

**ANALISIS DETERMINAN PENYEBAB *STUNTING* PROVINSI
DI INDONESIA: APLIKASI MODEL *RANDOM FOREST* DAN
INTERPRETASI SHAP PADA DATA SSGI TAHUN 2024**

TUGAS AKHIR

Diajukan sebagai syarat menyelesaikan jenjang strata Satu (S-1) di
Program Studi Teknik Informatika, Fakultas Teknologi Industri,
Institut Teknologi Sumatera

Oleh:

Cindy Nadila Putri

122140002



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SUMATERA
LAMPUNG SELATAN
2026**

LEMBAR PENGESAHAN

Saya menyatakan bahwa Tugas Akhir berjudul “Analisis Determinan Penyebab *Stunting* Provinsi di Indonesia: Aplikasi Model *Random Forest* dan Interpretasi SHAP pada Data SSGI Tahun 2024” merupakan hasil karya saya sendiri dan belum pernah diajukan, baik sebagian maupun seluruhnya, di Institut Teknologi Sumatera atau institusi pendidikan lain oleh saya maupun pihak lain.

Lampung Selatan, 18 Januari 2026
Penulis,

Cindy Nadila Putri
NIM. 122140002

Foto 2x3

Diperiksa dan disetujui oleh,
Pembimbing

1. Martin Clinton Tosima Manullang, Ph.D.
NIP. 19930109 2019 03 1 017
2. Martin Clinton Tosima Manullang, Ph.D.
NIP. 19930109 2019 03 1 017

.....

.....

Penguji

1. Dosen Penguji I
NIP. 19900000 2000 00 0 000
2. Dosen Penguji II
NIP. 19900000 2000 00 0 000

.....

.....

Disahkan oleh,
Koordinator Program Studi Teknik Informatika
Fakultas Teknologi Industri
Institut Teknologi Sumatera

Andika Setiawan, S.Kom., M.Cs.
NIP. 19911127 2022 03 1 007

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir dengan judul “Analisis Determinan Penyebab *Stunting* Provinsi di Indonesia: Aplikasi Model *Random Forest* dan Interpretasi SHAP pada Data SSGI Tahun 2024” adalah karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan benar.

Nama : Cindy Nadila Putri

NIM : 122140002

Tanda Tangan :

Tanggal :

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Institut Teknologi Sumatera, saya yang bertanda tangan di bawah ini:

Nama : Cindy Nadila Putri

NIM : 122140002

Program Studi : Teknik Informatika

Fakultas : Teknologi Industri

Jenis Karya : Tugas Akhir

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Institut Teknologi Sumatera **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Determinan Penyebab *Stunting* Provinsi di Indonesia: Aplikasi Model *Random Forest* dan Interpretasi SHAP pada Data SSGI Tahun 2024

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Institut Teknologi Sumatera berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Lampung Selatan
Pada tanggal : 18 Januari 2026

Yang menyatakan

Cindy Nadila Putri

KATA PENGANTAR

Pada halaman ini mahasiswa berkesempatan untuk menyatakan terima kasih secara tertulis kepada pembimbing dan pihak lain yang telah memberi bimbingan, nasihat, saran dan kritik, kepada mereka yang telah membantu melakukan penelitian, kepada perorangan atau lembaga yang telah memberi bantuan keuangan, materi dan/atau sarana. Cara menulis kata pengantar beraneka ragam, tetapi hendaknya menggunakan kalimat yang baku. Ucapan terima kasih agar dibuat tidak berlebihan dan dibatasi pada pihak yang terkait secara ilmiah (berhubungan dengan subjek/materi penelitian).

Puji syukur kehadiran Allah SWT/Tuhan Yang Maha Esa atas limpahan rahmat, karunia, serta petunjuk-Nya sehingga penyusunan tugas akhir ini telah terselesaikan dengan baik. Dalam penyusunan tugas akhir ini penulis telah banyak mendapatkan arahan, bantuan, serta dukungan dari berbagai pihak. Oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. [Rektor ITERA] selaku Rektor Institut Teknologi Sumatera.
2. [Dekan FTI] selaku Dekan Fakultas Teknologi Industri.
3. [Koor Prodi IF] selaku Ketua Program Studi Teknik Informatika.
4. [Dosen Pembimbing] selaku Dosen Pembimbing atas ide, waktu, tenaga, perhatian, dan masukan yang telah disumbangsihkan kepada penulis.
5. [Isi nama lainnya]

Akhir kata penulis berharap semoga tugas akhir ini dapat memberikan manfaat bagi kita semua.

RINGKASAN

Analisis Determinan Penyebab *Stunting* Provinsi di Indonesia: Aplikasi Model *Random Forest* dan Interpretasi SHAP pada Data SSGI Tahun 2024

Cindy Nadila Putri

Halaman Ringkasan berisi uraian singkat tentang latar belakang masalah, rumusan masalah, tujuan, metodologi penelitian, hasil dan analisis data, serta kesimpulan dan saran. Isi ringkasan tidak lebih dari 1000 kata (sekitar maksimal 2 halaman).

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

ABSTRAK

Analisis Determinan Penyebab *Stunting* Provinsi di Indonesia: Aplikasi Model *Random Forest* dan Interpretasi SHAP pada Data SSGI Tahun 2024

Cindy Nadila Putri

Halaman ABSTRAK berisi uraian tentang latar belakang, tujuan, metodologi penelitian, hasil / kesimpulan. Ditulis dalam BAHASA INDONESIA tidak lebih dari 250 kata, dengan jarak antar baris satu spasi. Pada akhir abstrak ditulis kata “Kata Kunci” yang dicetak tebal, diikuti tanda titik dua dan kata kunci yang tidak lebih dari 5 kata. Kata kunci terdiri dari kata-kata yang khusus menunjukkan dan berkaitan dengan bahan yang diteliti, metode/instrumen yang digunakan, topik penelitian. Kata kunci diketik pada jarak dua spasi dari baris akhir isi abstrak.

Kata Kunci: kunci1, kunci2

ABSTRACT

Determinant Analysis of Stunting Causes Across Indonesian Provinces:
Application of Random Forest Model and SHAP Interpretation on SSGI 2024

Data

Cindy Nadila Putri

Halaman ABSTRACT berisi uraian tentang latar belakang, tujuan, metodologi penelitian, hasil / kesimpulan. Ditulis dalam BAHASA INGGRIS tidak lebih dari 250 kata, dengan jarak antar baris satu spasi. Secara khusus, kata dan kalimat pada halaman ini tidak perlu ditulis dengan huruf miring meskipun menggunakan Bahasa Inggris, kecuali terdapat huruf asing lain yang ditulis dengan huruf miring (misalnya huruf Latin atau Greek, dll). Pada akhir abstract ditulis kata “Keywords” yang dicetak tebal, diikuti tanda titik dua dan kata kunci yang tidak lebih dari 5 kata. Keywords terdiri dari kata-kata yang khusus menunjukkan dan berkaitan dengan bahan yang diteliti, metode/instrumen yang digunakan, topik penelitian. Keywords diketik pada jarak dua spasi dari baris akhir isi abstrak.

Keywords: keywords1, keywords2

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR RUMUS

Rumus 2.1 Rumus sederhana 12

Rumus 2.2 Mean Absolute Error (MAE) 12

Rumus 2.3 Distribusi Normal 13

Rumus 2.4 Sistem persamaan linier 13

DAFTAR KODE

Kode 4.1 Akuisisi Gambar	21
--------------------------------	----

BAB I

PENDAHULUAN

1.1 Latar Belakang

Manusia memerlukan kondisi tubuh yang sehat untuk dapat beraktivitas secara maksimal. Saat ini salah satu masalah kesehatan yang banyak terjadi di Indonesia adalah kendala *stunting*. Masalah ini telah menjadi perhatian serius karena dampaknya yang meluas, terutama pada anak-anak sebagai generasi penerus bangsa. Berbagai faktor, seperti pola hidup, akses terhadap makanan bergizi, serta kualitas pelayanan kesehatan, turut berkontribusi terhadap tingginya prevalensi masalah ini.

Penyakit *stunting* merupakan kondisi di mana tubuh mengalami kekurangan gizi secara berlebihan dan terjadi pada rentang waktu yang cukup lama **said2024pencegahan**. Dampak dari masalah ini akan mengakibatkan kendala pertumbuhan pada anak sehingga tinggi badan anak cenderung menjadi lebih pendek. Tak hanya memengaruhi pertumbuhan fisik, penyakit *stunting* juga akan berpengaruh ke dalam aspek pertumbuhan lainnya seperti mental, intelektual, dan kognitif anak **rahagia2023upaya**. Oleh karena itu, penting untuk memahami faktor-faktor yang menyebabkan terjadinya *stunting* agar dapat mengambil langkah-langkah pencegahan yang efektif.

Di Indonesia, kasus *stunting* terjadi pada balita usia 0-5 tahun berada pada persentase sebesar 19,8% menurut Survei Status Gizi Indonesia (SSGI) **kemenkes2025ssgi**. Meskipun telah mengalami penurunan dari tahun sebelumnya yaitu angka 21,5%, hasil ini menunjukkan bahwa target pemerintah Indonesia yaitu menurunkan prevalensi *stunting* sampai 14,4% di tahun 2029 masih belum tercapai **bkpk2023ski**, **setneg2024strategi**. *Stunting* dapat berasal dari faktor-faktor yang sangat kompleks dilihat dari aspek sosial, biologis, maupun lingkungan. Biasanya, penyebab utama pada *stunting*

dapat terjadi karena kurangnya asupan gizi, buruknya sanitasi, serta akses yang rendah terhadap pelayanan kesehatan **supriyanto2023implementasi, saleh2023edukasi**. Terdapat faktor penting lain yang menyebabkan terjadinya stunting yaitu kurangnya edukasi ibu tentang betapa penting untuk menjaga gizi seimbang pada masa kehamilan, masa menyusui, dan masa pertumbuhan anak **saleh2023edukasi**. Pada beberapa kasus, ibu hamil yang melahirkan bayi dengan kondisi berat badan lahir rendah (BBLR) mengalami fase kekurangan gizi selama masa kehamilannya **saleh2023edukasi**. Hal ini dapat mengakibatkan terjadinya *stunting* di kemudian hari.

Tingkat *stunting* di Indonesia sendiri dinilai tidak merata di seluruh wilayah. Beberapa provinsi seperti Bali dan DI Yogyakarta menunjukkan angka prevalensi yang relatif rendah, berbanding terbalik dengan provinsi Nusa Tenggara Timur dan Sulawesi Barat yang mencatat angka jauh lebih tinggi **kemenkes2025ssgi**. Situasi yang timpang ini menunjukkan adanya variasi faktor determinan yang memengaruhi kejadian *stunting* di tiap daerah, baik dari segi sosial, ekonomi, pendidikan, maupun kondisi lingkungan **zemariam2025prediction**. Faktor lainnya seperti status gizi ibu, akses terhadap layanan kesehatan, dan sanitasi dasar memiliki kontribusi yang berbeda-beda pada setiap wilayah. Adanya variasi faktor risiko ini mengindikasikan bahwa pendekatan penanganan yang bersifat generalis memiliki keterbatasan dalam menjangkau akar masalah di tiap daerah, sehingga dibutuhkan landasan data yang spesifik per wilayah untuk mendukung perumusan strategi intervensi yang lebih presisi dan tepat sasaran. Maka dari itu, diperlukan analisis yang lebih mendalam untuk memahami faktor-faktor yang paling berpengaruh untuk situasi *stunting* tiap provinsi.

Penelitian sebelumnya menerapkan algoritma *machine learning* untuk memprediksi *stunting* pada kalangan remaja di Ethiopia dan menemukan bahwa metode konvensional kurang mampu menangkap interaksi kompleks antarvariabel **zemariam2025prediction**. Namun penelitian tersebut dilakukan

secara terbatas pada kelompok usia remaja dan wilayah tertentu, sehingga belum secara langsung dapat digeneralisasi dan masih memerlukan kajian tersendiri untuk ke skala nasional Indonesia yang mencakup berbagai karakteristik. Dengan demikian, penting untuk mengeksplorasi pendekatan serupa pada skala provinsi di Indonesia untuk memahami faktor determinan *stunting* berdasarkan kondisi lokal.

Dalam beberapa tahun terakhir, pendekatan *machine learning* mulai banyak digunakan dalam bidang kesehatan dan gizi untuk mengidentifikasi pola tersembunyi dalam data yang kompleks **tamanna2025identifying**. Salah satu algoritma yang populer adalah *Random Forest*, yang dikenal memiliki kemampuan tinggi dalam menangani variabel dalam jumlah besar serta menghasilkan prediksi yang akurat. Keunggulan lain dari *Random Forest* adalah kemampuannya dalam mengukur tingkat kepentingan setiap fitur (*feature importance*), sehingga dapat membantu memahami faktor-faktor mana yang paling berpengaruh terhadap suatu fenomena. Namun demikian, model ini sering dianggap sebagai *black box* karena sulit dijelaskan secara langsung oleh pengambil kebijakan atau peneliti non-teknis.

Untuk mengatasi keterbatasan interpretasi tersebut, digunakan metode interpretabilitas seperti SHAP (*SHapley Additive exPlanations*) yang mampu menjelaskan kontribusi masing-masing fitur terhadap hasil prediksi model **tamanna2025identifying**. Melalui pendekatan ini, setiap faktor dapat dinilai apakah ia meningkatkan atau menurunkan kemungkinan terjadinya *stunting* pada suatu wilayah. Dengan demikian, kombinasi antara *Random Forest* dan SHAP tidak hanya memberikan hasil prediksi yang akurat, tetapi juga penjelasan yang dapat dipahami secara intuitif oleh pengambil keputusan. Pendekatan ini memberikan nilai tambah dalam analisis kebijakan berbasis data, terutama untuk menentukan prioritas intervensi di daerah dengan tingkat *stunting* tinggi.

Menjawab permasalahan tersebut, penelitian ini diarahkan untuk

mengidentifikasi faktor-faktor determinan utama yang memengaruhi prevalensi *stunting* antarprovinsi di Indonesia menggunakan data SSGI tahun 2024 melalui pendekatan *Random Forest* yang diinterpretasikan dengan SHAP. Penelitian ini berfokus pada dua aspek utama, yaitu menilai performa model dalam memprediksi tingkat *stunting* serta memahami peran masing-masing variabel terhadap hasil prediksi. Hasil penelitian diharapkan dapat memberikan dasar empiris bagi pemerintah dan pemangku kebijakan dalam merancang strategi intervensi gizi yang lebih efektif dan berbasis bukti ilmiah.

1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang, dihasilkan rumusan masalah pada penelitian yang akan dijabarkan pada poin berikut:

1. Apa saja faktor-faktor yang menjadi determinan utama penyebab *stunting* antarprovinsi di Indonesia berdasarkan data SSGI tahun 2024?
2. Bagaimana performa model *Random Forest* dalam memprediksi prevalensi *stunting* antarprovinsi di Indonesia?
3. Bagaimana interpretasi SHAP digunakan untuk menjelaskan kontribusi masing-masing faktor terhadap hasil prediksi model *Random Forest*?

1.3 Tujuan Penelitian

Setelah penjabaran rumusan masalah, pada penelitian ini dibuat tujuan penelitian yaitu sebagai berikut:

1. Mengidentifikasi faktor-faktor determinan utama penyebab *stunting* antarprovinsi di Indonesia berdasarkan data SSGI tahun 2024.
2. Menganalisis performa model *Random Forest* dalam memprediksi prevalensi *stunting* antarprovinsi di Indonesia.
3. Menerapkan interpretasi SHAP untuk menjelaskan kontribusi masing-masing faktor terhadap hasil prediksi model *Random Forest*.

1.4 Batasan Masalah

Agar penelitian tidak terlalu luas dan tidak keluar dari pokok permasalahan, maka ditentukan batasan oleh beberapa poin berikut:

1. Data yang digunakan untuk penelitian ini terbatas pada Survei Status Gizi Indonesia (SSGI) tahun 2024 yang telah dipublikasikan oleh Kementerian Kesehatan RI. Perlu diketahui bahwa penelitian ini tidak menggunakan data tambahan dari sumber lain. Sehingga penelitian yang dilakukan bukan analisis *time-series* antartahun dan memiliki ruang lingkup spasial terbatas pada 34 provinsi di Indonesia sesuai ketersediaan data.
2. Populasi yang diambil untuk menjadi fokus penelitian ini adalah anak balita usia 0-59 bulan (5 tahun), sesuai dengan sasaran pengukuran prevalensi *stunting* pada SSGI.
3. Variabel yang digunakan untuk penelitian ini mencakup berbagai faktor yang memiliki potensi untuk memengaruhi prevalensi penyebab *stunting*, seperti karakteristik balita, karakteristik ibu, serta kondisi rumah tangga yang meliputi status gizi, pendidikan, ekonomi, dan akses terhadap fasilitas kesehatan, air bersih, serta sanitasi. variabel-variabel ini merupakan sebagian dari indikator yang tersedia dalam dataset SSGI tahun 2024 dan dipilih berdasarkan relevansi terhadap topik penelitian.
4. Analisis yang dilakukan pada penelitian ini difokuskan pada tingkat provinsi di Indonesia, sehingga hasilnya mencerminkan kondisi agregat dan tidak mempertimbangkan variasi di tingkat kabupaten/kota atau individu.
5. Model yang digunakan untuk penelitian ini adalah algoritma *Random Forest* untuk memprediksi prevalensi *stunting*, dengan interpretasi hasil menggunakan metode SHAP (*SHapley Additive exPlanations*) untuk memahami kontribusi setiap variabel terhadap prediksi prevalensi

stunting.

6. Penelitian ini berfokus pada analisis faktor determinan penyebab *stunting*, bukan pada pengembangan sistem prediksi *stunting* secara *real-time* atau aplikasi praktis lainnya.

1.5 Manfaat Penelitian

Penelitian ini dilaksanakan dengan harapan agar nantinya dapat memiliki manfaat sebagai berikut:

1. Bagi masyarakat, penelitian ini diharapkan dapat memberi gambaran empiris mengenai faktor-faktor utama yang memiliki kontribusi terhadap *stunting* di berbagai provinsi di Indonesia terutama di Sumatera. Hasil penelitian diharapkan dapat menjadi dasar bagi tenaga kesehatan dan pemerintah dalam menyusun strategi intervensi gizi yang lebih terarah dan sesuai dengan karakteristik wilayah masing-masing.
2. Bagi peneliti, penelitian ini bermanfaat untuk mengembangkan pemahaman mengenai penerapan algoritma *machine learning*, khususnya *Random Forest* dengan interpretasi SHAP dalam menganalisis data kesehatan masyarakat.
3. Bagi akademisi, dapat dijadikan referensi mahasiswa lain yang ingin meneliti terkait faktor-faktor penyebab *stunting* maupun penerapan model *machine learning* untuk analisis data survei kesehatan.

1.6 Sistematika Penulisan

Secara keseluruhan, struktur penulisan laporan ini disusun untuk memberikan gambaran umum setiap tahapan penelitian mengenai analisis determinan penyebab *stunting* di provinsi-provinsi di Indonesia. Setiap bab disusun secara sistematis agar pembahasan mengenai penerapan model *Random Forest* dan interpretasi SHAP pada data SSGI tahun 2024 dapat diikuti dengan runtut dan mudah dipahami.

Bab I

Bab I berisi pendahuluan. Pendahuluan menguraikan latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah dan sistematika penulisan dari penelitian.

Bab II

Bab II berisi tinjauan pustaka. Pada bab ini akan dibahas teori-teori mengenai *stunting*, faktor-faktor determinan penyebab *stunting*, Survei Status Gizi Indonesia (SSGI), serta penerapan algoritma *machine learning* khususnya *Random Forest* pada model prediksi risiko *stunting*. Selain itu, bab ini juga membahas hasil penelitian terdahulu yang relevan sebagai landasan konseptual dan pembeda penelitian ini dari studi yang telah dilakukan sebelumnya.

Bab III

Bab ini menjelaskan metode yang akan digunakan dalam penelitian, termasuk sumber dan jenis data, variabel penelitian, tahapan pengolahan data, serta rancangan model analisis menggunakan *Random Forest* dan SHAP. Selain itu, bab ini juga menjelaskan mengenai teknik evaluasi model serta prosedur interpretasi hasil untuk mengidentifikasi faktor determinan utama penyebab *stunting*.

Bab IV

Bab ini menguraikan hasil analisis data, performa model *Random Forest*, hasil interpretasi SHAP, serta pembahasan mengenai faktor-faktor determinan yang memiliki pengaruh prevalensi *stunting* antarprovinsi di Indonesia.

Bab V

Bab V berisi Kesimpulan dan saran. Pada bab ini akan diberikan kesimpulan dari hasil penelitian yang menjawab rumusan masalah berdasarkan hasil analisis yang telah dilakukan sebelumnya. Saran-saran

juga diberikan sebagai masukan bagi pihak terkait dan peneliti selanjutnya untuk pengembangan kebijakan dan penelitian lanjutan di bidang *stunting*.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Berisi penelitian terdahulu yang menjadi konsep / pendukung penelitian yang dilakukan. Lakukan pembahasan secara sistematis dengan menjelaskan masalah apa yang diangkat di penelitian terdahulu, metode yang digunakan, kontribusi yang diberikan, serta analisis penulis terkait dengan keunggulan atau keterbatasannya. Tuangkan perbandingan penelitian terdahulu dengan penelitian yang akan dikerjakan, minimal 5 jurnal pembandingan (4-5 tahun terakhir). Kemudian penulis sebaiknya melakukan rangkuman terutama terkait dengan peluang pengembangan atau tugas akhir yang akan dilakukan

Perujukan literatur dapat dilakukan dengan menambahkan entri baru dalam file `references.bib`. Cara merujuk sitasi menggunakan `\cite{nama label sitasi}`. Hasil sitasi seperti ini: **knuth2001art**, **Vogels2006Am**, atau **Suryanto2019MAE****Ivanny2014Banding**. Daftar Pustaka hanya akan memunculkan sitasi yang direferensikan menggunakan command `\cite{}`.

Tuliskan **judul jurnal, penulis jurnal, tahun jurnal terbit, penjelasan isi jurnal, metode penelitian, hasil penelitian & pengujian**.

1. Sistem Informasi Pendaftaran Haji dan Umroh Di PT Multazam Sriwijaya Barakah Palembang Menggunakan Metode Rapid Application Development. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur

adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2. Sistem Informasi Umroh Di PT XYZ Lampung Menggunakan Metode Rapid Application Development. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Tabel ringkasan tinjauan pustaka ditulis setelah penjelasan masing-masing jurnal.

Tabel 2.1 Literasi Penelitian Terdahulu

No.	Judul	Masalah	Metode	Hasil
1.	Sistem Informasi Pendaftaran Haji dan Umroh Di PT Multazam Sriwijaya Barakah Palembang Menggunakan Metode Rapid Application Development	Belum adanya sistem untuk pendaftaran haji & umrah	Rapid Application Development	Sistem Informasi Pendaftaran Haji dan Umroh di PT Multazam Sriwijaya Barakah Palembang
2.	Sistem Informasi Umroh Di PT XYZ Lampung Menggunakan Metode Rapid Application Development	Belum adanya sistem untuk pendaftaran haji & umrah	Rapid Application Development	Sistem Informasi Pendaftaran Umroh di PT XYZ Lampung

No.	Judul	Masalah	Metode	Hasil
3.	Sistem Informasi Umroh Di PT XYZ Lampung Menggunakan Metode Rapid Application Development	Belum adanya sistem untuk pendaftaran haji & umrah	Rapid Application Development	Sistem Informasi Pendaftaran Umroh di PT XYZ Lampung

2.2 Dasar Teori

Berisi teori/konsep yang berkaitan/digunakan dalam tugas akhir yang dikerjakan. Gunakanlah data melalui buku/jurnal referensi, publikasi tugas akhir, penelitian, buku, dan informasi web yang dapat dipertanggungjawabkan, hindari penggunaan dasar teori melalui tautan Wikipedia, surat kabar, atau portal berita, yang dapat memiliki isi yang tidak bersifat fakta.

2.2.1 Teori 1

Berikut adalah contoh penyisipan tabel menggunakan `\begin{longtable}{}{}`:

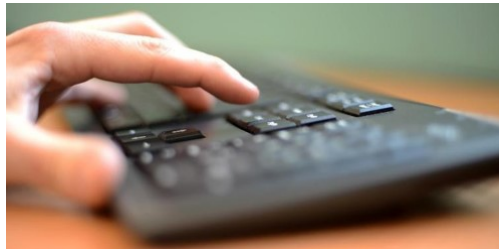
Tabel 2.2 Contoh Tabel

Col1	Col2	Col2	Col3
1	6	87837	787
2	7	78	5415
3	545	778	7507
4	545	18744	7560
5	88	788	6344

2.2.1.1 Subsubbab

Berikut adalah contoh subsubbab. Ini adalah level subbab maksimal dalam laporan Tugas Akhir, dan tidak boleh lebih dalam.

Gambar ?? adalah contoh Gambar yang diambil dari internet yang harus dicantumkan sumbernya dan memiliki lisensi Creative Common. Jika gambar adalah milik peneliti lain atau tidak dibuat atau diambil sendiri maka peneliti wajib meminta izin kepada peneliti lain tersebut untuk mencantumkan gambar. Gunakan `\begin{figure}` untuk memasukkan gambar. Gunakan `\caption{[nama caption]}` untuk memberikan caption gambar. Nomor caption akan diurutkan secara otomatis. Jangan lupa untuk melabeli setiap gambar dengan `\label{[nama label]}`, agar bisa direferensi menggunakan `\ref{[nama label]}`



Gambar 2.1 Contoh gambar dan caption
Sumber: Contoh

2.2.2 Teori 2

Untuk membuat sebuah rumus persamaan, gunakan kode `\begin{equationcaptioned}` seperti dibawah:

$$x + 1 = 2$$

(Rumus 2.1)

Teks caption rumus tidak akan muncul di teks, tetapi akan muncul di Daftar Rumus.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

(Rumus 2.2)

Berikut adalah contoh penulisan persamaan yang lebih kompleks, yaitu persamaan distribusi normal.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

(Rumus 2.3)

Jika menuliskan banyak persamaan secara berurutan, gunakan `\begin{split}`:

$$\begin{aligned} 2x - 5y &= 8 \\ 3x + 9y &= -12 \end{aligned}$$

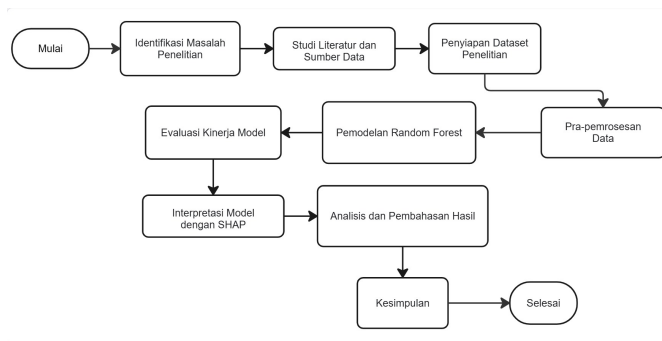
(Rumus 2.4)

BAB III

METODE PENELITIAN

3.1 Alur Penelitian

Penelitian ini dirancang untuk menganalisis faktor-faktor determinan penyebab *stunting* pada tingkat provinsi di Indonesia dengan memanfaatkan data sekunder *Annual Report Survei Status Gizi Indonesia (SSGI)* tahun 2024. Pendekatan yang digunakan mengombinasikan pemodelan *machine Learning* regresi menggunakan *Random Forest* dan metode interpretasi model SHAP (*SHapley Additive exPlanations*). Model ini dibangun menggunakan dataset agregat tingkat provinsi yang memuat berbagai indikator status gizi balita serta faktor determinan yang mencakup aspek kesehatan ibu dan anak, pola asuh, layanan kesehatan, penyakit, perlindungan sosial, dan kondisi sanitasi. Penelitian ini bertujuan untuk mengidentifikasi faktor determinan utama yang berkontribusi terhadap tingginya prevalensi *stunting* antar provinsi, serta memberikan pemahaman yang bersifat interpretatif sebagai dasar perumusan kebijakan dan intervensi gizi yang lebih tepat sasaran. Secara garis besar, tahapan penelitian digambarkan dalam diagram alir (*flowchart*) pada Gambar ??.



Gambar 3.1 Diagram Alir Penelitian

Berdasarkan Gambar ??, alur penelitian disusun secara sistematis yang diawali dengan tahap identifikasi masalah mengenai urgensi penanganan *stunting* di Indonesia dan kompleksitas faktor determinan yang mempengaruhinya di tingkat provinsi. Tahap ini didukung oleh studi literatur mendalam untuk memahami variabel-variabel determinan multidimensi yang mencakup berbagai aspek, serta metode pengolahan data dan pemodelan *machine learning* yang relevan. Berlandaskan pemahaman tersebut, selanjutnya dilakukan pengumpulan dataset yang bersumber dari Laporan Tahunan Survei Status Gizi Indonesia (SSGI) tahun 2024. Laporan Tahunan SSGI 2024 kemudian digunakan untuk menyiapkan dataset agregat tingkat provinsi yang memuat variabel target (prevalensi *stunting*) dan variabel fitur (faktor determinan). Data mentah yang diperoleh selanjutnya diproses melalui tahapan pra-pemrosesan data yang meliputi pembersihan data, penyaringan data, dan pemisahan variabel target dan fitur.

Setelah data siap, tahap berikutnya adalah pembangunan model regresi menggunakan algoritma *Random Forest*. Proses ini mencakup penentuan parameter terbaik melalui *Grid Search* dan validasi silang (*Cross Validation*) untuk memastikan performa model yang optimal. Model yang telah dilatih kemudian dievaluasi kinerjanya menggunakan metrik statistik seperti *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), serta koefisien determinasi (R^2 Score). Langkah terakhir adalah menginterpretasikan hasil prediksi model *Random Forest* menggunakan SHAP untuk mengetahui kontribusi setiap fitur determinan terhadap angka *stunting*, baik secara global skala nasional maupun lokal per provinsi.

3.2 Identifikasi Masalah

Tahap awal penelitian ini dilakukan dengan mengidentifikasi permasalahan mengenai prevalensi *stunting* di Indonesia serta kompleksitas

faktor determinan yang mempengaruhinya. Identifikasi masalah dalam penelitian ini dilakukan melalui tahapan observasi statistik awal terhadap distribusi dataset SSGI 2024 yang menyajikan data *stunting* dalam berbagai aspek determinan, seperti kesehatan ibu, pelayanan kesehatan, pola asuh, konsumsi pangan, lingkungan, serta faktor sosial ekonomi. Keberagaman indikator tersebut memperlihatkan bahwa permasalahan *stunting* bersifat multidimensional dan tidak hanya dipengaruhi oleh faktor tunggal. Temuan ini mendorong munculnya pertanyaan penelitian mengenai kesesuaian strategi intervensi gizi yang bersifat seragam dalam menurunkan prevalensi *stunting* di seluruh provinsi di Indonesia. Berdasarkan kondisi tersebut, penelitian ini ditujukan untuk mengidentifikasi faktor determinan utama penyebab *stunting* pada tingkat provinsi menggunakan pendekatan analisis data interpretasi berbasis *machine learning*.

3.3 Studi Literatur

Pada tahap studi literatur, dilakukan identifikasi mengenai pendekatan, metode, dan jenis data yang digunakan dalam penelitian terkait *stunting* terdahulu. Hal ini dilakukan untuk menemukan celah penelitian yang dapat diisi melalui pendekatan baru yang diusulkan. Hasil tinjauan literatur menunjukkan bahwa sebagian besar studi menerapkan metode prediksi *stunting* dengan algoritma *machine learning*, terutama algoritma *Random Forest*, *Gradient Boosting*, dan *XGBoost* dengan fokus utama pada peningkatan akurasi prediksi. Selain itu, sebagian besar penelitian *machine learning* mengenai *stunting* dilakukan menggunakan data individu atau wilayah yang terbatas dengan fokus utama pada peningkatan akurasi prediksi. Penelitian yang menggunakan data agregat tingkat provinsi masih sangat minim, begitu pula dengan penggunaan metode interpretasi model untuk memahami kontribusi faktor determinan penyebab *stunting*.

Berdasarkan temuan tersebut, penelitian ini mengadopsi pendekatan

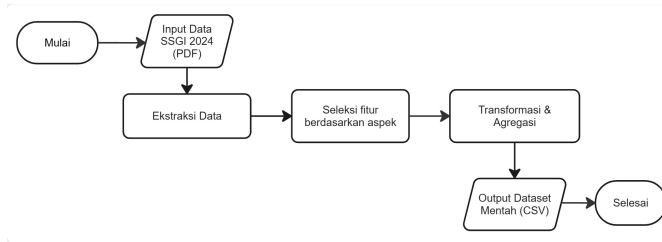
pemodelan *machine learning* regresi menggunakan algoritma *Random Forest* yang dipadukan dengan metode interpretasi SHAP untuk mengisi celah penelitian yang ada. Pendekatan ini diharapkan dapat memberikan wawasan baru mengenai faktor-faktor determinan utama penyebab *stunting* pada tingkat agregat provinsi menggunakan data SSGI 2024 masih terbatas. Tahap studi literatur juga dilakukan sebagai dasar pemilihan teori dan metode yang akan diterapkan dalam penelitian.

3.4 Penyiapan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari Laporan Publikasi Survei Status Gizi Indonesia (SSGI) tahun 2024 yang diterbitkan oleh Kementerian Kesehatan Republik Indonesia **kemenkes2025ssgi**. Laporan ini disajikan dalam format dokumen PDF dan memuat informasi statistik yang sangat komprehensif terkait status gizi balita serta faktor-faktor yang mempengaruhinya. Secara keseluruhan, SSGI 2024 berisikan ±5.189 tabel statistik yang menggambarkan indikator kesehatan ibu dan anak, pola asuh, akses pelayanan kesehatan, kondisi sanitasi dan lingkungan, serta karakteristik sosial ekonomi. Informasi dalam laporan SSGI 2024 juga disajikan dalam berbagai tingkat agregasi wilayah, mulai dari tingkat nasional, provinsi, hingga kabupaten/kota. Selain itu, tabel-tabel yang tersedia juga mencakup beragam kategori faktor determinan *stunting*, baik yang berkaitan dengan aspek kesehatan ibu dan balita, lingkungan dan sanitasi, maupun aspek sosial ekonomi dan akses layanan kesehatan.

Kompleksitas struktur data ini menyebabkan data tidak dapat langsung digunakan sebagai dataset penelitian kuantitatif tanpa melalui proses penyiapan data terlebih dahulu. Oleh karena itu, diperlukan suatu tahapan penyiapan data untuk mengekstraksi, menyeleksi, dan menyusun kembali informasi yang relevan agar diperoleh dataset terstruktur yang sesuai dengan tujuan penelitian. Alur penyiapan data yang dilakukan dalam penelitian ini ditunjukkan pada

Gambar ??.



Gambar 3.2 Alur Penyiapan Data

3.5 Pra-pemrosesan Data (*Data Preprocessing*)

Sebelum dilakukan pemodelan, data mentah diproses melalui tahapan:

1. **Pembersihan Data (*Data Cleaning*):** Mengonversi tipe data numerik dan menangani nilai yang hilang (*missing values*) untuk menjaga kualitas input.
2. **Seleksi Fitur:** Memilih variabel yang relevan berdasarkan domain pengetahuan dan membuang variabel *confounding* (seperti status gizi *underweight* dan *wasting*) agar fokus pada faktor determinan penyebab.
3. **Penyaringan Wilayah:** Menggunakan data dari 36 provinsi yang memiliki kelengkapan atribut dan mengecualikan provinsi pemekaran baru yang datanya belum stabil untuk menjaga validitas pelatihan model.

3.6 Pembangunan Model *Random Forest*

Tahap inti penelitian adalah penerapan algoritma *Random Forest Regressor*. Proses ini terdiri dari:

1. **Optimasi Parameter (*Hyperparameter Tuning*):** Menggunakan metode *Grid Search* untuk mencari kombinasi optimal pada parameter *n_estimators* (jumlah pohon) dan *max_depth* (kedalaman pohon).
2. **Validasi Model (*LOOCV*):** Menerapkan skema *Leave-One-Out Cross Validation*, dimana model dilatih dan diuji berulang kali sebanyak jumlah provinsi (N) untuk mencegah *overfitting* akibat terbatasnya jumlah

sampel.

3.7 Evaluasi Performa Model

Kinerja model diukur menggunakan metrik evaluasi regresi untuk melihat tingkat kesalahan prediksi:

1. **Mean Absolute Error (MAE):** Mengukur rata-rata kesalahan absolut antara nilai prediksi dan aktual.
2. **Root Mean Squared Error (RMSE):** Mengukur tingkat kesalahan dengan memberikan bobot lebih pada kesalahan besar.
3. **R^2 Score:** Mengukur seberapa baik model dapat menjelaskan variabilitas data stunting.

3.8 Interpretasi Hasil dengan SHAP

Untuk menjelaskan model *black-box*, digunakan metode SHAP yang menghasilkan:

1. **Interpretasi Global:** Mengidentifikasi fitur dominan yang mempengaruhi stunting di skala nasional melalui *Summary Plot*.
2. **Interpretasi Lokal:** Menganalisis faktor penyebab spesifik pada provinsi tertentu menggunakan *Waterfall Plot*, sehingga rekomendasi intervensi dapat lebih tepat sasaran.

3.9 Alat dan Bahan

3.9.1 Alat Penelitian

Perangkat keras dan lunak yang digunakan dalam penelitian ini adalah:

1. **Perangkat Keras:** Laptop (Processor Intel Core i5/AMD Ryzen 5, RAM 8GB, SSD).
2. **Software:** Visual Studio Code, Jupyter Notebook, Python 3.10.
3. **Library:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, SHAP.

3.9.2 Bahan Penelitian

Bahan utama adalah dataset SSGI 2024 yang terdiri dari variabel dependen (prevalensi stunting) dan variabel independen yang dikelompokkan menjadi:

- Praktik Pengasuhan & Kesehatan Ibu
- Pemberian Makan Bayi & Anak (PMBA)
- Pelayanan Kesehatan Balita
- Kesehatan Lingkungan (Air & Sanitasi)
- Faktor Sosial-Ekonomi (JKN, Bansos)

3.10 Ilustrasi Perhitungan Evaluasi

Evaluasi model regresi dihitung menggunakan persamaan berikut:

$$1. \text{ Mean Absolute Error (MAE): } MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (\text{Rumus 3.1})$$

$$2. \text{ Root Mean Squared Error (RMSE): } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (\text{Rumus 3.2})$$

$$3. \text{ Koefisien Determinasi (R}^2\text{): } R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (\text{Rumus 3.3})$$

Dimana y_i adalah nilai aktual, \hat{y}_i adalah nilai prediksi, dan \bar{y} adalah rata-rata nilai aktual.

3.11 Rancangan Pengujian

Pengujian dilakukan untuk membuktikan kinerja model dalam memprediksi stunting dan kemampuan SHAP dalam mengisolasi faktor determinan. Validitas dijamin melalui skema LOOCV dimana setiap provinsi diuji secara independen sebagai *unseen data* untuk memastikan objektivitas hasil.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Berisi hasil penelitian berdasarkan rancangan yang sudah dijelaskan pada Bab ??, terutama dari Subbab ??. Bagi yang membuat alat, jelaskan alat yang jadi dalam bentuk apa. Bagi yang membuat aplikasi, jelaskan aplikasi yang jadi dalam bentuk seperti apa. Jabarkan dalam bentuk pseudocode dan dijelaskan per bagian kodenya. Gunakan gambar dan tabel sebagai alat bantu menjelaskan hasil.

Contoh implementasi kode dapat ditulis menggunakan `\begin{lstlisting}`. Contoh kode dapat dilihat pada Kode ??.

Kode 4.1 Akuisisi Gambar

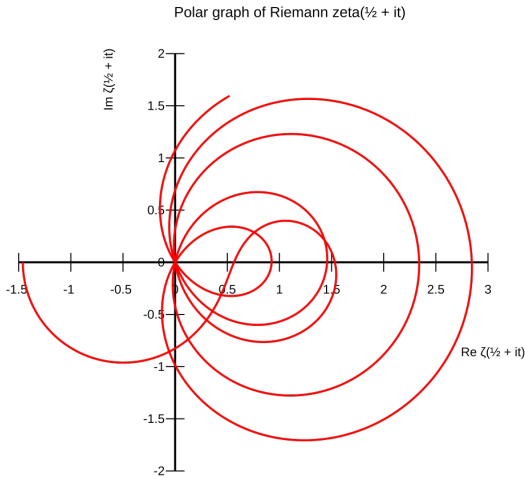
```
1 def process_dataset(dataset_path):
2     image_files = glob(os.path.join(dataset_path, '*.png'))
3     image_files.sort()
4     for image_file in image_files:
5         frame = cv2.imread(image_file)
6         if frame is None:
7             continue
8         frame_rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
9         cv2.imshow('Frame', frame)
10        if cv2.waitKey(1) & 0xFF == ord('q'):
11            break
12        cv2.destroyAllWindows()
13 def main():
14     datasets = get_all_dataset_folders(DATASET_ROOT)
15     for dataset in datasets:
16         process_dataset(dataset)
17         print("print string")
```

4.2 Hasil Pengujian

Berikan hasil pengujian berdasarkan rancangan & skenario yang sudah direncanakan sebelumnya pada Subbab ??.

Tabel 4.1 Data *dummy* Pengujian

Subjek	Hasil Prediksi (BPM)							GT
	F	NA	NO	RC	LC	M	C	
1	68	69	68	70	68	71	69	68
2	69	69	68	70	68	71	69	69
3	70	70	69	71	68	73	69	70
4	71	70	70	72	69	73	70	71
5	72	72	70	72	70	74	70	72



Gambar 4.1 Contoh Graf Pengujian

4.3 Analisis Hasil Penelitian

Berikan analisis hasil penelitian & pengujian, berupa data yang didapatkan dari penelitian & pengujian Tugas Akhir yang sudah anda kerjakan. Gunakan gambar dan tabel sebagai alat bantu menjelaskan analisis hasil. Data luaran penelitian yang dapat dianalisis berupa:

- 1. Hasil pengujian

2. Hasil kuesioner

3. Aplikasi yang dikembangkan

Analisis dapat membandingkan dengan hasil penelitian sebelumnya yang memiliki kemiripan topik.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berisi kesimpulan dari hasil dan pembahasan terkait penelitian yang dilakukan, dapat juga berupa temuan yang Anda dapatkan setelah melakukan penelitian atau analisis terhadap tugas akhir Anda. Memberikan jawaban dari poin pada subbab ?? dan ??.

5.2 Saran

Berisi saran mengenai aspek tugas akhir atau temuan yang dapat dikembangkan dan diperkaya di tugas akhir selanjutnya. Saran dapat berkaitan erat pada subbab ??.

LAMPIRAN

A Dataset

B Hasil Wawancara

C Rincian Kasus Uji