalam program pencegahan stunting. Pemerintah perlu mengkomunikasikan informasi tentang air bersih, sanitasi, dan dampaknya terhadap pertumbuhan anak secara efektif. Pencegahan stunting dapat berhasil jika faktor ini mendapat perhatian yang memadai dari pemerintah.

Menurut hasil penelitian Mashar et al. (2021), beberapa faktor yang meningkatkan risiko terjadinya stunting pada anak antara lain meliputi pola pengasuhan orang tua, kelengkapan imunisasi, kondisi

higiene dan sanitasi, riwayat penyakit diare, paparan asap rokok, dan riwayat Infeksi Saluran Pernapasan Atas (ISPA). Kemenkes RI Tahun 2016 dikutip Yuwita (2024) mengidentifikasi dampak stunting dalam dua kategori. Dampak jangka pendek termasuk gangguan pada perkembangan otak, penurunan tingkat kecerdasan, hambatan pertumbuhan fisik, dan gangguan metabolisme tubuh. Sedangkan dampak jangka panjang mencakup kemunduran kemampuan kognitif dan prestasi belajar, penurunan daya tahan tubuh yang menyebabkan kerentanan terhadap penyakit, serta peningkatan risiko berbagai penyakit kronis seperti diabetes, obesitas, penyakit kardiovaskular, kanker, stroke, dan disabilitas pada usia lanjut. Lebih jauh, stunting juga mengakibatkan rendahnya kualitas kerja yang tidak kompetitif yang berdampak pada penurunan produktivitas ekonomi.

Akses yang terbatas terhadap pelayanan kesehatan, termasuk sanitasi dan air bersih, merupakan salah satu faktor yang sangat signifikan dalam kasus stunting (Komalasari et al., 2020). Susilawati & Ginting (2023) mengklasifikasikan faktor-faktor yang mempengaruhi stunting menjadi faktor langsung dan tidak langsung. Faktor langsung mencakup asupan nutrisi makanan dan status kesehatan. Sementara itu, faktor tidak langsung meliputi akses terhadap pelayanan kesehatan, kondisi lingkungan rumah tangga, tingkat pendidikan, serta kondisi perekonomian yang rendah yang juga berkontribusi sebagai pemicu stunting.

2.3.1.2. Dampak Stunting

Berdasarkan pendapat Choliq et al. (2020), dampak stunting dapat diklasifikasikan menjadi dampak jangka pendek dan jangka panjang.

Dampak Jangka Pendek:

- a. Peningkatan kejadian kesakitan dan kematian
- b. Perkembangan kognitif, motorik, dan verbal anak yang tidak optimal
- c. Gangguan perkembangan otak

Dampak Jangka Panjang:

- a. Postur tubuh yang tidak ideal saat dewasa (lebih pendek dibandingkan standar)
- b. Peningkatan risiko obesitas dan berbagai penyakit lainnya
- c. Penurunan kesehatan reproduksi
- d. Kapasitas belajar dan performa akademik yang suboptimal
- e. Produktivitas dan kapasitas kerja yang tidak maksimal

Sebagai tambahan, Widiastuti (2019) dikutip Afdira (2023) menguraikan dampak stunting secara lebih terperinci:

Dampak jangka pendek:

- a. Gangguan perkembangan otak
- b. Penurunan kemampuan berpikir pada anak
- c. Gangguan pertumbuhan fisik
- d. Gangguan sistem metabolisme tubuh

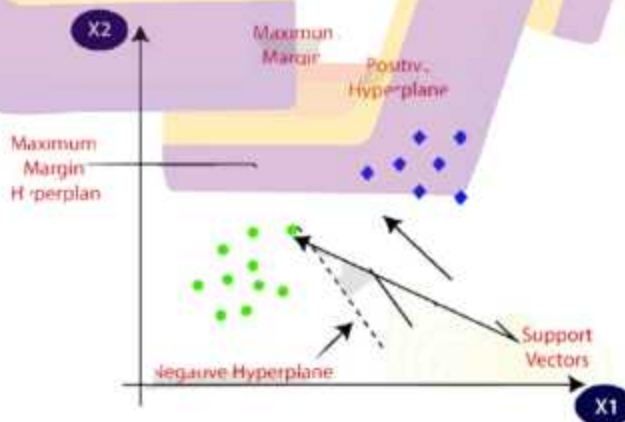
Dampak jangka panjang:

- a. Penurunan kemampuan kognitif dan prestasi belajar

- b. Menurunnya imunitas atau kekebalan tubuh
- c. Peningkatan risiko penyakit kronis seperti diabetes, obesitas, penyakit jantung dan pembuluh darah, kanker, stroke, serta disabilitas pada usia lanjut

2.3.2. Algoritma SVM

Prinsip kerja utama Algoritma SVM adalah mengidentifikasi hyperplane terbaik yang berfungsi sebagai pembatas antara dua kelas dalam ruang input (Rasyida et al., 2020). Hyperplane ini dapat berbentuk garis linear maupun non-linear, bergantung pada karakteristik dataset. Hyperplane optimal diperoleh berdasarkan margin, yaitu jarak minimal dari titik-titik data pada tiap kelas. Titik-titik data yang tepat berada pada margin tersebut disebut sebagai support vector, sebagaimana ditampilkan pada Gambar berikut.



Gambar 2.1. Struktur SVM Berdasarkan Garis Hyperplane

Dewi et al. (2023) menjelaskan bahwa SVM dapat mengimplementasikan klasifikasi dengan memanfaatkan Persamaan 2.1 dengan rumus sebagai berikut.

$$f(x) = \text{sign}(w \cdot x + b) \quad 2.1$$

Keterangan:

$f(x)$ = Fungsi prediksi

w = Vektor normal hyperplane

x = Vektor fitur input

b = Bias atau intercept

Terdapat 3 jenis kernel yang diimplementasikan dalam Algoritma SVM untuk menyelesaikan permasalahan klasifikasi: Kernel Linear, Polynomial, dan Radial Basis Function (RBF). Fungsi dan rumus dari ketiga kernel tersebut dijelaskan oleh Rabbani et al. (2023) sebagai berikut:

a. Kernel Linear

Kernel linear digunakan untuk menghitung produk titik dari dua vektor input dalam ruang 17 asli tanpa melakukan transformasi ke ruang fitur dimensi lebih tinggi. Perhitungan ini mengacu pada Persamaan 2.2 sebagai berikut.

$$K(x, y) = (x, y) \quad 2.2$$

Keterangan :

$K(x, y)$ = Fungsi kernel yang mengukur kesamaan (similarity) antara dua vektor

(x, y) = Notasi untuk **dot product** (produk titik) antara vektor x dan y

b. Kernel Polynomial

Fungsi kernel polynomial adalah mengukur relasi polynomial antara dua vektor input dalam ruang asli. Pengukuran ini mengimplementasikan Persamaan 2.3 sebagai berikut.

$$K(x, y) = (x \cdot y + c)^d \quad 2.3$$

Keterangan :

$K(x, y)$ = Fungsi kernel polynomial

$x \cdot y$ = Dot product antara vektor x dan y

c = Parameter bias/konstanta

d = Derajat polynomial

c. Kernel Radial Basis Function (RBF)

Kernel RBF mengadopsi fungsi Gaussian, yang juga disebut fungsi radial, untuk mengkalkulasi kesamaan antara dua vektor input dalam ruang fitur. Kalkulasi ini dapat dilakukan melalui Persamaan 2.4 sebagai berikut.

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad 2.4$$

Keterangan :

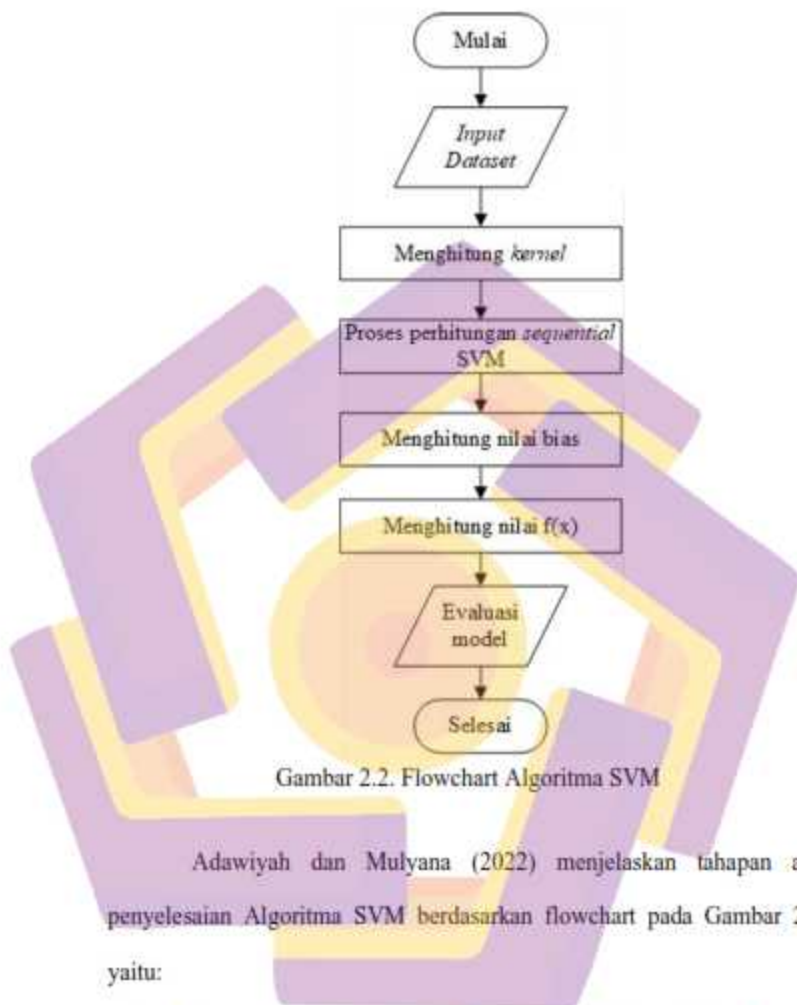
K = Fungsi Kernel (mengukur kesamaan antara dua data)

x, x' = Dua vektor/data yang akan dibandingkan

\exp = Fungsi eksponensial

γ = Parameter yang mengontrol lebar pengaruh titik

$\|x - x'\|^2$ = Kuadrat jarak Euclidean antara dua vektor



Gambar 2.2. Flowchart Algoritma SVM

Adawiyah dan Mulyana (2022) menjelaskan tahapan alur penyelesaian Algoritma SVM berdasarkan flowchart pada Gambar 2.2, yaitu:

- Input dataset yang sudah dinormalisasi, terdiri dari data uji dan data latih.
- Kalkulasi kernel SVM.
- Perhitungan dataset menggunakan metode sequential SVM.
- Penghitungan nilai bias.

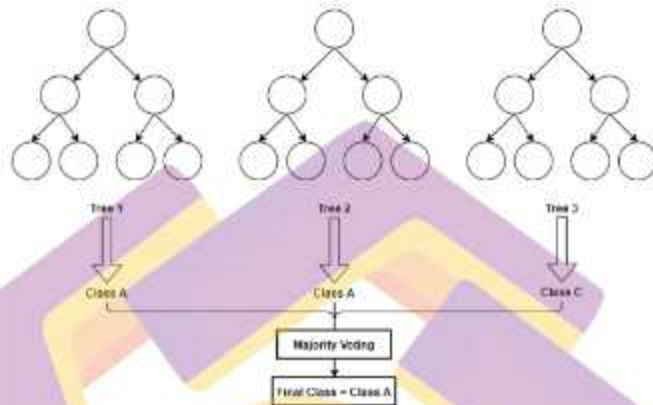
- e. Kalkulasi nilai $f(x)$ berdasarkan data uji dan data latih.
- f. Perhitungan nilai akurasi.

2.3.3. *Random Forest*

Random Forest merupakan sebuah teknik pembelajaran *ensemble* dalam analisis data dan modern. Teknik ini menggunakan beberapa pohon keputusan yang dibangun dari berbagai subset data yang berbeda-beda dan kemudian melakukan proses *voting* atau pengambilan suara terbanyak terhadap hasil prediksi dari masing-masing pohon keputusan tersebut terhadap untuk menghasilkan *output* final dari pohon *random* (Sun et al., 2020). Dengan kata lain, *Random Forest* dari sejumlah besar pohon keputusan yang bekerja secara simultan. Kumpulan atau himpunan pohon keputusan yang beragam ini kemudian digunakan secara kolektif untuk mengklasifikasikan data ke dalam suatu kelas tertentu, sehingga menghasilkan prediksi yang lebih akurat dan robust dibandingkan dengan penggunaan pohon keputusan tunggal.

Di antara berbagai metode CART (Classification and Regression Tree), *Random forest* hadir sebagai teknik data mining yang tidak membutuhkan asumsi. Metode ini mengembangkan konsep pohon keputusan menjadi struktur yang lebih kompleks. Modelnya terdiri dari kumpulan banyak pohon yang secara metaforis membentuk hutan dalam analisis data. Pembentukan model memanfaatkan metode aggregating bootstrap (bagging) dan proses pemilihan fitur acak. Berikut gambar yang ditampilkan memperlihatkan representasi pohon keputusan dan cara

Random forest menghasilkan kesimpulan dari kumpulan pohon tersebut (Iman & Wijayanto, 2021).



Gambar 2.3 Ilustrasi Metode *Random Forest*

2.3.3.1. Algoritma *Random Forest*

Lestari & Sirodj (2022) mengemukakan algoritma *Random Forest* untuk proses klasifikasi sebagai berikut:

1. Lakukan bootstrap sampling dengan mengambil N sampel acak dari data training secara replacement.
2. Gunakan sampel bootstrap (BS_j) untuk konstruksi pohon ke- j ($j = 1, 2, \dots, J$), dengan pertumbuhan pohon dibiarkan maksimal tanpa pemangkasan. Bangun pohon klasifikasi menggunakan teknik pemilihan fitur acak, ulangi langkah-langkah berikut pada setiap simpul terminal sampai ukuran simpul minimum tercapai.
 - (i) Pilih secara acak m variabel dari p variabel prediktor di mana $m < p$, dengan menggunakan persamaan (13), (14), dan (15).

- (ii) Mengidentifikasi variabel/pemisah terbaik dari semua pilihan fitur mtry sebagai penyekat.
 - (iii) Membagi node menjadi dua sub-node berdasarkan solusi split dari langkah (ii).
3. Ulangi langkah 1-3 sebanyak j kali untuk menciptakan hutan dengan j pohon. Setiap pohon klasifikasi memberikan satu hasil prediksi, menghasilkan j prediksi.
 4. Tentukan prediksi final dengan mengagregasi hasil dari setiap pohon menggunakan majority vote.

Random forest adalah pengklasifikasi yang terdiri dari sekumpulan pengklasifikasi terstruktur pohon $\{h(x, \Theta_k), k = 1, \dots\}$ di mana $\{\Theta_k\}$ adalah distribusi vektor acak independen, dan setiap pohon menyumbangkan satu suara untuk kelas yang paling umum pada input x . Random forest adalah kombinasi pohon prediktif di mana setiap pohon bergantung pada nilai vektor acak yang diambil secara independen dengan distribusi yang sama untuk semua pohon di dalam hutan. Random forest adalah kombinasi dari pohon-pohon prediktif dimana setiap pohon bergantung pada nilai vektor acak yang disampel secara independen dengan distribusi yang identik untuk semua pohon di dalam hutan (Aryani & Wijayanto, 2021).

Guna menghasilkan keputusan akhir, RF melakukan klasifikasi dengan metode pengelompokan berdasarkan kemunculan mayoritas dari berbagai pohon. Pada algoritma RF, dataset pelatihan didefinisikan

sebagai $S = \{(x_i, y_j), i = 1, 2, \dots, N; j = 1, 2, \dots, M\}$ dengan x sebagai sampel dan y sebagai variabel fitur S . Jumlah total sampel pelatihan adalah N , dan setiap sampel memiliki M variabel fitur (Religia et al., 2021).

Proses klasifikasi, algoritma RF dibangun melalui tiga langkah:

- 1) mengumpulkan k subset pelatihan,
- 2) membangun setiap model pohon keputusan, dan
- 3) menggabungkan k pohon ke dalam model RF. Algoritma RF memiliki hasil kinerja yang baik dan waktu pemrosesan yang cepat, sehingga dapat diterapkan pada data yang tidak seimbang dalam jumlah besar.

2.3.3.2. Parameter *Random Forest*

Random forest hadir dengan beberapa parameter yang dapat digunakan sesuai dengan kebutuhan pengguna. Parameter-parameter ini menyediakan berbagai pilihan yang, ketika ditentukan, akan mengatur proses pembelajaran sesuai dengan kebutuhan pengguna. Berikut adalah parameter-parameter penting dalam Random forest.

a. *n_estimator*

Parameter *n_estimator* mengatur jumlah pohon di hutan. Nilai ini dapat disesuaikan berdasarkan kebutuhan pengguna.

b. Kriteria

Kriteria digunakan sebagai ukuran kualitas pembagian. Ada dua kriteria yang didukung, yaitu "gini" untuk mengukur ketidakmurnian gini, dan "entropy" untuk menghitung information gain.

c. *Max_depth*

`Max_depth` menetapkan kedalaman maksimum dalam pohon. Jika tidak ditentukan, node akan terus tumbuh hingga semua daun murni atau hingga daun mengandung kurang dari `min_samples_splits` sampel.

d. `Min_samples_split`

`Min_samples_split` adalah jumlah minimum sampel yang dibutuhkan untuk membagi node internal.

e. `min_samples_leaf`

Parameter ini menentukan jumlah minimum sampel yang diperlukan pada node daun. Pemisahan pada kedalaman berapa pun hanya dipertimbangkan jika `min_samples_leaf` terpenuhi pada cabang kiri dan kanan.

f. `Max_features`

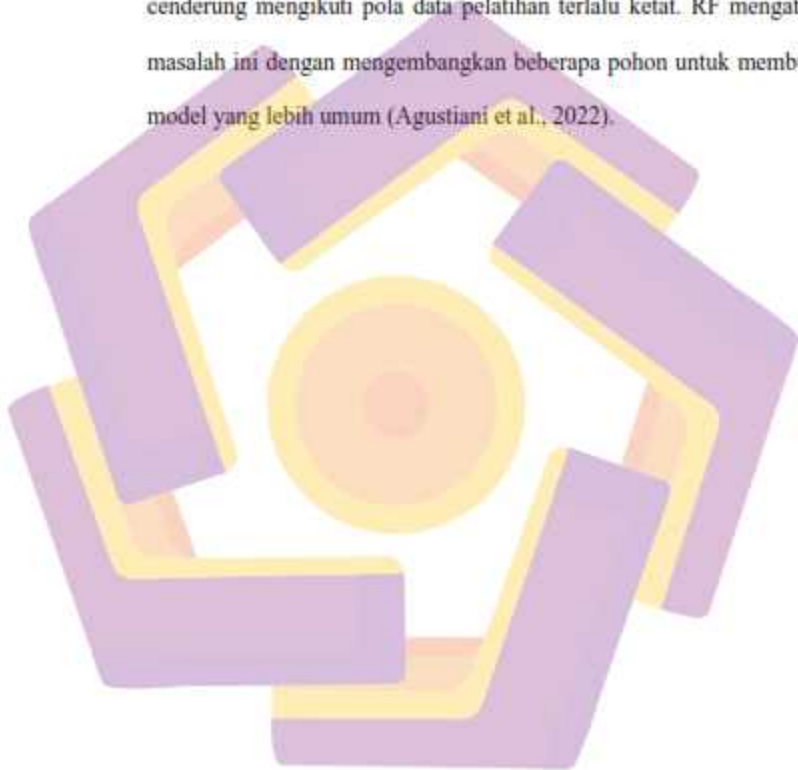
`Max_features` mengindikasikan jumlah fitur yang dipertimbangkan saat mencari pemisahan terbaik. Nilainya termasuk `auto`, `log2`, dan `sqrt`. Pencarian pemisahan akan terus berlanjut sampai setidaknya satu partisi yang valid dari sampel node ditemukan (Yolanda, 2024).

2.3.3.3. Kelebihan *Random Forest*

Beberapa kelebihan algoritme *Random Forest* antara lain:

- a. Tingkat kesalahan yang relatif rendah
- b. Kemampuan klasifikasi yang sangat akurat
- c. Kapasitas untuk memproses data pelatihan dalam jumlah besar

- d. Pendekatan yang efektif untuk estimasi nilai yang hilang dengan menggunakan sampel pelatihan dan subset variabel input yang dipilih secara acak di setiap node.
- e. Mengatasi masalah kritis overfitting pada pohon keputusan yang cenderung mengikuti pola data pelatihan terlalu ketat. RF mengatasi masalah ini dengan mengembangkan beberapa pohon untuk membuat model yang lebih umum (Agustiani et al., 2022).



BAB III

METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

Penelitian ini merupakan jenis penelitian eksperimen dengan pendekatan kuantitatif. Penelitian eksperimen dipilih karena peneliti secara aktif menguji dan membandingkan kinerja dari dua algoritma machine learning, yaitu Support Vector Machine (SVM) dan Random Forest, dalam mengklasifikasi kasus stunting. Sifat penelitian adalah komparatif-analitis karena tujuan utamanya adalah membandingkan tingkat akurasi, presisi, recall, dan F1-score dari kedua algoritma tersebut untuk menentukan mana yang lebih unggul.

Pendekatan kuantitatif digunakan karena seluruh data yang dikumpulkan berupa angka dan diolah menggunakan teknik statistik serta data mining. Hasil analisis disajikan dalam bentuk numerik serta divisualisasikan menggunakan tabel, diagram, dan confusion matrix untuk memudahkan interpretasi dan penarikan kesimpulan secara objektif.

3.2. Metode Pengumpulan Data

Metode pengumpulan data adalah cara yang digunakan peneliti untuk mengumpulkan informasi yang dibutuhkan dalam penelitian. Dalam penelitian ini, ada dua cara utama yang digunakan untuk mengumpulkan data:

a. Studi Literatur

Studi literatur dilakukan dengan membaca dan mempelajari berbagai bahan bacaan yang berhubungan dengan topik penelitian. Bahan bacaan ini berupa jurnal ilmiah, buku referensi, artikel akademik, dan publikasi terpercaya lainnya. Fokus studi literatur meliputi pemahaman mengenai stunting dan faktor penyebabnya, prevalensi stunting di Indonesia dan dunia, teknik pengolahan data menggunakan data mining, metode klasifikasi dalam machine learning, tahapan pengolahan data, prinsip kerja algoritma Support Vector Machine (SVM) dan Random Forest, serta teknik evaluasi hasil klasifikasi menggunakan confusion matrix dan metrik-metrik terkait.

b. Data Sekunder

Data yang digunakan dalam penelitian ini terdiri dari dua sumber utama, yaitu data sekunder dari instansi kesehatan dan dataset publik untuk kebutuhan validasi eksternal. Data sekunder utama diperoleh dari Dinas Kesehatan Kota Banjarmasin tahun 2024 yang terdiri dari 2.231 record data balita. Dataset ini mencakup variabel pertumbuhan antropometri seperti Tinggi Badan menurut Umur (TB/U), Berat Badan menurut Umur (BB/U), dan Berat Badan menurut Tinggi Badan (BB/TB).

Selanjutnya, untuk menguji kemampuan generalisasi model melalui tahapan validasi eksternal, digunakan dataset publik berjudul "*Dataset Stunting and Nutritional Status of Toddler from Jeneponto Regency, South Sulawesi, Indonesia*". Dataset ini dipublikasikan pada 17 April 2025

(Versi 2) oleh Asmaul Husna RS, Syahrani Lonang, dan Ahmad Fatoni Dwi Putratolong . Penggunaan data dari wilayah Kabupaten Jeneponto ini bertujuan untuk memastikan bahwa model klasifikasi yang telah dibangun memiliki performa yang stabil saat diimplementasikan pada karakteristik data yang berbeda di luar wilayah penelitian utama.

3.3. Metode Analisis Data

Penelitian ini menggunakan dua algoritma utama yaitu Support Vector Machine (SVM) dan Random Forest untuk dibandingkan kemampuannya dalam mengklasifikasi stunting pada anak-anak di Kota Banjarmasin. Evaluasi yang menghasilkan nilai accuracy (akurasi), precision (presisi), recall (sensitivitas), dan F1-score sebagai dasar perbandingan kinerja model. Tahapan analisis data mengikuti kerangka CRISP-DM (Cross-Industry Standard Process for Data Mining) yang terdiri dari:

1. Pemahaman Bisnis/Pemahaman Penelitian

Tahap awal ini difokuskan pada pemahaman tujuan penelitian yang ingin dicapai. Tujuan utama adalah membandingkan performa algoritma SVM dan Random Forest dalam klasifikasi kasus stunting di Kota Banjarmasin. Peneliti berusaha mengidentifikasi algoritma mana yang memberikan akurasi dan keandalan klasifikasi yang lebih tinggi berdasarkan berbagai faktor yang mempengaruhi stunting.

2. Pemahaman Data

Tahap ini meliputi pengumpulan data dan eksplorasi karakteristik data yang akan digunakan. Data yang dikumpulkan berasal dari Dinas Kesehatan Kota Banjarmasin yang berisi informasi kesehatan anak-anak meliputi variabel antropometri (tinggi badan/umur/TB/U, berat badan/umur/BB/U, berat badan/tinggi badan/BB/TB), lingkaran kepala, lingkaran lengan, riwayat imunisasi, pola makan, kondisi ekonomi keluarga, pendidikan orang tua, dan variabel pendukung lainnya. Eksplorasi data dilakukan untuk memahami pola dan hubungan antar variabel yang relevan untuk proses klasifikasi

3. Persiapan Data

Tahap persiapan data dilakukan untuk memastikan data siap digunakan dalam pembuatan model klasifikasi melalui beberapa proses:

a. Pembersihan Data

Proses pembersihan data dilakukan untuk mengatasi masalah data yang dapat mengganggu analisis, antara lain missing values (data kosong), duplikasi data, outlier (data ekstrem), dan data tidak valid. Penanganan masalah data ini diperlukan untuk memastikan keakuratan hasil analisis.

b. Normalisasi Data

Normalisasi data merupakan proses penyeragaman nilai-nilai data yang memiliki skala berbeda. Teknik min-max normalization digunakan untuk mengubah semua nilai menjadi rentang antara 0 dan 1. Proses ini penting untuk menyeimbangkan pengaruh semua variabel dalam proses

pembelajaran algoritma, khususnya untuk algoritma SVM yang sensitif terhadap perbedaan skala.

c. Pembagian Data

Data yang telah dibersihkan dibagi menjadi dua subset yaitu training set (data latih) dan testing set (data uji). Pembagian data menggunakan dua skenario yaitu rasio 70:30 dan 80:20. Pembagian ini bertujuan untuk melatih model dengan data yang cukup sekaligus menguji performa model dengan data yang belum pernah dilihat sebelumnya.

4. Pemodelan

Tahap pemodelan merupakan proses pembangunan sistem klasifikasi menggunakan algoritma SVM dan Random Forest. Implementasi kedua algoritma dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan platform Google Colab. Model dilatih untuk mengklasifikasi kemungkinan terjadinya stunting berdasarkan berbagai faktor yang telah diidentifikasi sebelumnya.

5. Evaluasi

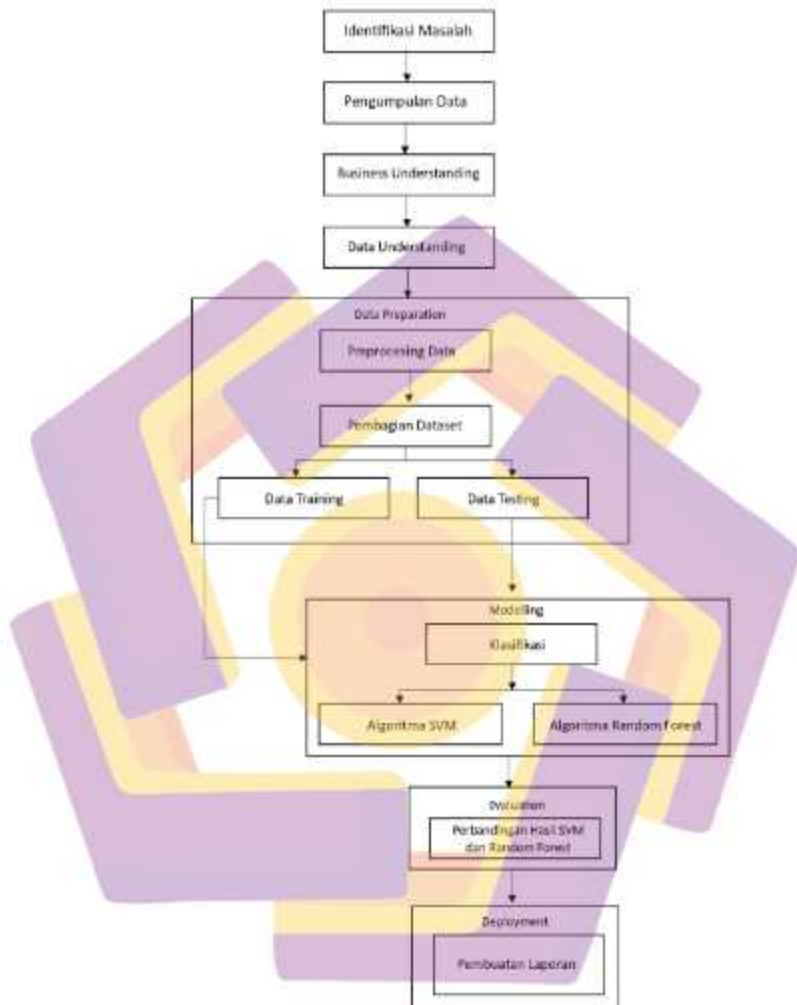
Tahap evaluasi dilakukan untuk menilai kualitas model klasifikasi yang telah dibangun. Penilaian performa model menggunakan confusion matrix yang mengukur berbagai metrik evaluasi antara lain accuracy (akurasi keseluruhan), precision (ketepatan prediksi kasus positif), recall (kemampuan deteksi kasus positif), dan F1-score (keseimbangan antara precision dan recall). Hasil evaluasi ini menjadi dasar penentuan algoritma terbaik untuk mengklasifikasi stunting di Kota Banjarmasin.

3.4. Alur Penelitian

Alur penelitian merupakan rangkaian tahapan yang dilaksanakan secara sistematis dan berurutan untuk mencapai tujuan penelitian. Alur ini berperan penting dalam memastikan penelitian berjalan terarah, terstruktur, dan sesuai dengan perencanaan yang telah ditetapkan. Setiap tahapan memiliki tujuan spesifik dan menghasilkan output yang menjadi input bagi tahapan berikutnya.

Penelitian ini mengadopsi kerangka CRISP-DM (Cross-Industry Standard Process for Data Mining) yang terdiri dari enam tahapan utama, dimulai dari pemahaman masalah penelitian, pemahaman data, persiapan data, pemodelan, evaluasi, hingga penerapan hasil penelitian. Hubungan antar tahapan dan urutan pelaksanaannya ditunjukkan dalam Gambar 3.1.





Gambar 3.1 Alur Penelitian

1. Pemahaman Penelitian

Tahap awal ini difokuskan pada pendefinisian tujuan dan ruang lingkup penelitian. Fokus utama adalah membandingkan kinerja algoritma *Support Vector Machine* (SVM) dan *Random Forest* dalam mengklasifikasi kasus stunting pada balita di Kota Banjarmasin. Tahap ini menghasilkan rumusan masalah yang jelas dan spesifikasi kebutuhan analisis.

2. Pemahaman Data

Tahap ini meliputi akuisisi data dan eksplorasi karakteristik dataset. Data sekunder diperoleh dari Dinas Kesehatan Kota Banjarmasin kemudian dilakukan analisis eksploratori untuk memahami distribusi, korelasi, dan kualitas data. Hasil tahap ini adalah dokumentasi karakteristik data dan identifikasi masalah data yang perlu ditangani.

3. Persiapan Data

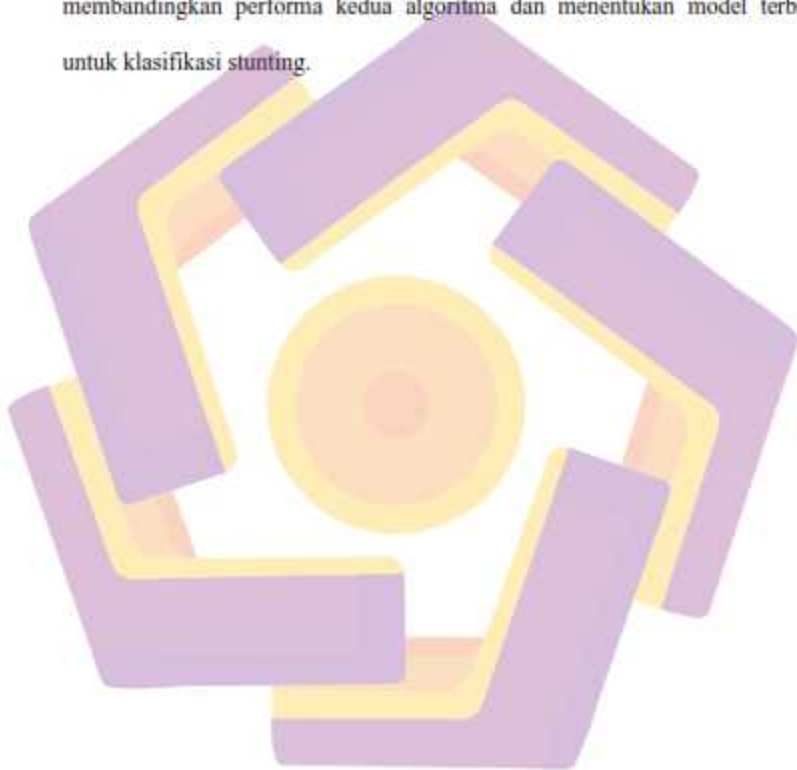
Tahap persiapan data meliputi proses transformasi data mentah menjadi format yang siap untuk pemodelan. Aktivitas yang dilakukan meliputi pembersihan data (*data cleaning*) untuk menangani *missing values* dan outlier, normalisasi data untuk menyamakan skala variabel, serta pembagian dataset menjadi data latih dan data uji dengan rasio 70:30 dan 80:20.

4. Pemodelan

Tahap pemodelan melibatkan implementasi algoritma SVM dan *Random Forest* menggunakan bahasa pemrograman Python pada platform Google Colab. Kedua model dilatih menggunakan data training untuk mempelajari pola hubungan antara variabel prediktor dan status stunting.

5. Evaluasi

Tahap evaluasi dilakukan untuk mengukur kinerja model yang telah dibangun. Evaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-score berdasarkan *confusion matrix*. Hasil evaluasi digunakan untuk membandingkan performa kedua algoritma dan menentukan model terbaik untuk klasifikasi stunting.



BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Bagian ini menyajikan hasil penelitian sesuai dengan alur proses data mining yang meliputi pemahaman data, eksplorasi awal, persiapan data, pemodelan, evaluasi model, serta analisis hasil komparatif antara algoritma Support Vector Machine (SVM) dan Random Forest (RF).

4.2 Deskripsi Data Awal

4.2.1 Struktur dan Karakteristik Dataset

Dataset awal terdiri dari 2.321 baris data dan 17 variabel, yang mencakup informasi antropometrik (berat, tinggi, usia), demografi (jenis kelamin, desa/kelurahan), serta nilai z-score indikator gizi. Struktur ringkas dataset awal ditunjukkan pada Tabel 4.1.

Tabel 4.1. Struktur Dataset Awal

No	Variabel	Tipe Data	Keterangan
1	JK	Kategorik	Jenis Kelamin
2	Tgl Lahir	Date	Tanggal Lahir
3	Status	Kategorik	
4	Desa/kel	Kategorik	Lokasi Desa/Kelurahan
5	Titik Koordinat (Longitude-Latitude)	Numerik	Koordinat Geografis
6	Status Tempat Tinggal	Kategorik	
7	WC di Tempat Tinggal	Kategorik	Fasilitas Sanitasi
8	Usia Saat Ukur	Numerik	Dalam bulan
9	Tanggal Pengukuran	Date	
10	Berat	Numerik	Dalam kg
11	Tinggi	Numerik	Dalam cm
12	BB/U	Numerik	Berat Badan menurut Umur

13	ZS BB/U	Numerik	Z-score BB/U
14	TB/U	Numerik	Tinggi Badan menurut Umur
15	ZS TB/U	Numerik	Z-score TB/U (Indikator Stunting)
16	BB/TB	Numerik	Berat Badan menurut Tinggi Badan
17	ZS BB/TB	Numerik	Z-score BB/TB

4.3 Persiapan Data

Proses persiapan data dilakukan melalui serangkaian tahapan sistematis untuk memastikan kualitas dan konsistensi data sebelum pemodelan. Tahapan ini meliputi penanganan *missing values* dan split data.

4.3.1 Penanganan Missing Values

Penanganan *missing values* pada dataset dilakukan dengan beberapa teknik yang disesuaikan dengan karakteristik data. Untuk variabel numerik digunakan *mean imputation*, sedangkan untuk variabel kategorik diterapkan *mode imputation*. Variabel koordinat geografis diimputasi dengan nilai rata-rata berdasarkan kelompok desa atau kelurahan.

```
Missing Values After Imputation:
ID                0
Tgl_Lahir        0
Status           0
Dosa/No         0
Tinggi_Kordinat (Long/Lat/Local) 0
Status_Tempat_Tinggal 0
No_dk_Tempat_Tinggal 0
Data_Seast_Ukur  0
Tanggal_Hrgukuran 0
Rasus           0
Tinggi          0
BB/U            0
ZS_BB/U         0
TBA            0
ZS_TB/U         0
BB/TB          0
ZS_BB/TB       0
```

Gambar 4.1 Hasil *missing value*

Pada Gambar 4.1 Dataset telah melalui proses imputasi, di mana data-data yang sebelumnya kosong telah diisi dengan nilai-nilai tertentu. Hasilnya, seperti yang terlihat pada gambar, tidak ada lagi *missing values* yang tersisa.

4.3.2 Feature Engineering

Feature Engineering dilakukan untuk menyiapkan variabel agar sesuai untuk proses pemodelan. Tahapan ini meliputi konversi format data, pembersihan nilai tidak valid, pengkodean variabel kategorik, pembentukan variabel target, normalisasi skala fitur, serta penanganan ketidakseimbangan kelas.

Tabel 4.2 *Feature Engineering*

Fitur Asli	Kondisi Awal	Proses FE	Hasil Akhir
Usia Saat Ukur	Teks (X Tahun - Y Bulan - Z Hari)	Konversi usia bulan menggunakan fungsi usia	Usia_Bulan (numerik)
Berat	String, beberapa missing	Konversi ke numerik dengan <code>pd.to_numeric</code>	Berat_num (numerik)
Tinggi	Angka dengan missing	Membersihkan missing (row-drop)	Tinggi (numerik, clean)
JK	Kategori (L/P)	Label encoding {L:1, P:0}	JK_num (0/1)
ZS BB/U	Angka	Dipertahankan	Numerik siap pakai
ZS TB/U	Angka	Dibinary-kan Stunting (1 jika ZS < -2)	Label Stunting (0/1)
ZS BB/TB	Angka	Dipertahankan	Numerik siap pakai

Target Stunting	Tidak tersedia langsung	Dibentuk dari ZS TB/U berdasarkan WHO	Fitur target Stunting
Perbedaan Skala Fitur	Fitur beda skala (bulan, cm, kg)	StandardScaler	Fitur ternormalisasi (mean=0, std=1)

Pada Tabel 4.2 Proses *feature engineering* yang mengubah data mentah menjadi fitur siap model. Usia dikonversi ke *Usia_Bulan*, berat dan tinggi dibersihkan menjadi numerik, *JK* dikodekan, dan label *Stunting* dibentuk dari *ZS TB/U*. Fitur dinormalisasi dengan *StandardScaler* agar skala data seragam untuk pemodelan.

4.3.3 Split Data

Pada penelitian ini akan dilakukan proses split data menjadi 2 bagian yaitu data training dan data testing. Variasi pembagian data akan dilakukan menjadi beberapa bagian antara lain dengan rasio 70% data training dan 30% data testing dan 80% data training dan 20% data testing.

4.3.4 Distribusi Variabel Target

Setelah proses preprocessing dan feature engineering, distribusi variabel target stunting menunjukkan komposisi yang seimbang antara kelas positif dan negatif. Hasil distribusi target dapat dilihat pada Gambar 4.2

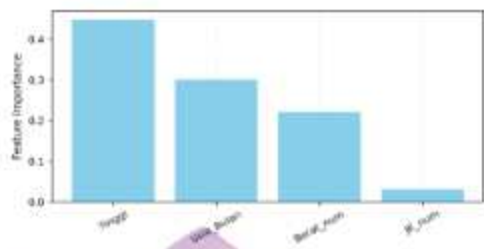


Gambar 4.2 Distribusi data

Pada Gambar 4.2 menunjukkan distribusi data setelah preprocessing dengan total 2.318 sampel. Kelas tidak stunting (0) berjumlah 1.030 sampel (44,43%) dan kelas stunting (1) berjumlah 1.288 sampel (55,57%). Distribusi setelah Smote berjumlah 1.288 sampel kelas stunting dan tidak stunting.

4.3.5 Feature Importance

Analisis *feature importance* dilakukan untuk mengevaluasi kontribusi masing-masing variabel dalam proses klasifikasi status *stunting*. Skor kepentingan ini menunjukkan tingkat signifikansi suatu fitur bagi model dalam membedakan antara kelas *stunting* dan tidak *stunting*. Hasil analisis tingkat kepentingan fitur dapat dilihat pada Gambar 4.3



Gambar 4.3 Feature Importance

Berdasarkan Gambar 4.3, fitur Tinggi dan Usia_Bulan merupakan kontributor paling signifikan dengan skor masing-masing sebesar 0,45 dan 0,30. Sementara itu, fitur Berat_num dan JK_num memiliki pengaruh yang lebih rendah terhadap model dengan skor 0,22 dan 0,03.

4.4 Hasil Tahap Pemodelan (*Modeling*)

Hasil Klasifikasi ini merupakan output yang diperoleh dari pemodelan menggunakan algoritma Support Vector Machine (SVM) dan Random Forest. Pada penelitian ini, model SVM diimplementasikan menggunakan kernel RBF (Radial Basis Function) dengan hyperparameter $C=1.0$ dan $\text{gamma}=\text{'scale'}$. Pemilihan kernel RBF didasarkan pada kemampuannya untuk menangani hubungan non-linear antara fitur-fitur dalam data.

Pada Gambar 4.2 menampilkan contoh hasil klasifikasi model SVM dan Random Forest dengan pembagian data 80:20 terhadap data testing, sedangkan Gambar 4.3 menampilkan hasil untuk pembagian data 70:30.

Untuk mengukur sejauh mana model mampu mengklasifikasikan data dengan tepat, evaluasi dilakukan menggunakan tiga metrik utama, yaitu akurasi, presisi, dan *recall*. Akurasi dihitung berdasarkan persentase total data (baik stunting maupun tidak stunting) yang berhasil diklasifikasikan secara benar dibandingkan dengan seluruh data uji. Sementara itu, presisi digunakan untuk melihat tingkat ketepatan antara data yang diklasifikasikan positif oleh model dengan kondisi aktual di lapangan. Di sisi lain, *recall* atau sensitivitas digunakan untuk mengukur kemampuan model dalam menjaring kembali seluruh data yang secara nyata masuk dalam kategori positif agar tidak terlewatkan dalam hasil klasifikasi.

4.4.1 Perbandingan 80:20

X1_min	X1a_Batas	Berat_x1	Tinggi	True_Label	FF_80_No_SPOTE	FF_80_SPOTE	SFF_80_No_SPOTE	SFF_80_SPOTE
262	-1.221448	0.827117	0.935701	1.131202	0	0	0	0
1815	0.816700	-1.569632	-1.514227	-1.881818	0	0	0	0
784	-1.221448	-0.940242	-0.842865	-1.087858	1	1	1	1
1428	0.816700	0.741380	0.109975	0.112580	1	1	1	1
1854	0.816700	-0.347875	-0.020830	-0.189925	1	1	1	1

Gambar 4.4 Hasil Klasifikasi Model 80:20

Gambar 4.4 menyajikan beberapa baris sampel dari dataset pengujian setelah melalui proses normalisasi, bersama dengan hasil klasifikasi yang dihasilkan oleh model Support Vector Machine (SVM) dan Random Forest dengan perbandingan data training dan testing 80:20.

4.4.2 Perbandingan 70:30

	ix_num	Ukura_Duas	Berat_num	Tinggi	Trax_Label	RF_70_No_SNOTC	RF_70_SNOTC	SWF_70_No_SNOTC	SWF_70_SNOTC
1747	0.825124	-0.304774	-0.306314	-0.257430	0	0	0	1	0
229	0.825124	-0.515282	-0.467487	-0.548864	1	1	1	1	1
1443	-1.211940	0.980500	0.127654	0.327380	0	1	1	1	1
175	0.825124	0.216831	0.470980	0.254170	1	1	1	1	1
2132	-1.211940	-0.107882	0.229146	-0.111380	1	1	1	0	1

Gambar 4.5 Hasil Klasifikasi Model 70:30

Gambar 4.5 menyajikan beberapa baris sampel dari dataset pengujian setelah melalui proses normalisasi, bersama dengan hasil klasifikasi yang dihasilkan oleh model Support Vector Machine (SVM) dan Random Forest dengan perbandingan data training dan testing 70:30.

4.5 Evaluasi

Pada tahap ini dilakukan penilaian akurasi dari kedua metode untuk mengetahui metode terbaik dalam mengidentifikasi stunting menggunakan algoritma Support Vektor Machine dan Random Forest.

4.5.1 Algoritma Support Vector Machine

Actual 0	190	16
Actual 1	28	230
	Pred 0	Pred 1

Gambar 4.6 Confusion Matrix Support Vector Machine 80:20

Gambar 4.6 merupakan hasil dari *confusion matrix* dari *Support Vector Machine* dengan perbandingan data training dan data testing sebesar 80%:20% tanpa SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

MODE: Test - Tanpa SMOTE (split 80:20)				
	precision	recall	f1-score	support
Tidak starting (0)	0.872	0.921	0.896	206
Starting (1)	0.835	0.891	0.913	258
accuracy			0.905	464
macro avg	0.853	0.907	0.904	464
weighted avg	0.857	0.905	0.905	464

Gambar 4.7 Hasil *Confusion Matrix Support Vector Machine 80:20*

Actual 0	196	10
Actual 1	37	221
	Pred 0	Pred 1

Gambar 4.8 *Confusion Matrix Support Vector Machine 80:20*

Gambar 4.8 merupakan hasil dari *confusion matrix* dari *Support Vector Machine* dengan perbandingan data training dan data testing sebesar 80%:20% dengan SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

```

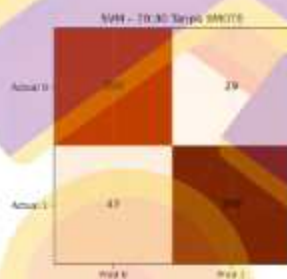
-----
MODEL: SVM - SMOTE (Split 80:20)
-----
              precision    recall  f1-score   support

Tidak Stunting (0)    0.841    0.951    0.893     286
    Stunting (1)      0.957    0.857    0.904     258

   accuracy              0.899     464
  macro avg              0.899    0.904    0.898     464
 weighted avg              0.905    0.899    0.899     464

```

Gambar 4.9 Hasil *Confusion Matrix Support Vector Machine 80:20*



Gambar 4.10 *Confusion Matrix Support Vector Machine 80:20*

Gambar 4.10 merupakan hasil dari *confusion matrix* dari *Support Vector Machine* dengan perbandingan data training dan data testing sebesar 70%:30% tanpa SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

```

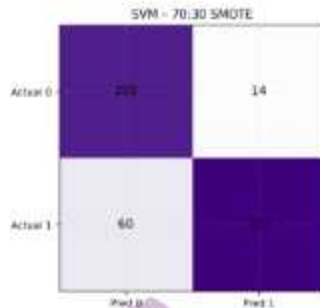
-----
MODEL: SVM - Tanpa SMOTE (Split 70:30)
-----
              precision    recall  f1-score   support

Tidak Stunting (0)    0.856    0.906    0.881     389
    Stunting (1)      0.921    0.879    0.899     387

   accuracy              0.891     696
  macro avg              0.889    0.892    0.890     696
 weighted avg              0.892    0.891    0.891     696

```

Gambar 4.11 Hasil *Confusion Matrix Support Vector Machine 70:30*



Gambar 4.12 *Confusion Matrix Support Vector Machine 70:30*

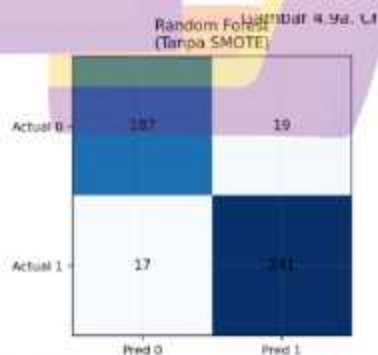
Gambar 4.12 merupakan hasil dari *confusion matrix* dari *Support Vector Machine* dengan perbandingan data training dan data testing sebesar 70%:30% dengan SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

 MODEL: SVM - SMOTE (split 70:30)

	precision	recall	F1-score	support
Tidak Stunting (0)	0.831	0.955	0.893	269
Stunting (1)	0.929	0.845	0.886	187
accuracy			0.894	456
macro avg	0.885	0.900	0.889	456
weighted avg	0.902	0.894	0.894	456

Gambar 4.13 Hasil *Confusion Matrix Support Vector Machine 70:30*

4.5.2 Algoritma Random Forest



Gambar 4.14 *Confusion Matrix Random Forest 80:20*

Gambar 4.15 merupakan hasil dari *confusion matrix* dari *Random Forest* dengan perbandingan data training dan data testing sebesar 80%:20% tanpa SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

```

=====
MODEL: Random Forest - Tanpa SMOTE (Split 80:20)
=====

```

	precision	recall	f1-score	support
Tidak Stunting (0)	0.917	0.908	0.912	206
Stunting (1)	0.927	0.934	0.931	258
accuracy			0.922	464
macro avg	0.922	0.921	0.921	464
weighted avg	0.922	0.922	0.922	464

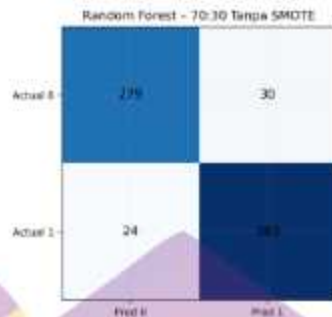
Gambar 4.15 Hasil *Confusion Matrix Random Forest* 80:20

Random Forest (SMOTE)

	Actual 0	Actual 1
Pred 0	170	16
Pred 1	18	26

Gambar 4.16 *Confusion Matrix Random Forest* 80:20

Gambar 4.16 merupakan hasil dari *confusion matrix* dari *Random Forest* dengan perbandingan data training dan data testing sebesar 80%:20% dengan SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

Gambar 4.17 *Confusion Matrix Random Forest 70:30*

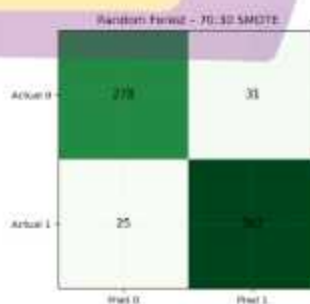
Gambar 4.17 merupakan hasil dari *confusion matrix* dari *Random Forest* dengan perbandingan data training dan data testing sebesar 70%:30% tanpa SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

=====

MODEL: Random Forest - Tanpa SMOTE (Split 70:30)

=====

	precision	recall	f1-score	support
Tidak Stunting (0)	0.921	0.993	0.932	309
Stunting (1)	0.924	0.935	0.931	187
accuracy			0.922	496
macro avg	0.922	0.920	0.921	496
weighted avg	0.922	0.922	0.922	496

Gambar 4.18 Hasil *Confusion Matrix Random Forest 70:30*

Gambar 4.19 *Confusion Matrix Random Forest 70:30*

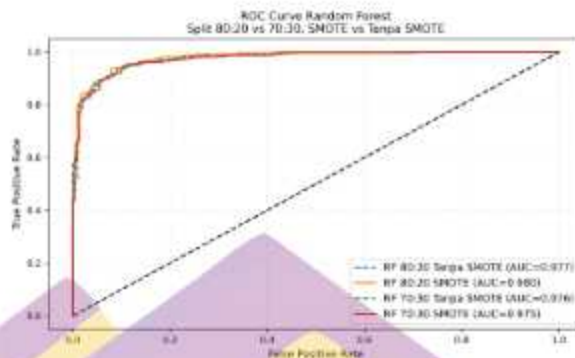
Gambar 4.19 merupakan hasil dari *confusion matrix* dari *Random Forest* dengan perbandingan data training dan data testing sebesar 70%:30% dengan SMOTE. Sehingga didapatkan hasil akurasi, presisi, recall dan F1-score sebagai berikut:

MODEL: Random Forest - SMOTE (Split 70:30)				
	precision	recall	f1-score	support
Tidak Stunting (0)	0.917	0.900	0.908	309
Stunting (1)	0.921	0.935	0.928	387
accuracy			0.920	696
macro avg	0.919	0.918	0.918	696
weighted avg	0.920	0.920	0.919	696

Gambar 4.20 Hasil *Confusion Matrix Random Forest 70:30*

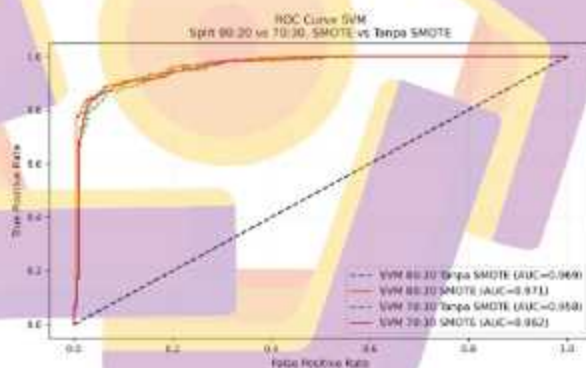
4.5.3. Analisis Kurva ROC dan Cross-Validation

Pada tahap ini dilakukan analisis evaluasi model secara komprehensif melalui dua pendekatan utama, yaitu analisis kurva ROC (Receiver Operating Characteristic) dan validasi silang (cross-validation). Analisis kurva ROC bertujuan untuk mengukur kemampuan diskriminasi model dalam membedakan kedua kelas dengan menganalisis trade-off antara true positive rate dan false positive rate pada berbagai threshold klasifikasi, sementara validasi silang 5-fold dilakukan untuk menguji stabilitas dan reliabilitas model dengan memvalidasi performa pada berbagai subset data. Kedua metode evaluasi ini saling melengkapi dalam memberikan gambaran yang tuntas mengenai konsistensi dan keandalan model klasifikasi yang dikembangkan.



Gambar 4.21 Kurva ROC Random Forest

Gambar 4.21 membandingkan kurva ROC untuk model Random Forest dengan berbagai konfigurasi pembagian data dan penggunaan SMOTE



Gambar 4.22 Kurva ROC SVM

Gambar 4.22 Pada Gambar Kurva ROC SVM menunjukkan performa klasifikasi sangat baik dengan AUC 0.958-0.971. Konfigurasi 80:20 dengan SMOTE mencapai AUC tertinggi (0.971). SMOTE konsisten meningkatkan performa model pada semua rasio data.

```

=== Hasil 5-fold Cross Validation (Mean Accuracy) ===
Model      Setting      CV_Mean_Acc
0      Random Forest Tanpa SMOTE      0.921480
1      Random Forest SMOTE      0.936731
2      SVM Tanpa SMOTE      0.893442
3      SVM SMOTE      0.899860

```

Gambar 4.23 Hasil validasi silang 5-fold

Gambar 4.23 menunjukkan hasil validasi silang 5-fold dimana Random Forest dengan SMOTE mencapai akurasi tertinggi (93.67%), diikuti oleh SVM (89.98%).

4.6 Hasil Validasi Eksternal

Validasi eksternal dilakukan untuk menguji sejauh mana kemampuan model dalam melakukan klasifikasi pada dataset yang belum pernah digunakan sebelumnya. Tahapan ini bertujuan untuk memastikan bahwa model memiliki kemampuan generalisasi yang baik dan tetap akurat saat dihadapkan pada data baru di luar data pelatihan. Hasil evaluasi performa pada data eksternal disajikan dalam Gambar 4.24 dan Gambar 4.25

```

Hasil SVM pada Data Eksternal:
1. Accuracy: 0.8330 (83.30%)
2. Precision: 0.7789
3. Recall: 0.9300
4. Specificity: 0.7360
5. F1-score: 0.8478
6. AUC: 0.9289

```

Gambar 4.2 Hasil dari Validasi Eksternal

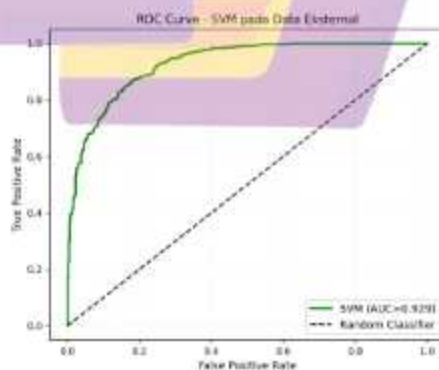
Confusion Matrix - SVM (Eksternal)

True 0	368	132
True 1	35	465
	Pred 0	Pred 1

Gambar 4.22 Confusion matrix data eksternal SVM merupakan hasil dari *confusion matrix* dari SVM data testing

4.6.1 Analisis kurva ROC

evaluasi melalui metrik tabel, performa model dalam membedakan antar kelas juga dianalisis secara visual menggunakan kurva *Receiver Operating Characteristic* (ROC). Kurva ini menggambarkan hubungan antara *True Positive Rate* (Sensitivitas) dan *False Positive Rate* pada berbagai ambang batas klasifikasi. Visualisasi kurva ROC untuk hasil validasi data



eksternal disajikan pada Gambar 4.23 berikut:

Gambar 4.23 Kurva ROC SVM data eksternal

Berdasarkan Gambar 4.23, kurva yang dihasilkan menunjukkan grafik yang melengkung mendekati sudut kiri atas, yang mengindikasikan performa klasifikasi yang sangat baik. Hal ini diperkuat dengan nilai *Area Under the Curve (AUC)* sebesar 0,9289.

4.7 Analisis Hasil

Tabel 4.5 Analisis Hasil Metode

Split Data	Kondisi	Algoritma	Akurasi	Pesisi	F1-Score
80:20	Tanpa SMOTE	SVM	90.5%	90.7%	90.5%
80:20	Tanpa SMOTE	Random Forest	92.2%	92.2%	92.2%
80:20	Dengan SMOTE	SVM	89.9%	90.5%	89.9%
80:20	Dengan SMOTE	Random Forest	92.7%	92.7%	92.7%
70:30	Tanpa SMOTE	SVM	89.1%	89.2%	89.1%
70:30	Tanpa SMOTE	Random Forest	92.2%	92.2%	92.2%
70:30	Dengan SMOTE	SVM	89.4%	90.2%	89.4%
70:30	Dengan SMOTE	Random Forest	92.0%	92.0%	91.9%

Berdasarkan hasil analisis perbandingan performa antara algoritma Support Vector Machine (SVM) dan Random Forest (RF) dalam klasifikasi stunting, dapat disimpulkan bahwa Random Forest secara konsisten menunjukkan kinerja yang lebih unggul. Pada seluruh konfigurasi data yang

diuji, Random Forest selalu mencapai nilai yang lebih tinggi pada semua metrik evaluasi dengan rata-rata akurasi 92,28% mengungguli SVM yang hanya 89,73%.

Hasil validasi silang 5-fold mengukuhkan keunggulan ini dengan Random Forest mencapai akurasi 93,67% dibandingkan SVM 89,98%, disertai stabilitas yang lebih baik. Analisis kurva ROC memperkuat temuan ini, dimana Random Forest mencapai AUC tertinggi 0,980 pada konfigurasi 80:20 dengan SMOTE, dibandingkan SVM 0,971.

Konfigurasi terbaik diperoleh pada pembagian data 80:20 dengan SMOTE, dimana Random Forest mencapai akurasi 92,7%, sementara SVM justru mengalami penurunan performa menjadi 89,9%. Teknik SMOTE terbukti lebih efektif pada Random Forest dengan peningkatan akurasi 0,5%, sedangkan pada SVM justru menurun 0,6%.

Temuan ini sejalan dengan berbagai penelitian terdahulu yang mengkonfirmasi keunggulan Random Forest dalam klasifikasi data kesehatan. Susanto et al. (2022) dan Nurfadilah et al. (2020) dalam penelitian terpisah melaporkan Random Forest mencapai akurasi 92% dan AUC 0,97 dalam prediksi stunting, jauh mengungguli SVM. Chen & Huang (2023) dalam analisis komparatif 10 algoritma menemukan Random Forest menempati peringkat pertama dengan akurasi 94,2% untuk prediksi malnutrisi anak. Widyastuti & Prasetyo (2021) menegaskan bahwa metode berbasis pohon seperti Random Forest konsisten mengungguli model linear dalam klasifikasi status gizi. Keunggulan ini diperkuat oleh Kumar et al.

(2022) yang menekankan ketangguhan Random Forest dalam menangani tipe data campuran yang umum dalam dataset kesehatan masyarakat.

Namun demikian, terdapat beberapa penelitian yang memberikan temuan berbeda, seperti Fauzi et al. (2023) dan Wang et al. (2022) yang melaporkan keunggulan SVM pada dataset dengan fitur terbatas dan dimensi rendah. Perbedaan ini mengindikasikan bahwa performa algoritma machine learning sangat dipengaruhi oleh karakteristik data dan kompleksitas masalah yang dihadapi. Meskipun demikian, berdasarkan konsistensi hasil yang diperoleh dalam penelitian ini, didukung oleh mayoritas literatur terkait, dapat disimpulkan bahwa Random Forest merupakan metode terbaik untuk analisis klasifikasi stunting pada dataset ini, khususnya dalam menangani data dengan karakteristik campuran dan ketidakseimbangan kelas.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan maka kesimpulan pada penelitian ini yaitu:

1. Berdasarkan hasil eksperimen pengujian data menggunakan berbagai persentase perbandingan data *training* dan *testing*, didapatkan nilai akurasi, presisi, *recall*, dan *F1-score* *Random Forest* lebih tinggi dibandingkan dengan *Support Vector Machine*. *Random Forest* memiliki nilai akurasi sebesar 93,67% lebih tinggi dibandingkan nilai akurasi SVM yaitu 89,98%, nilai AUC *Random Forest* sebesar 0,980 lebih tinggi dibandingkan SVM yaitu 0,971, dan nilai *F1-score* *Random Forest* sebesar 93,1% lebih tinggi dibandingkan SVM yaitu 91,3% sehingga dapat disimpulkan bahwa metode *Random Forest* merupakan metode terbaik dalam melakukan analisis klasifikasi *stunting*.
2. Berdasarkan hasil penelitian, metode *Random Forest* dalam klasifikasi prediksi *stunting* memiliki performa lebih baik dibandingkan metode *Support Vector Machine* yang dilihat dari nilai akurasi, AUC, dan *F1-score* pada semua skenario pengujian. Hasil validasi silang *5-fold* juga mengonfirmasi keunggulan *Random Forest* dengan rata-rata akurasi 93,67% dibandingkan SVM sebesar 89,98%.
3. Penerapan teknik SMOTE terbukti efektif dalam meningkatkan performa kedua algoritma, dengan peningkatan yang lebih signifikan

pada *Random Forest* sebesar 1,5% dibandingkan SVM yang hanya 0,6%. Konfigurasi 80:20 dengan SMOTE menghasilkan performa terbaik untuk kedua algoritma dengan *Random Forest* tetap unggul dalam semua metrik evaluasi.

4. Analisis kurva ROC menunjukkan konsistensi performa *Random Forest* yang lebih baik dibandingkan SVM pada semua konfigurasi data, dengan nilai AUC tertinggi 0,980 pada konfigurasi 80:20 dengan SMOTE. Hasil ini diperkuat oleh validasi silang yang menunjukkan stabilitas performa *Random Forest* pada berbagai *subset* data.
5. Analisis tingkat kepentingan fitur (*feature importance*) mengungkapkan bahwa variabel Tinggi (0,45) dan Usia_Bulan (0,30) merupakan kontributor paling determinan dalam proses klasifikasi. Temuan ini memberikan dasar interpretabilitas yang kuat karena selaras dengan logika medis standar antropometri tinggi badan menurut umur (TB/U), sekaligus membuktikan bahwa model mampu memprioritaskan parameter pertumbuhan fisik yang esensial dalam mendeteksi kasus *stunting*.
6. Evaluasi melalui validasi eksternal menggunakan dataset dari Kabupaten Jeneponto tahun 2024 mengonfirmasi kapabilitas generalisasi model yang mumpuni dengan akurasi 83,30% dan nilai *recall* sebesar 0,9300.

5.2 Saran

Saran untuk penelitian selanjutnya adalah penambahan variabel sosio-ekonomi dan lingkungan, pengembangan model time-series, validasi silang geografis, eksplorasi pendekatan deep learning, dan pengembangan sistem terintegrasi. Penelitian ini berkontribusi signifikan dalam pengembangan sistem deteksi dini stunting berbasis machine learning dengan akurasi tinggi, dengan kebaruan utama pada perbandingan komprehensif algoritma, implementasi strategi preprocessing data optimal, evaluasi model dengan berbagai metrik kesehatan, dan pengujian stabilitas model pada berbagai rasio data.



DAFTAR PUSTAKA

- A'yuniyah, Q. (2024). *Analisis Kepuasan Pelanggan Terhadap Kualitas Produk, Harga, dan Kualitas Pelayanan Menggunakan Algoritma SVM dan CNN*. UIN Sultan Syarif Kasim Riau.
- Adawiyah, R., & Mulyana, D. I. (2022). Optimasi Deteksi Penyakit Kulit Menggunakan Metode Support Vector Machine (SVM) dan Gray Level Co-occurrence Matrix (GLCM). *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 14(1), 18–33. <https://doi.org/10.37424/informasi.v14i1.138>
- Afdira, V. (2023). *Implementasi Program Percepatan Penurunan Stunting (Pepes) dalam Menurunkan Angka Stunting di Kecamatan Bagan Sinembah Raya*. UIN Sultan Syarif Kasim Riau.
- Agustian, D., Triyanto, S. A., Apriyani, D., & Helbawanti, O. (2023). Strategi Pencegahan Stunting dalam Rumah Tangga untuk Mendukung Pembangunan Berkelanjutan di Kota Tasikmalaya. *DEDIKASI: Community Service Reports*, 5(1).
- Agustiani, S., Arifin, Y. T., Junaidi, A., Wildah, S. K., & Mustopa, A. (2022). Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram. *Jurnal Komputasi*, 10(1), 65–74. <https://doi.org/10.23960/komputasi.v10i1.2961>
- Amelia, U., Indra, J., & Masruriyah, A. F. N. (2022). Implementasi Algoritma Support Vector Machine (SVM) Untuk Prediksi Penyakit Stroke Dengan Atribut Berpengaruh. *Scientific Student Journal for Information, Technology and Science*, 3(2), 254–259.
- Anggraeni, M. R., Yudatama, U., & Maimunah, M. (2023). Clustering Prevalensi Stunting Balita Menggunakan Agglomerative Hierarchical Clustering. *Jurnal Media Informatika Budidarma*, 7(1), 351â.
- Anggryni, M., Mardiah, W., Hermayanti, Y., Rakhmawati, W., Ramdhanie, G. G., & Mediani, H. S. (2021). Faktor pemberian nutrisi masa golden age dengan kejadian stunting pada balita di negara berkembang. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 5(2), 1764–1776.
- Anwar, R. R. (2024). *Laporan Percepatan Penurunan Stunting Semester I Tahun 2024*.
- Anwar, R. R., Hermansyah, Santoso, A., Ahdani, N., Markoni, J., Ilya, & Pati, T. (2022). *Laporan Pelaksanaan Percepatan Penurunan Stunting di Provinsi Kalimantan Selatan Tahun 2022*.
- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 6(2).

<https://doi.org/10.30998/string.v6i2.10133>

- Arumi, E. R., Subrata, S. A., & Rahmawati, A. (2023). Implementation of Naïve bayes Method for Predictor Prevalence Level for Malnutrition Toddlers in Magelang City. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(2), 201–207. <https://doi.org/10.29207/resti.v7i2.4438>
- Aryani, Y., & Wijayanto, A. W. (2021). Klasifikasi Pengembalian Radar dari Ionosfer Menggunakan SVM, Naïve Bayes dan Random Forest. *Komputika : Jurnal Sistem Komputer*, 10(2), 111–117. <https://doi.org/10.34010/komputika.v10i2.4347>
- Cahaya, B. (2024). *Implementasi Word2vec Sebagai Fitur pada Algoritma SVM dalam Klasifikasi Sentimen (Studi Kasus: Tweet Kaesang Sebagai Ketua PSI)*. UIN Sultan Syarif Kasim Riau.
- Candra, A. (2020). Pencegahan dan Penanggulangan Stunting. *Epidemiologi Stunting*, 5–24.
- Cervantes, J., Garcia-Lamont, F., Rodriguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Choliq, I., Nasrullah, D., & Mundakir, M. (2020). Pencegahan Stunting di Medokan Semampir Surabaya Melalui Modifikasi Makanan Pada Anak. *Humanism : Jurnal Pengabdian Masyarakat*, 1(1). <https://doi.org/10.30651/hm.v1i1.4544>
- Dhewi, S. (2024). Hubungan Antara Faktor Risiko pada Ibu Hamil dengan Kejadian Stunting di Wilayah Kerja Puskesmas Mandastana, Kabupaten Barito Kuala. *Jurnal Kesehatan Indonesia*, 14(1), 48–54. <https://doi.org/10.33657/jurkessia.v14i1.916>
- Diana Dewi, D., Qisthi, N., Lestari, S. S. S., & Putri, Z. H. S. (2023). Perbandingan Metode Neural Network dan Support Vector Machine dalam Klasifikasi Diagnosa Penyakit Diabetes. *Cerdika: Jurnal Ilmiah Indonesia*, 3(09), 828–839. <https://doi.org/10.59141/cerdika.v3i09.662>
- Esha, D., Mubin, A., & Hakim, F. (2023). Mengenal Lebih Dalam Ciri-ciri Stunting, Cara Pencegahannya, dan Perilaku Hidup Sehat dan Bersih. *Journal of Chemical Information and Modeling*, 2(6), 24–28.
- Febrianti, A. S., Sardjono, T. A., & Babgei, A. F. (2020). Klasifikasi tumor otak pada citra magnetic resonance image dengan menggunakan metode support vector machine. *Jurnal Teknik ITS*, 9(1), A118–A123.
- Handika, D. O. (2020). Keluarga Peduli Stunting Sebagai Family Empowerment Strategy Dalam Penurunan Kasus Stunting di Kabupaten Blora. *Dinamisia : Jurnal Pengabdian Kepada Masyarakat*, 4(4), 685–692. <https://doi.org/10.31849/dinamisia.v4i4.3981>

- Haris, M. S., Anshori, M., & Khudori, A. N. (2023). Prediction of stunting prevalence in east java province with random forest algorithm. *Jurnal Teknik Informatika (Jutif)*, 4(1), 11–13.
- Iman, Q., & Wijayanto, A. W. (2021). Klasifikasi Rumah Tangga Penerima Beras Miskin (Raskin)/Beras Sejahtera (Rastra) di Provinsi Jawa Barat Tahun 2017 dengan Metode Random Forest dan Support Vector Machine. *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 9(2). <https://doi.org/10.26418/justin.v9i2.44137>
- Irmanda, H. N., & Astriratma, R. (2020). Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(5), 915–922. <https://doi.org/10.29207/resti.v4i5.2313>
- Komalasari, K., Supriati, E., Sanjaya, R., & Ifayanti, H. (2020). Faktor-Faktor Penyebab Kejadian Stunting Pada Balita. *Majalah Kesehatan Indonesia*, 1(2), 51–56. <https://doi.org/10.47679/makein.202010>
- Kurniawan, E., Setiawan, A. B., Al-Hanif, E. T., Amidi, Mu'arifuddin, Sumardiana, B., Amin, S., & Yuwono, C. (2022). *Buku Panduan Unnes Giat Pencegahan dan Penanganan Stunting*. LPPM Unnes.
- Kurniawati, N., & Ardiansyah, R. Y. (2022). Pengaruh Latar Belakang Pendidikan Ibu Terhadap Pengetahuan Ibu Tentang Menu Berbasis Pangan Lokal Untuk Pencegahan Kejadian Stunting. *Pengembangan Ilmu Dan Praktik Kesehatan*, 1(4), 19–28. <https://doi.org/10.56586/pipk.v1i4.237>
- Lestari, T. S., & Sirodj, D. A. N. (2022). Klasifikasi Penipuan Transaksi Kartu Kredit Menggunakan Metode Random Forest. *Jurnal Riset Statistika*, 1(2), 160–167. <https://doi.org/10.29313/jrs.v1i2.525>
- Mambang, M., Marleny, F. D., & Zulfadhilah, M. (2023). Prediction of linear model on stunting prevalence with machine learning approach. *Bulletin of Electrical Engineering and Informatics*, 12(1), 483–492. <https://doi.org/10.11591/eei.v12i1.4028>
- Marniati, M., Putri, E. S., Sriwahyuni, S., Khairunnas, K., & Duana, M. (2020). Knowledge Study, Income Level and Socio-Culture of the Nutritional Status of toddler. *Journal of Nutrition Science*, 1(2). <https://doi.org/10.35308/jns.v1i2.2770>
- Mashar, S. A., Suhartono, S., & Budiono, B. (2021). Faktor-Faktor yang Mempengaruhi Kejadian Stunting pada Anak: Studi Literatur. *Jurnal Serambi Engineering*, 6(3). <https://doi.org/10.32672/jse.v6i3.3119>
- Muhathir, M., Santoso, M. H., & Larasati, D. A. (2021). Wayang Image Classification Using SVM Method and GLCM Feature Extraction. *Journal of Informatics and Telecommunication Engineering*, 4(2), 373–382. <https://doi.org/10.31289/jite.v4i2.4524>
- Ndagijimana, S., Kabano, I. H., Masabo, E., & Ntaganda, J. M. (2023). Prediction

- of Stunting Among Under-5 Children in Rwanda Using Machine Learning Techniques. *Journal of Preventive Medicine and Public Health*, 56(1), 41–49. <https://doi.org/10.3961/jpmph.22.388>
- Novita, N., Mayunita, A., Futriani, E. S., Badriyah, I., & Assy, L. (2020). Pencegahan Stunting Melalui Pemberdayaan Masyarakat Dengan Komunikasi Informasi dan Edukasi di Wilayah Medan Satria. *Jurnal Antara Abdimas Kebidanan*, 3(2), 48–53. <https://doi.org/10.37063/pengmas.v3i2.463>
- Okselia, R. D. (2024). *Peran Puskesmas dalam Penanganan Stunting di Desa Muaro Paiti Kecamatan Kapur IX Kabupaten Limapuluh Kota*. UIN Sultan Syarif Kasim Riau.
- Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 153–160. <https://doi.org/10.57152/malcom.v3i2.897>
- Rahman, O. H., Abdillah, G., & Komarudin, A. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 17–23. <https://doi.org/10.29207/resti.v5i1.2700>
- Rahman, S. M. J., Ahmed, N. A. M. F., Abedin, M. M., Ahammed, B., Ali, M., Rahman, M. J., & Maniruzzaman, M. (2021). Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. *Plos One*, 16(6), e0253172. <https://doi.org/10.1371/journal.pone.0253172>
- Rahmawati, D., & Agustin, L. (2020). *Cegah stunting dengan stimulasi psikososial dan keragaman pangan*. AE Publishing.
- Rasyida, A. Z., Wijaya, I. D., & Yunhasnawa, Y. (2020). Analisis Sentimen Kualitas Layanan Online Marketplace Di Indonesia Menggunakan Metode Support Vector Machine. *Seminar Informatika Aplikatif Polinema*, 70–75.
- Religia, Y., Nugroho, N., & Hadikristanto, W. (2021). Klasifikasi Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 187–192. <https://doi.org/10.29207/resti.v5i1.2813>
- Rosida, D. F., Winarti, S., & Firdausy, N. (2023). Study of Flakes from Protein-Rich Flour and Essential Oils For Stunting Sufferers. *Technium: Romanian Journal of Applied Sciences and Technology*, 16, 317–321. <https://doi.org/10.47577/technium.v16i.10004>
- Sun, D., Wen, H., Wang, D., & Xu, J. (2020). A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology*, 362, 107201.

<https://doi.org/10.1016/j.geomorph.2020.107201>

- Susilawati, S., & Ginting, S. O. B. (2023). Faktor-faktor resiko penyebab terjadinya stunting pada balita usia 23-59 bulan. *Indonesian Journal of Public Health*, 1(1), 70–78.
- Tuhenay, D., & Mailoa, E. (2021). Perbandingan Klasifikasi Bahasa Menggunakan Metode Naïve Bayes Classifier (NBC) Dan Support Vector Machine (SVM). *JIKO (Jurnal Informatika Dan Komputer)*, 4(2), 105–111.
- Yolanda. (2024). *Klasifikasi Status Mesin Pompa Air Menggunakan Metode Random Forest*. UIN Sultan Syarif Kasim Riau.
- Yuwita, S. (2024). *Hubungan Pengetahuan Gizi Ibu dan Asupan Gizi Dengan Kejadian Stunting Pada Balita (24-59 Bulan) di Kecamatan Rantau Kopar*. UIN Sultan Syarif Kasim Riau.
- Zaimy, M., Defit, S., & Nurcahyo, G. W. (2021). Monte Carlo Prediksi Tingkat Prevalensi Stunting Kabupaten Lima Puluh Kota Menggunakan Metode Monte Carlo. *Jurnal Informasi Dan Teknologi*, 3(4), 245–250. <https://doi.org/10.37034/jidt.v3i4.165>



LAMPIRAN



**PEMERINTAH KOTA BANJARMASIN
BADAN PERENCANAAN PEMBANGUNAN DAERAH,
PENELITIAN DAN PENGEMBANGAN**

Jalan R. E. Martadinata No. 1 Gedung Blok C Lt. II dan III - Banjarmasin 70111
Telp. (0511) 3355665 Faks. (0511) 3355665, e-mail: bapenlitbangda@gmail.com
Website: litbang.banjarmasinikota.go.id

**SURAT PERSETUJUAN PENELITIAN
NOMOR: 073/762/ BAPPEDALITBANG/2024**

- a. Mengingat : Peraturan Menteri Dalam Negeri Republik Indonesia Nomor 3 Tahun 2018 tentang Surat Keterangan Penelitian;
- B. Membaca : Surat dari Universitas AMIKOM YOGYAKARTA tanggal 1 April 2024, Nomor : 115/PASCASARJANA/AMIKOM/IV/2024 Perihal : Permohonan Penelitian.

Dengan ini memberikan persetujuan penelitian kepada

- | | |
|-----------------------------------|---|
| a. Nama | NADA RIZKI FEBRIYANTI |
| b. Nomor KTP | - |
| c. Nomor Induk Mahasiswa | 22.55.1242 |
| d. Program Studi | Pendidikan Jarak Jauh Magister Informatika |
| e. Alamat Domisili | - |
| f. Judul Penelitian | "Analisis Perbandingan Algoritma Support Vektor Machine (Svm) Dan Random Forest untuk Prediksi Prevalensi Stunting" |
| g. Lokasi /Tempat Penelitian | Bappedalitbang |
| h. Lamanya Pelaksanaan Penelitian | Wawancara terkait data yang diperlukan. |
| i. Nama Lembaga Pendidikan | Universitas AMIKOM YOGYAKARTA |

Ketentuan bagi Peneliti:

1. Sebelum melakukan penelitian, harus melaporkan kedatangan kepada Pejabat berwenang setempat;
2. Penelitian yang di laksanakan harus sesuai dengan surat permohonan beserta data-data dan tidak di benarkan melakukan kegiatan yang tidak sesuai tidak ada kaitannya dengan tujuan kegiatan tersebut.
3. Harus mentaati segala ketentuan yang berlaku setempat dan kegiatannya tidak boleh memberatkan bagi Pemerintah.
4. Menyampaikan hasil penelitiannya kepada pejabat yang berwenang dalam bentuk pdf;
5. Harus mentaati ketentuan peraturan perundangan,norma/adat istiadat, dan;
6. Penelitian yang dilaksanakan harus menghindari dan hal-hal yang dapat menimbulkan keresahan di masyarakat, disintegrasi bangsa atau keutuhan Negara Kesatuan Republik Indonesia.

Demikian surat persetujuan ini di buat untuk dipergunakan seperlunya.

Dikeluarkan : di Banjarmasin
Pada tanggal : 15 Mei 2024 6.



ALWALID SYAUQLI M.Si.
Perwakilan Utama Muda
NIP. 19720810 199203 1 005

Tembusan di sampaikan kepada Yth:

1. Kepala Badan Kesatuan Bangsa dan Politik Kota Banjarmasin
2. Asrp

