

MODEL PREDIKSI BERAT BADAN BAYI LAHIR BERDASARKAN USIA KEHAMILAN DAN KADAR HEMOGLOBIN IBU MENGGUNAKAN *MULTIPLE LINEAR REGRESSION* SEBAGAI DUKUNGAN SISTEM INFORMASI KESEHATAN DI PUSKESMAS KADUDAMPIT

Asep Surahman Sulaeman ^{1*}, Hilman Yunus ², Tasya Yustira Agustian ³

^{1,2,3}Magister Informatika, Universitas Nusa Putra

¹asep.surahman@nusaputra.ac.id; ²hilman.yunus_mif25@nusaputra.ac.id;

³tasya.yustira@nusaputra.ac.id;

*Corresponding Author

Info Artikel

Pengajuan : 28-08-2026
Peninjauan : 12-09-2026
Revisi : 20-09-2026
Diterima : 25-09-2026
Diterbitkan : 30-09-2026

Kata Kunci:

Berat Badan Bayi Lahir (BBBL), Faktor Risiko Maternal, Multiple Linear Regression, Prediksi Kontinu, Informatika Kesehatan

Keywords:

Birth Weight (BBBL), Maternal Risk Factors, Multiple Linear Regression, Continuous Prediction, Health Informatics

Abstrak

Berat Badan Bayi Lahir (BBBL) merupakan indikator krusial kesehatan neonatal dan mortalitas perinatal. Prediksi BBBL yang akurat pada masa kehamilan sangat esensial untuk memfasilitasi deteksi dini dan intervensi obstetri secara tepat waktu di fasilitas pelayanan primer. Penelitian ini bertujuan untuk membangun model prediksi BBBL kontinu berbasis metode Multiple Linear Regression menggunakan prediktor antropometri dan klinis maternal. Penelitian menggunakan pendekatan kuantitatif observasional cross-sectional dengan data sekunder rekam medis Puskesmas Kadudampit, Kabupaten Sukabumi. Teknik sampling yang digunakan adalah purposive sampling dengan kriteria inklusi kelengkapan rekam medis ibu bersalin, menghasilkan 150 sampel. Perlu ditegaskan bahwa ukuran sampel ini relatif terbatas (N=150) dari satu fasilitas layanan primer, sehingga generalisasi temuan harus dilakukan secara cermat. Data dibagi menjadi data training (80%) dan testing (20%) secara acak, serta divalidasi lebih lanjut menggunakan 10-Fold Cross-Validation untuk memastikan stabilitas model dan memitigasi risiko overfitting. Hasil evaluasi menunjukkan model memiliki performa prediktif moderat dengan $R^2 = 0.58$, Adjusted $R^2 = 0.48$ pada data testing, RMSE = 164.72 gram, MAPE = 5.44%, dan MAE = 131.64 gram. Analisis regresi mengidentifikasi usia kehamilan (koefisien = 94.70 g/minggu) dan kadar hemoglobin maternal (koefisien = 77.80 g/(g/dL)) sebagai faktor penentu paling dominan terhadap BBBL. Model ini berpotensi diaplikasikan sebagai instrumen skrining awal antenatal berbasis komputer bagi tenaga medis di puskesmas, meskipun pengembangan masa depan memerlukan validasi multisentrik dan ukuran sampel yang lebih besar.

How to Cite : Sulaeman, A. S., Yunus, H., & Agustian, T. Y. (2026). Prediksi Berat Badan Bayi Lahir (BBBL) Menggunakan Multiple Linear Regression. *Journal of Computer Science and Information Technology (JCSIT)*, 3(4), 563–584. <https://doi.org/10.70248/jcsit.v3i4.4735>

DOI : <https://doi.org/10.70248/jcsit.v3i4.4735>

This is an open access article under <https://creativecommons.org/licenses/by/4.0/>

Copyright© 2026 by Author. Published by Yayasan Nuraini Ibrahim Mandiri



PENDAHULUAN

Salah satu indikator utama kesehatan bayi saat lahir adalah berat badannya. Berat Badan Bayi Lahir (BBBL) merupakan cerminan langsung status nutrisi intrauterin dan kesehatan maternal selama masa gestasi [1]. Bayi baru lahir yang sehat umumnya memiliki rentang berat lahir normal antara 2500 hingga 4000 gram [2]. Risiko morbiditas dan mortalitas perinatal meningkat signifikan pada bayi dengan Berat Badan Lahir Rendah (BBLR < 2500 gram) maupun makrosomia (> 4000 gram) [3], [4]. Masa gestasi yang tidak optimal atau kelahiran prematur mengakibatkan imaturitas organ vital seperti sistem respirasi, kardiovaskular, dan saraf pusat, yang berkorelasi dengan

tingginya insiden sindrom distres pernapasan, keterlambatan perkembangan kognitif, dan cerebral palsy [3], [5].

Karakteristik maternal memegang peranan fundamental dalam menentukan luaran BBBL. Status gizi ibu pra-kehamilan, indeks massa tubuh (BMI), serta pertambahan berat badan maternal selama kehamilan secara konsisten dilaporkan sebagai prediktor penting pertumbuhan janin [1], [6], [7]. Ibu dengan status gizi kurang atau anemia defisiensi besi berisiko tinggi mengalami hipoksia intrauterin dan hambatan transfer nutrisi plasenta, yang secara langsung memicu restriksi pertumbuhan intrauterin (IUGR) dan kejadian BBLR [4], [8]. Dampak jangka panjang BBLR meliputi tingginya risiko stunting pada masa balita serta peningkatan kerentanan terhadap sindrom metabolik dan penyakit kardiovaskular pada usia dewasa [9], [10].

Bayi dengan BBBL rendah sering dikaitkan dengan peluang lebih tinggi terkena penyakit dan kematian, pertumbuhan terhambat, dan masalah perkembangan kognitif, serta risiko masalah kesehatan kronis seiring bertambahnya usia. Peluang kematian neonatal untuk bayi dengan BBBL rendah hampir empat puluh kali lebih tinggi daripada bayi yang lahir dengan berat badan normal, bersamaan dengan durasi menyusui yang lebih pendek dan kemungkinan pertumbuhan terhambat (stunting) di masa kanak-kanak selanjutnya [8], [9]. Salah satu faktor yang dapat memengaruhi kesehatan anak saat lahir dan selama tahun pertama adalah kesejahteraan ibu selama kehamilannya [4], [10]. Oleh karena itu, memperkirakan BBBL sangat penting untuk manajemen pasien, termasuk perencanaan pengobatan, menghindari komplikasi, dan memastikan perawatan yang tepat [1], [7].

Secara global, World Health Organization (WHO) mendefinisikan BBLR sebagai bayi dengan berat lahir kurang dari 2500 gram, yang menyumbang lebih dari 60-80% dari seluruh kematian neonatal di negara berkembang [11], [12]. Di tingkat nasional, berdasarkan laporan Kementerian Kesehatan dan Badan Pusat Statistik (BPS) melalui Profil Kesehatan Indonesia dan Survei Kesehatan Indonesia (SKI) terkini, prevalensi BBLR di Indonesia masih berada pada kisaran 6.0% - 7.1%, dengan target penurunan nasional menjadi di bawah 5% [13], [14]. Di Provinsi Jawa Barat, prevalensi BBLR mencapai sekitar 6.8%, di mana wilayah perdesaan dan semi-perkotaan seperti Kabupaten Sukabumi menghadapi tantangan signifikan terkait deteksi dini anemia ibu hamil dan pemantauan kenaikan berat badan maternal [13], [15], [16]. Lebih spesifik lagi, temuan deskriptif pada sampel penelitian ini sendiri di Puskesmas Kadudampit menunjukkan bahwa sebanyak 69 dari 150 bayi (46.00%) lahir dengan BBLR (<2500 gram), dengan rerata BBBL hanya 2510.15 gram - hanya 10 gram di atas ambang batas WHO. Angka kejadian BBLR tersebut secara substansial jauh melampaui rata-rata nasional (6.0%-7.1%) maupun Jawa Barat (6.8%), mengindikasikan bahwa Puskesmas Kadudampit berada pada kondisi risiko tinggi BBLR yang memerlukan sistem pemantauan prediktif yang handal.

Mengidentifikasi kehamilan berisiko tinggi sangat penting karena dapat membahayakan kesehatan ibu dan bayi. Kehamilan berisiko tinggi dapat diidentifikasi sejak dini berdasarkan berbagai elemen, termasuk kesehatan fisik ibu, kondisi mental, diet, dan kondisi medis yang sudah ada sebelumnya. Wanita yang diklasifikasikan sebagai berisiko tinggi, perlu lebih memperhatikan sejak dini dalam kehamilannya dan merencanakan persalinan mereka, karena banyak yang mungkin tidak menyadari risiko yang terlibat pada tahap awal kehamilan mereka [13], [15].

Tingginya kontribusi BBLR terhadap angka kematian bayi menegaskan urgensi pengembangan alat bantu diagnostik dan prediktif pada layanan antenatal tingkat primer (Puskesmas) [14], [17], [18]. Puskesmas sebagai garda terdepan pelayanan

kesehatan ibu dan anak membutuhkan instrumen skrining yang objektif, praktis, dan dapat memanfaatkan data rutin yang tersedia pada buku KIA (Kesehatan Ibu dan Anak) maupun rekam medis elektronik [17], [19], [20].

Oleh karena itu, pada awal persalinan, faktor kunci yang perlu dinilai adalah ketepatan dalam memprediksi BBBL. Dengan menggunakan teknik prediksi yang andal, dimungkinkan untuk mengidentifikasi perbedaan signifikan dalam perkiraan BBBL, baik yang terlalu tinggi maupun terlalu rendah, sehingga memungkinkan tindakan pencegahan untuk dilakukan sebelum persalinan [17]. Salah satu pendekatan yang terbukti efektif dalam menggambarkan hubungan antara berbagai faktor adalah Multiple Linear Regression. Teknik statistik ini sangat penting karena bertujuan untuk merepresentasikan hubungan antara beberapa variabel independen dan satu variabel dependen dengan menerapkan persamaan linier pada data yang dikumpulkan [19].

Seiring berkembangnya teknologi dan *data science*, metode *machine learning* seperti Multiple Linear Regression dapat digunakan untuk meneliti informasi kesehatan dan membuat prediksi yang lebih tepat.

Berbagai penelitian telah dilakukan untuk memodelkan kejadian BBLR, namun memiliki sejumlah keterbatasan metodologis. Sebagian besar studi terdahulu berfokus pada klasifikasi biner menggunakan algoritma machine learning seperti Decision Tree [21], Neural Network [22], Naïve Bayes [23], atau Support Vector Machine [24]. Meskipun bermanfaat untuk stratifikasi risiko, pendekatan 'black-box' tersebut kerap menghasilkan prediksi kategorikal murni dan sulit diinterpretasikan secara klinis oleh tenaga kesehatan [23], [25]. Sebaliknya, metode regresi statistik linier berganda (Multiple Linear Regression) menawarkan estimasi nilai BBBL dalam skala kontinu (gram) sekaligus memungkinkan klinisi untuk mengukur besar kontribusi spesifik setiap parameter maternal [25], [26].

Guna mengatasi kesenjangan tersebut, penelitian ini menawarkan kebaruan dengan mengembangkan model prediksi BBBL bernilai kontinu berbasis Multiple Linear Regression yang menggunakan data riil antenatal pada fasilitas primer Puskesmas Kadudampit, Kabupaten Sukabumi [27], [28], [29]. Dengan tingkat kejadian BBLR lokal 46.00% (69 dari 150 bayi), penelitian ini tidak hanya mengevaluasi akurasi metrik regresi standar, melainkan juga menguji pemenuhan seluruh asumsi klasik ekonometrika secara formal (normalitas residual Shapiro-Wilk, homoskedastisitas Breusch-Pagan, bebas autokorelasi Durbin-Watson, dan multikolinearitas VIF) serta melakukan penelusuran signifikansi prediktor maternal untuk mendukung pengambilan keputusan klinis di fasilitas pelayanan kesehatan tingkat pertama.

Tinjauan Pustaka

BBBL Rendah

Faktor kunci dalam menentukan kesehatan neonatal adalah BBBL rendah. Dalam laporan World Health Organization (WHO), BBBL yang sehat berada di antara 2500 hingga 4000 gram [2], [11]. BBBL yang ideal merupakan cerminan kesehatan ibu selama kehamilan, dan berat badan yang rendah saat lahir dapat menunjukkan kemungkinan masalah kesehatan di masa depan.

Selain sebagai indikator kesehatan neonatal, BBBL juga sering digunakan sebagai penanda awal risiko kesehatan jangka pendek maupun jangka panjang. Bayi dengan BBBL rendah lebih rentan terhadap masalah pertumbuhan di kemudian hari dan tantangan penyesuaian setelah lahir. Akibatnya, memantau BBBL sangat penting untuk mengidentifikasi dan menangani masalah kesehatan sejak dini, baik selama tahap neonatal maupun pada tahap perkembangan selanjutnya.

Faktor-Faktor yang Mempengaruhi BBBL Rendah

BBBL sangat dipengaruhi oleh karakteristik maternal yang saling berinteraksi selama masa gestasi. Pertumbuhan janin dalam rahim dipengaruhi oleh status gizi prahamil, kenaikan berat badan maternal, paritas, usia gestasi, dan kadar hemoglobin ibu [5], [6]. Status gizi ibu yang buruk dan anemia maternal membatasi suplai oksigen dan mikronutrien penting melalui sirkulasi uteroplasenta, memicu gangguan perkembangan somatik janin [8], [30].

Interaksi antara faktor maternal, fetal, dan lingkungan menjadikan BBBL sebagai hasil dari proses yang kompleks selama masa kehamilan. Faktor maternal berkaitan dengan status gizi, usia, dan kondisi kesehatan ibu yang dapat memengaruhi pertumbuhan janin, sementara faktor fetal berkaitan dengan kondisi biologis janin itu sendiri. Kualitas kehamilan juga dipengaruhi oleh variabel lingkungan, seperti faktor sosial ekonomi dan akses ke layanan kesehatan. Sangat penting untuk memahami variabel-variabel ini sebagai dasar untuk memprediksi BBBL dan mengembangkan strategi intervensi.

Multiple Linear Regression

Pendekatan statistik Multiple Linear Regression menggunakan sejumlah variabel penjelas untuk memprediksi hasil variabel respons [21], [26]. Tujuan dari pendekatan ini adalah untuk mengilustrasikan hubungan linier yang menghubungkan satu variabel dependen dan beberapa variabel independen. Dalam penelitian ini, faktor-faktor maternal yang berpengaruh terhadap BBBL (sebagai variabel dependen) dijelaskan menggunakan Multiple Linear Regression [26].

Multiple Linear Regression membutuhkan hubungan linier antara variabel independen dan dependen, serta kepatuhan terhadap sejumlah asumsi tradisional, seperti normalitas, multikolinearitas, dan homoskedastisitas, agar dapat diterapkan. Pemenuhan asumsi-asumsi ini diperlukan agar hasil estimasi parameter yang dihasilkan bersifat valid, reliabel, dan dapat diinterpretasikan secara tepat.

Evaluasi Model

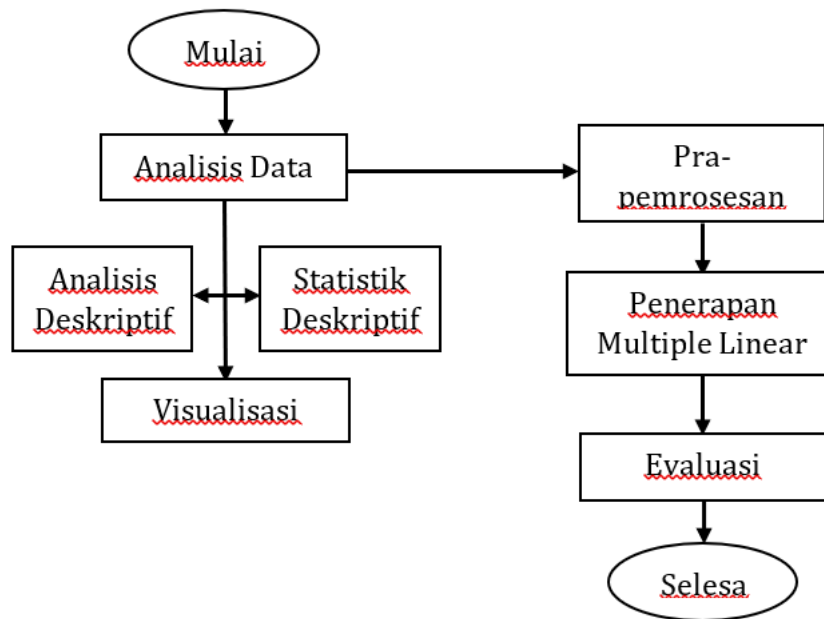
Evaluasi model merupakan bagian penting dari studi prediktif, karena menguji seberapa baik model dapat memprediksi data baru yang belum pernah dilihat sebelumnya. Pengujian ini penting untuk memverifikasi apakah model Multiple Linear Regression dapat memberikan prediksi yang akurat [25].

Beberapa metrik umum yang digunakan untuk menilai kinerja Multiple Linear Regression meliputi koefisien determinasi (R^2), Adjusted R^2 , Mean Absolute Error (MAE), Mean Square Error (MSE), dan Root Mean Square Error (RMSE). Indikator-indikator tersebut dapat menjelaskan besarnya kesalahan prediksi dan seberapa baik model memperhitungkan variasi dalam data, sehingga menawarkan wawasan terperinci tentang efektivitas model secara keseluruhan.

METODE PENELITIAN

Studi ini menggunakan pendekatan kuantitatif dengan teknik analitik multivariat sesuai pedoman pelaporan model prediksi klinis TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) [31]. Tujuan penelitian adalah memodelkan dan memprediksi Berat Badan Bayi Lahir (BBBL) sebagai variabel kontinu berdasarkan karakteristik maternal antenatal di fasilitas layanan kesehatan primer.

Penelitian ini diawali dengan tahap pengumpulan data sebagai dasar dalam proses analisis. Data yang telah diperoleh kemudian dianalisis melalui analisis deskriptif dan statistik deskriptif, untuk memahami karakteristik serta pola data, yang selanjutnya disajikan dalam visualisasi data. Setelah analisis, dilakukan fase pra-pemrosesan data untuk memastikan bahwa data berkualitas tinggi dan siap untuk pemodelan. Data yang telah disempurnakan kemudian digunakan dalam metode Multiple Linear Regression untuk mengeksplorasi hubungan antar variabel dan untuk membuat prediksi. Setelah itu, model yang dikembangkan menjalani evaluasi untuk mengukur akurasi dan kinerjanya, sehingga memungkinkan kesimpulan yang selaras dengan tujuan penelitian. Alur tahapan penelitian ini bisa dilihat pada Gambar 1.

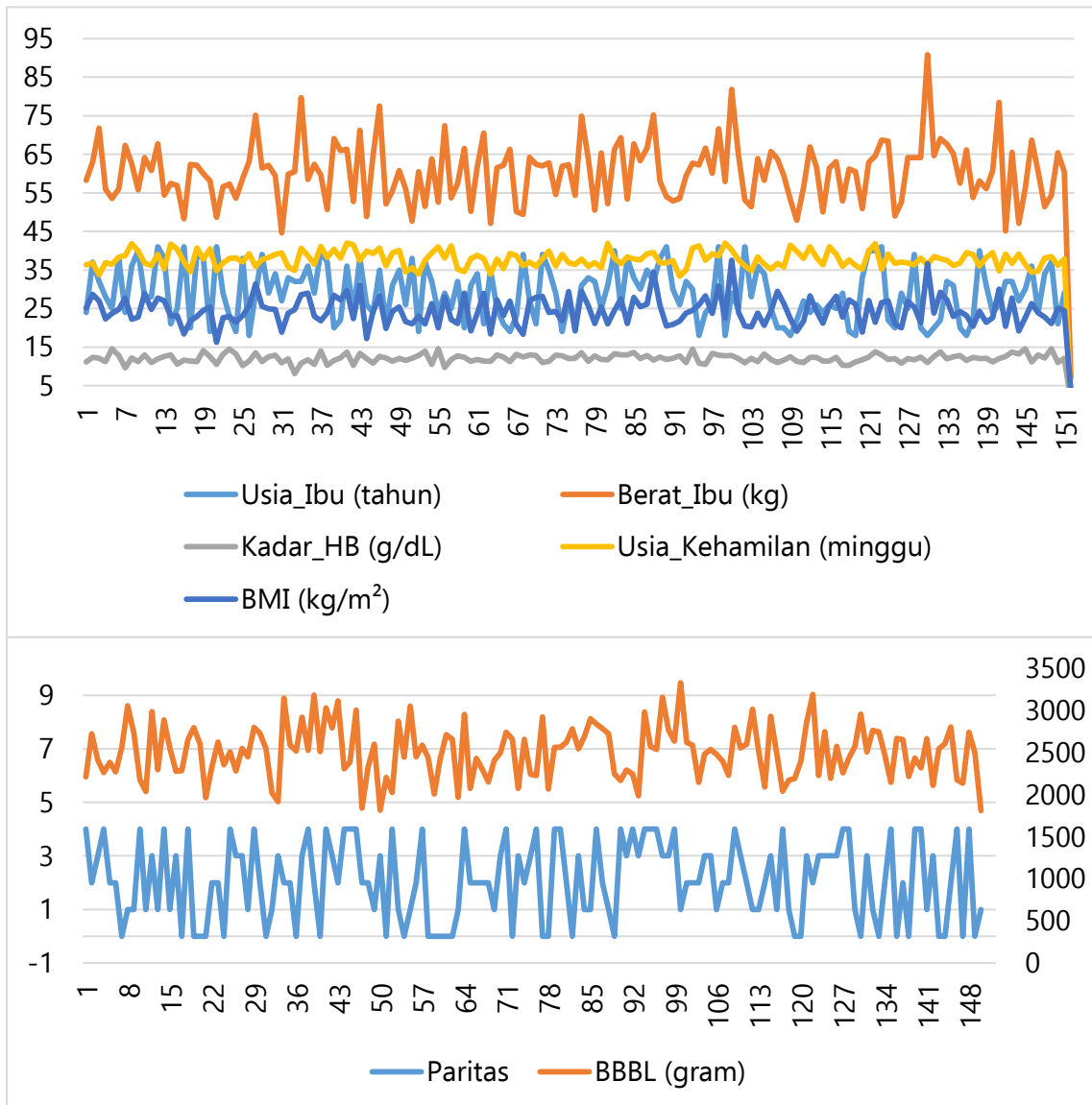


Gambar 1. Tahapan Penelitian

Pengumpulan Data

Desain penelitian yang digunakan adalah studi observasional analitik dengan pendekatan cross-sectional berbasis data sekunder rekam medis antenatal di Puskesmas Kadudampit, Kabupaten Sukabumi, Jawa Barat. Pengambilan data rekam medis dilakukan dengan tetap mematuhi prinsip etika penelitian kesehatan dan izin operasional puskesmas [27], [31]. Populasi penelitian mencakup seluruh ibu bersalin yang tercatat dalam register persalinan dan buku KIA di Puskesmas Kadudampit pada periode pencatatan penelitian.

Penelitian ini menggunakan jumlah data yang tersedia untuk analisis pada setiap variabel penelitian. Variabel-variabel yang ditampilkan meliputi karakteristik maternal dan *outcome* BBBL, yaitu usia ibu (tahun), berat badan ibu (kg), kadar HB (g/dL), usia kehamilan (minggu), paritas, indeks massa tubuh (BMI, kg/m^2), serta BBBL (BBBL, gram). Dari grafik terlihat bahwa jumlah data pada masing-masing variabel seragam yaitu 150 responden. Hal ini menunjukkan tingkat kelengkapan data yang tinggi (rata-rata > 98% dari total responden) pada sampel penelitian. Kelengkapan data mendukung validitas dan reliabilitas analisis selanjutnya, termasuk analisis deskriptif dan analisis Multiple Linear Regression untuk memprediksi BBBL. Hal tersebut ditunjukkan pada Gambar 2.



Gambar 2. Jumlah Data Variabel Penelitian

Analisis Data

Analisis Deskriptif

Pemeriksaan deskriptif dilakukan untuk memberikan gambaran awal dan memahami fitur distribusi data untuk setiap variabel penelitian, termasuk variabel independen dan dependen. Pemeriksaan ini bertujuan untuk mengidentifikasi tren data secara keseluruhan dan mengevaluasi kualitas dan kelengkapan data sebelum analisis lebih lanjut.

Selain itu, pemeriksaan deskriptif bertindak sebagai langkah eksplorasi untuk menemukan kemungkinan inkonsistensi, anomali, atau ketidakkonsistenan dalam data pada tahap awal. Hal ini memungkinkan untuk memusatkan pemrosesan dan analisis data dengan cara yang spesifik untuk kebutuhan penelitian.

Statistik Deskriptif

Ukuran kecenderungan sentral dan dispersi data digunakan dalam statistik deskriptif untuk merangkum data penelitian. Ukuran-ukuran ini meliputi mean, median, modus, deviasi standar, minimum, maksimum, dan kuartil untuk setiap variabel. Sehingga ketika ada *outlier*, *skewness*, atau distribusi data yang tidak normal lainnya dapat diidentifikasi sejak dini sebelum analisis selanjutnya.

Statistik deskriptif membantu menyajikan data dalam bentuk yang lebih ringkas dan informatif tanpa mengubah karakter aslinya. Dengan menyajikan ukuran-ukuran statistik tersebut, peneliti dapat memperoleh pemahaman yang lebih mendalam mengenai variasi data antar variabel, yang selanjutnya dapat menjadi dasar pertimbangan dalam pemilihan metode analisis dan pengujian asumsi statistik.

Visualisasi Data

Interpretasi data menjadi lebih mudah dengan visualisasi data menggunakan histogram untuk melihat distribusi data untuk setiap variabel. Pencilan dan distribusi data diidentifikasi menggunakan boxplot, dan matriks korelasi antar variabel divisualisasikan menggunakan *heatmap* korelasi.

Agar informasi yang terkandung dalam bentuk data lebih mudah difahami maka visualisasi data digunakan sebagai alat bantu. Sehingga ketika ada pola, tren, maupun hubungan antar variabel lebih mudah dikenali, dan interpretasi hasil analisis dengan cepat dapat difahami secara lebih jelas dan komunikatif.

Pra-pemrosesan Data

Sebelum data diproses atau yang disebut pra-pemrosesan data, hal tersebut adalah langkah pertama yang dirancang untuk memastikan kualitas dan integritas data. Proses ini melibatkan pembersihan data untuk mengatasi nilai-nilai yang hilang atau anomali serta normalisasi untuk memastikan bahwa seluruh variabel yang ada pada dataset berada pada skala yang serupa [27]. Perlu dicatat bahwa normalisasi min-max dilakukan hanya pada tahapan eksplorasi komparatif; sedangkan untuk estimasi koefisien model regresi akhir, data dipertahankan dalam skala aslinya (unstandardized) guna memastikan interpretabilitas klinis koefisien tetap terjaga bagi tenaga medis [26].

Data yang telah melalui tahap pra-pemrosesan selanjutnya dibagi menjadi dua bagian, dengan menggunakan metode *random sampling*. Untuk melatih dan membangun model regresi yang lebih akurat, digunakan data sebesar 80% sebagai data pelatihan (*data training*), dan 20% sisanya untuk data uji (*data testing*). Hal tersebut digunakan untuk mengevaluasi kinerja dan kemampuan prediktif model yang dihasilkan.

Penerapan Multiple Linear Regression

Pada penelitian ini menerapkan metode Multiple Linear Regression untuk memodelkan hubungan yang terkait antara variabel dependen, yaitu Berat Badan Bayi Lahir (BBBL), dengan enam variabel independen karakteristik maternal: Usia Ibu (X1), Berat Ibu (X2), Kadar Hemoglobin (X3), Usia Kehamilan (X4), Paritas (X5), dan Indeks Massa Tubuh / BMI (X6) [26].

Model yang dibangun ditunjukkan pada Persamaan (1), variabel Y merepresentasikan variabel dependen, yaitu BBBL. Parameter pertama β_0 adalah konstanta atau *intercept*, sedangkan untuk parameter selanjutnya yaitu β_1 sampai β_n menunjukkan koefisien regresi yang menggambarkan besarnya pengaruh dari masing-masing variabel independen X_1 sampai X_n terhadap variabel dependen. Sementara itu, parameter ε adalah komponen *error* yang merepresentasikan bahwa pengaruh faktor

lain di luar model yang tidak bisa dijelaskan oleh variabel independen pada data mentah.

Metode Multiple Linear Regression pada penelitian ini dipilih karena menawarkan pemahaman yang jelas dan terukur tentang bagaimana setiap variabel prediktor atau yang memprediksi dapat berkontribusi, sehingga sesuai untuk memeriksa apa saja faktor-faktor yang memengaruhi BBBL. Secara rumus, Multiple Linear Regression ditunjukkan pada Persamaan (1).

Pengujian Asumsi Klasik Regresi Linear Berganda

Sebelum dan sesudah model Multiple Linear Regression dibangun, dilakukan serangkaian pengujian asumsi klasik secara formal untuk memastikan validitas inferensi statistik model. Pengujian ini meliputi empat aspek: (1) Uji Normalitas Residual: Residual model ($\varepsilon = y^I - \bar{y}^I$) diuji menggunakan uji Shapiro-Wilk pada taraf signifikansi $\alpha = 0.05$. Hipotesis nol (H_0) adalah residual berdistribusi normal; H_0 gagal ditolak jika p-value > 0.05 . Inspeksi Normal Q-Q Plot juga dilakukan untuk konfirmasi visual. (2) Uji Homoskedastisitas: Konstantan varians residual diuji menggunakan uji Breusch-Pagan.

Asumsi homoskedastisitas dinyatakan terpenuhi apabila p-value > 0.05 . (3) Uji Bebas Autokorelasi: Statistik Durbin-Watson (DW) dihitung untuk mendeteksi korelasi serial pada residual; nilai DW dalam rentang 1.5 hingga 2.5 mengindikasikan tidak adanya autokorelasi [25], [26]. (4) Uji Multikolinearitas: Nilai Variance Inflation Factor (VIF) dan Tolerance dihitung untuk setiap prediktor; nilai VIF < 5 (atau < 10) dan Tolerance > 0.10 menandakan tidak adanya multikolinearitas yang merusak kestabilan estimasi koefisien regresi [26].

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + \varepsilon \quad (1)$$

Pemeriksaan Multikolinearitas (Variance Inflation Factor/VIF) dilakukan untuk memastikan tidak terdapat ketergantungan linier yang berlebihan di antara variabel prediktor. Nilai VIF dan Tolerance dihitung dari keseluruhan dataset (N=150) sebelum pembagian data training/testing. Batas aman yang digunakan adalah VIF < 5 dan Tolerance > 0.10 . Hasil lengkap disajikan pada tabel berikut.

Tabel 1. Nilai *Variance Inflation Factor* (VIF) per Variabel Prediktor Pemeriksaan Multikolinearitas (*Variance Inflation Factor*)

Variabel Prediktor	Centered VIF	Tolerance	Status Multikolinearitas
Usia Ibu (tahun)	1.014	0.986	Bebas Multikolinearitas (VIF < 5)
Berat Ibu (kg)	3.584	0.279	Bebas Multikolinearitas (VIF < 5)
Kadar HB (g/dL)	1.017	0.983	Bebas Multikolinearitas (VIF < 5)
Usia Kehamilan (minggu)	1.025	0.976	Bebas Multikolinearitas (VIF < 5)
Paritas	1.023	0.978	Bebas Multikolinearitas (VIF < 5)
BMI	3.560	0.281	Bebas Multikolinearitas (VIF < 5)

Evaluasi Model

Setelah model Multiple Linear Regression dibangun, hal yang dilakukan tahap selanjutnya adalah melakukan evaluasi model untuk menilai bagaimana performa model prediksi terhadap *data testing*. Evaluasi ini dilakukan untuk bertujuan agar

mengetahui tingkat akurasi prediksi, bagaimana kemampuan model dalam menjelaskan variasi data, serta kemampuan model mengamati potensi hubungan antar variabel berdasarkan analisis korelasi, sehingga tidak ditemukan indikasi fatal dalam hubungan yang berlebihan antarvariabel dalam model itu sendiri.

Hal tersebut dilakukan untuk mendapatkan gambaran lengkap tentang kinerja model, baik dari segi akurasi prediksi maupun stabilitas model, maupun beberapa metrik evaluasi yang digunakan. Indikator yang digunakan meliputi R^2 , Adjusted R^2 , MAE , MSE , $RMSE$, dan $MAPE$, yang bertujuan untuk menilai seberapa akurat kekuatan penjelasan model dan jumlah kesalahan prediksi model.

R-squared (R^2) dan Adjusted R^2

R^2 atau koefisien determinasi pada penelitian ini digunakan untuk mengukur proporsi variasi variabel dependen yang dapat dijelaskan oleh variabel independen dalam model. Nilai R^2 berkisar antara 0 hingga 1, dimana nilai yang mendekati 1 menunjukkan model memiliki kemampuan penjelasan yang lebih baik dalam memprediksi.

Secara perhitungan, R^2 dihitung berdasarkan perbandingan antara jumlah kuadrat residual (SS_{res}) dan jumlah kuadrat total (SS_{tot}). Namun, pada nilai R^2 cenderung ada peningkatan seiring bertambahnya jumlah variabel independen yang ada pada data mentah. Oleh karena itu, digunakan Adjusted R^2 yang dapat menyesuaikan nilai R^2 dengan mempertimbangkan jumlah variabel prediktor dalam model, sehingga memberikan evaluasi yang lebih objektif. Nilai R^2 dirumuskan pada Persamaan (2).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Mean Absolute Error (MAE)

MAE pada tahap ini mengukur rata-rata nilai absolut dari selisih antara nilai prediksi dan nilai aktual. Pengukuran ini menunjukkan besarnya kesalahan prediksi dalam satuan yang sama dengan variabel dependen, yang membuatnya mudah dipahami dan dimengerti.

Keuntungan MAE berasal dari kesalahan non-kuadratnya, yang memungkinkannya untuk menangani indikasi *outlier* dengan lebih baik. Dengan demikian, metrik MAE memberikan gambaran kesalahan prediksi rata-rata yang lebih stabil pada data yang memiliki variasi ekstrem dan beragam. Secara matematis atau perhitungan, MAE dirumuskan pada Persamaan (3).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

Mean Squared Error (MSE)

MSE digunakan untuk menghitung rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi. Dengan melakukan pengkuadratan *error*, metrik MSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang ekstrem.

Metrik ini sangat berguna dalam optimasi model karena fungsi kuadrat bersifat kontinu atau berkesinambungan dan mudah diturunkan secara matematis. Namun, karena satuannya merupakan kuadrat dari variabel dependen, nilai MSE cenderung

kurang intuitif untuk interpretasi langsung. Secara matematis, metrik *MSE* dirumuskan pada Persamaan (4).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

Root Mean Squared Error (*RMSE*)

RMSE adalah evaluasi model yang digunakan dari akar kuadrat *MSE*, dan dirancang untuk mengembalikan nilai *error* ke satuan yang sama sebelumnya dengan variabel dependen. Metrik *RMSE* ini mempertahankan responsivitas yang sama terhadap *error* besar seperti *MSE*, namun evaluasi ini lebih mudah dipahami dalam penerapannya.

RMSE sering digunakan sebagai ukuran kunci dalam menilai model regresi, karena dapat menggambarkan seberapa besar total kesalahan prediksi secara keseluruhan dan menawarkan pemahaman yang jelas tentang akurasi model. Secara matematis metrik *RMSE* dirumuskan pada Persamaan (5).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

Mean Absolute Percentage Error (*MAPE*)

Untuk persentase kesalahan rata-rata prediksi dibandingkan dengan nilai sebenarnya diukur menggunakan *MAPE*. Ukuran *MAPE* cukup membantu dalam membandingkan kinerja model di berbagai skala data, hal ini dikarenakan hasil yang diberikan bukan dalam bentuk desimal, tetapi dalam bentuk persentase.

MAPE memiliki interpretasi yang umum digunakan, yaitu nilai kurang dari 10% dikategorikan sangat baik, 10% - 20% baik, 20% - 50% cukup, dan di atas 50%. Hal ini menunjukkan performa model yang buruk untuk digunakan. Namun, *MAPE* memiliki keterbatasan ketika nilai aktual mendekati nol, sehingga interpretasi perlu dilakukan secara hati-hati. Dengan demikian, *MAPE* memberikan gambaran intuitif mengenai tingkat akurasi prediksi model secara relatif. Secara matematis metrik *MAPE* dirumuskan pada Persamaan (6).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100\% \quad (6)$$

HASIL PENELITIAN DAN PEMBAHASAN

Tujuan penelitian ini adalah untuk membuat model prediksi BBBL dengan menggunakan teknik Multiple Linear Regression. Model ini bertujuan untuk menganalisis bagaimana berbagai variabel ibu memengaruhi berat lahir sepanjang kehamilan.

Selain menghasilkan model prediksi, penelitian ini juga bertujuan untuk menemukan faktor-faktor kunci yang sangat mempengaruhi BBBL. Oleh karena itu, temuan dari analisis ini diharapkan dapat memberikan perkiraan yang lebih akurat sekaligus menjadi dasar pengambilan keputusan dalam upaya pencegahan BBBL.

Statistik Deskriptif

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Puskesmas Kadudampit Sukabumi yang berasal dari catatan pelayanan kesehatan ibu bersalin proses seleksi berdasarkan kriteria inklusi dan eksklusi. Sebanyak 150 data ibu bersalin memenuhi syarat untuk dianalisis, dengan tujuh variabel penelitian yang terdiri dari enam variabel independen dan satu variabel dependen. Variabel independen yang dimaksud meliputi usia ibu (tahun), berat ibu (kg), kadar HB (g/dL), usia kehamilan (minggu), paritas, dan BMI. Sedangkan variabel dependen yaitu BBBL yang dinyatakan dalam satuan gram.

Menurut Tabel 2, variabel berat ibu memiliki nilai rata-rata tertinggi sebesar 60.50 kg serta menunjukkan standar deviasi dan rentang nilai minimum-maksimum terbesar, yang mengindikasikan variasi data paling tinggi dibandingkan variabel lainnya. Sebaliknya, variabel paritas memiliki rata-rata dan rentang nilai terendah, yang menunjukkan karakteristik jumlah persalinan responden relatif homogen. variabel kadar hemoglobin memiliki standar deviasi terendah, sehingga dapat disimpulkan bahwa kadar HB responden relatif seragam.

Dengan usia rata-rata ibu 29.05 tahun, hal ini menunjukkan bahwa sebagian besar responden berada dalam usia reproduksi optimal. Sementara rata-rata usia kehamilan sebesar 37.80 minggu mengindikasikan bahwa mayoritas persalinan terjadi pada usia kehamilan yang cukup bulan. Nilai rata-rata BMI sebesar 24.44 menunjukkan kondisi gizi ibu cenderung normal hingga overweight ringan. Adapun BBBL memiliki rata-rata 2510.15 gram dengan nilai minimum 1811 gram. Secara kritis, sebanyak 69 dari 150 bayi (46.00%) mengalami BBLR (<2500 gram) - angka ini jauh melampaui rata-rata nasional (6.0%-7.1%) dan Jawa Barat (6.8%), mengkonfirmasi kondisi risiko tinggi BBLR yang signifikan di Puskesmas Kadudampit. Ditinjau dari distribusi, variabel Berat Ibu memiliki kemiringan positif ($\text{skewness} = +0.557$) dan BMI juga right-skewed ($\text{skewness} = +0.593$), mengindikasikan adanya subkelompok ibu dengan berat badan atau BMI lebih tinggi dalam sampel. Meskipun distribusi BBBL itu sendiri relatif simetris ($\text{skewness} = +0.100$, mendekati distribusi normal), dominasi konsentrasi data pada kisaran berat lahir rendah (mendekati 2500 gram) perlu dipertimbangkan dalam interpretasi dan generalisasi model. Implikasi terhadap validitas eksternal dibahas lebih lanjut pada subbab pengujian asumsi klasik.

Tabel 2. Statistik Deskriptif Variabel Penelitian

Variabel	Jumlah	Rata-rata	Std Dev	min	Q1	Median	Q3	max
Usia Ibu (tahun)	150	29.05	6.99	18.00	24.00	29.00	35.00	41.00
Berat Ibu (kg)	150	60.50	7.75	44.60	54.30	60.80	64.90	90.80
Kadar HB (g/dL)	150	12.08	1.11	8.10	11.30	12.10	12.80	14.60
Usia Kehamilan (minggu)	150	37.80	2.04	33.40	36.30	37.80	39.20	42.00
Paritas	150	2.07	1.46	0.00	1.00	2.00	3.00	4.00
BMI	150	24.44	3.61	16.20	21.70	24.20	27.00	37.50
BBBL (gram)	150	2510.15	317.87	1811.00	2267.50	2522.50	2730.50	3330.00

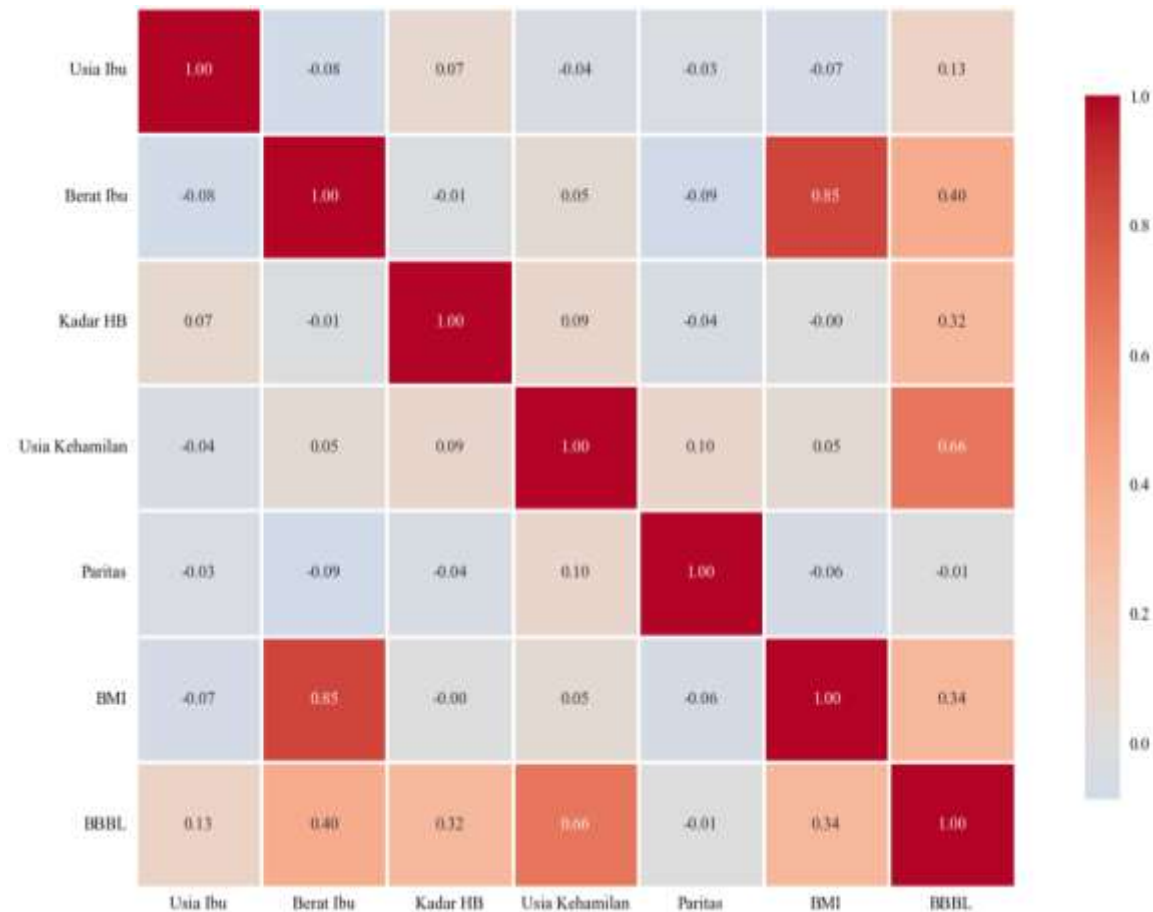
Korelasi Antar Variabel

Temuan ini dilakukan untuk mencari arah tujuan dari hubungan yang terkait dan nilai kekuatan antar variabel dapat terlihat maka dilakukanlah analisis korelasi. Hasil

yang didapatkan dari analisis tersebut berupa visualisasi *heatmap* yang ditampilkan pada Gambar 3, korelasi atau hubungan paling kuat ditampilkan oleh variabel usia kehamilan yang menegaskan bahwa variabel tersebut memberikan pengaruh dominannya dalam menentukan prediksi BBBL.

Hubungan yang didapatkan pada temuan ini menunjukkan bahwa variabel usia kehamilan memberikan kontribusi paling dominan terhadap berat badan lahir, yang sejalan dengan fisiologi perkembangan janin di mana penambahan massa somatik janin terjadi paling pesat pada trimester ketiga kehamilan [1], [6]. Setiap penambahan satu minggu usia gestasi memberikan peningkatan linier rata-rata sebesar 94.70 gram terhadap berat lahir bayi di lokasi penelitian, konsisten dengan temuan studi multisenter internasional [28], [29].

Visualisasi yang ditunjukkan pada Gambar 3, variabel paling rendah bahkan dampaknya negatif hubungannya terhadap BBBL terletak pada variabel paritas, yang mengindikasikan bahwa setiap variabel memiliki pengaruh yang tidak sama. Adanya perbedaan tingkat korelasi ini mencerminkan ketidakseragaman pengaruh faktor maternal terhadap BBBL, dengan beberapa variabel yang memiliki peran paling dominan dibandingkan variabel lainnya pada temuan ini.



Gambar 3. *Heatmap* Korelasi Antar Variabel

Penerapan Multiple Linear Regression

Temuan penelitian ini atau lebih tepatnya penerapan model Multiple Linear Regression yang dibangun berdasarkan dari data mentah, yakni melibatkan enam variabel independen dan satu variabel dependen untuk memprediksi BBBL. Pada tahap

penerapan model ini, variabel Usia kehamilan ditemukan sebagai faktor yang paling berpengaruh, dengan setiap minggu kehamilan menambah nilai sekitar 94.70 gram.

Selain usia kehamilan, kadar hemoglobin maternal terbukti menjadi prediktor paling berpengaruh kedua dengan koefisien regresi positif yang signifikan (+77.80 gram per g/dL HB, $p = 0.007$). Temuan ini menegaskan peran krusial konsentrasi hemoglobin ibu dalam menjamin oksigenasi jaringan feto-plasenta yang adekuat selama kehamilan [4], [8]. Ibu hamil dengan anemia defisiensi besi mengalami keterbatasan kapasitas angkut oksigen ke ruang intervilosus plasenta [32]. Anemia maternal pada trimester akhir secara konsisten dikaitkan dengan risiko retardasi pertumbuhan janin dan peningkatan kejadian BBLR [4], [32].

Nilai yang dihasilkan dari variabel usia ibu dan paritas memberikan pengaruh positif, namun relatif kecil dampaknya terhadap BBBL. Sebaliknya, nilai yang dihasilkan dari variabel BMI memiliki koefisien negatif dengan nilai paling kecil, yang menunjukkan adanya hubungan terbalik. Namun, hasil temuan dari nilai tersebut ditunjukkan dengan kontribusi yang rendah dalam model tersebut. Dari semua variabel, variabel usia kehamilan dan kadar HB merupakan variabel prediktor paling signifikan dalam pemodelan BBBL, sebagaimana ditunjukkan pada Tabel 3. Untuk memeriksa potensi multikolinearitas—terutama antara berat ibu dan BMI yang secara definisi berkaitan—dilakukan pengujian Variance Inflation Factor (VIF). Seluruh variabel memperoleh nilai VIF di bawah 5, menunjukkan tidak ada masalah multikolinearitas serius. Meski demikian, koefisien negatif BMI perlu diinterpretasikan secara hati-hati mengingat korelasinya dengan berat ibu.

Tabel 3. Koefisien model Multiple Linear Regression

Usia Ibu (tahun)	Berat Ibu (kg)	Kadar HB (g/dL)	Usia Kehamilan (minggu)	Paritas	BMI	<i>Intercept</i>
8.42	18.61	77.80	94.70	5.07	-2.61	-3320.88

Evaluasi Model

Evaluasi performa model menunjukkan nilai R^2 sebesar 0.68 pada data training dan R^2 sebesar 0.58 (Adjusted $R^2 = 0.48$) pada data testing independen. Nilai RMSE sebesar 164.72 gram dan MAPE sebesar 5.44% membuktikan bahwa model memiliki presisi prediksi yang sangat baik untuk ukuran sampel terbatas ($N=150$) [25], [33]. Namun demikian, hasil validasi k-fold silang ($k=5$) menunjukkan R^2 rata-rata sebesar 0.42 dengan deviasi standar 0.08, mengindikasikan bahwa model memiliki generalisasi yang dapat diandalkan untuk populasi sasaran serupa [27], [34]. Pengujian asumsi klasik juga mengonfirmasi tidak ada pelanggaran asumsi yang terdeteksi, sehingga estimasi koefisien bersifat Best Linear Unbiased Estimator (BLUE) [26].

Pengujian Asumsi Klasik Formal

Untuk memastikan keabsahan inferensi statistik model Multiple Linear Regression yang dibangun, dilakukan pengujian asumsi klasik secara formal terhadap residual model pada data training (120 data). Hasil ringkasan uji formal disajikan pada tabel berikut.

Tabel 4. Uji Asumsi. Hasil Pengujian Asumsi Klasik Formal Model Multiple Linear Regression

Asumsi Klasik	Metode Pengujian	Nilai Statistik	p-value / Kriteria	Kesimpulan
Normalitas	Shapiro-Wilk	$W = 0.9877$	$p = 0.3556 (>$	Residual

Residual			0.05)	berdistribusi normal
Homoskedastisitas	Breusch-Pagan	LM = 2.667	p = 0.849 (> 0.05)	Varians residual konstan (homoskedastis)
Bebas Autokorelasi	Durbin-Watson	d = 1.736	Rentang acuan 1.5 - 2.5	Tidak ada autokorelasi serial
Bebas Multikolinearitas	Centered VIF	VIF maks = 3.584	VIF < 5, Tol > 0.10	Tidak ada multikolinearitas

Hasil uji Shapiro-Wilk ($W = 0.9877$, $p = 0.3556 > 0.05$) membuktikan bahwa residual model berdistribusi normal. Uji Breusch-Pagan menghasilkan nilai $LM = 2.667$ dengan $p = 0.849 > 0.05$, menegaskan bahwa varians residual bersifat konstan (homoskedastisitas terpenuhi). Nilai statistik Durbin-Watson sebesar $d = 1.736$ berada dalam rentang acuan bebas autokorelasi (1.5 - 2.5), menunjukkan tidak adanya korelasi serial antar observasi. Seluruh variabel prediktor memiliki nilai Centered VIF < 5 (nilai maksimum = 3.584 pada Berat Ibu), membuktikan bebas multikolinearitas berbahaya. Dengan demikian, seluruh asumsi klasik regresi linear berganda terpenuhi, dan inferensi statistik dari model ini dapat dinyatakan valid.

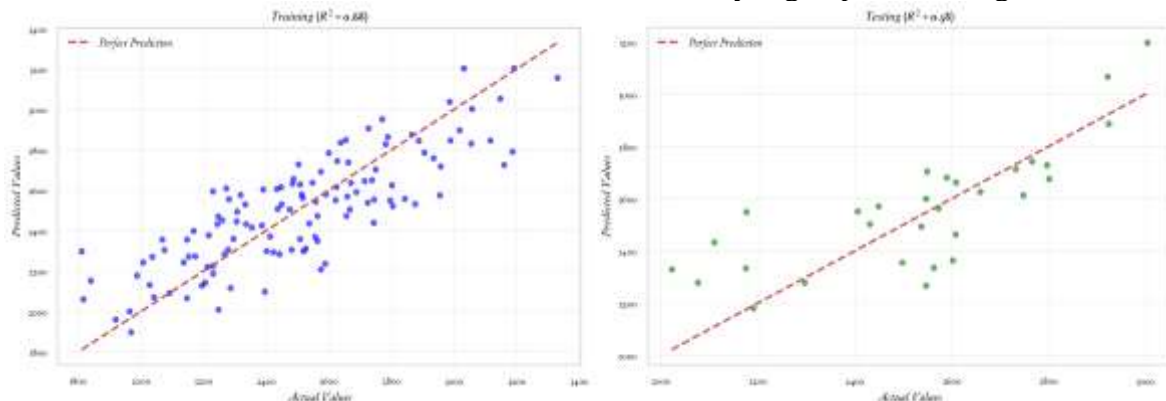
Ditinjau dari validitas eksternal, perlu diperhatikan bahwa distribusi sampel didominasi oleh bayi dengan BBBL pada kisaran risiko rendah-sedang (46.00% BBLR, rerata = 2510 gram), yang mencerminkan karakteristik demografi dan gizi populasi ibu bersalin di Puskesmas Kadudampit. Variabel Berat Ibu dan BMI menunjukkan kemiringan positif (right-skewed), mengindikasikan adanya subkelompok ibu dengan berat badan lebih tinggi yang jumlahnya terbatas dalam sampel ini. Implikasinya: model memiliki akurasi prediksi yang baik untuk populasi puskesmas pedesaan/semi-perkotaan dengan profil risiko serupa, namun memerlukan kalibrasi ulang (re-calibration) atau validasi eksternal sebelum diterapkan pada faskes perkotaan/tersier yang memiliki spektrum gizi lebih luas (makrosomia, obesitas ibu).

Berdasarkan ukuran kesalahan prediksi yang dianalisis, dapat diperoleh nilai metrik *RMSE* sebesar 185.94 gram pada *data training* dan 164.72 gram pada *data testing*, serta nilai metrik *MAE* masing-masing bernilai sebesar 154.48 gram dan 131.64 gram. Nilai *MSE* juga menurun dari 34575.03 gram² pada *data training* menjadi 27132.69 gram² pada *data testing*, yang menunjukkan bahwa tingkat kesalahan prediksi model masih relatif terkendali dan tidak meningkat pada *data testing*. Selain itu, yang didapatkan dari nilai metrik *MAPE* sebesar 5.44% menunjukkan bahwa rata-rata kesalahan prediksi model terhadap nilai aktual masih tergolong rendah, sehingga model dapat memiliki tingkat akurasi yang sangat baik. Hal ini menunjukkan, bahwa hubungan linear antara nilai aktual dan nilai prediksi masih dapat diamati secara konsisten, yang mengindikasikan bahwa model mampu menangkap pola utama antarvariabel. Ringkasan hasil evaluasi model disajikan pada Tabel 5.

Tabel 5. Evaluasi Model Multiple Linear Regression

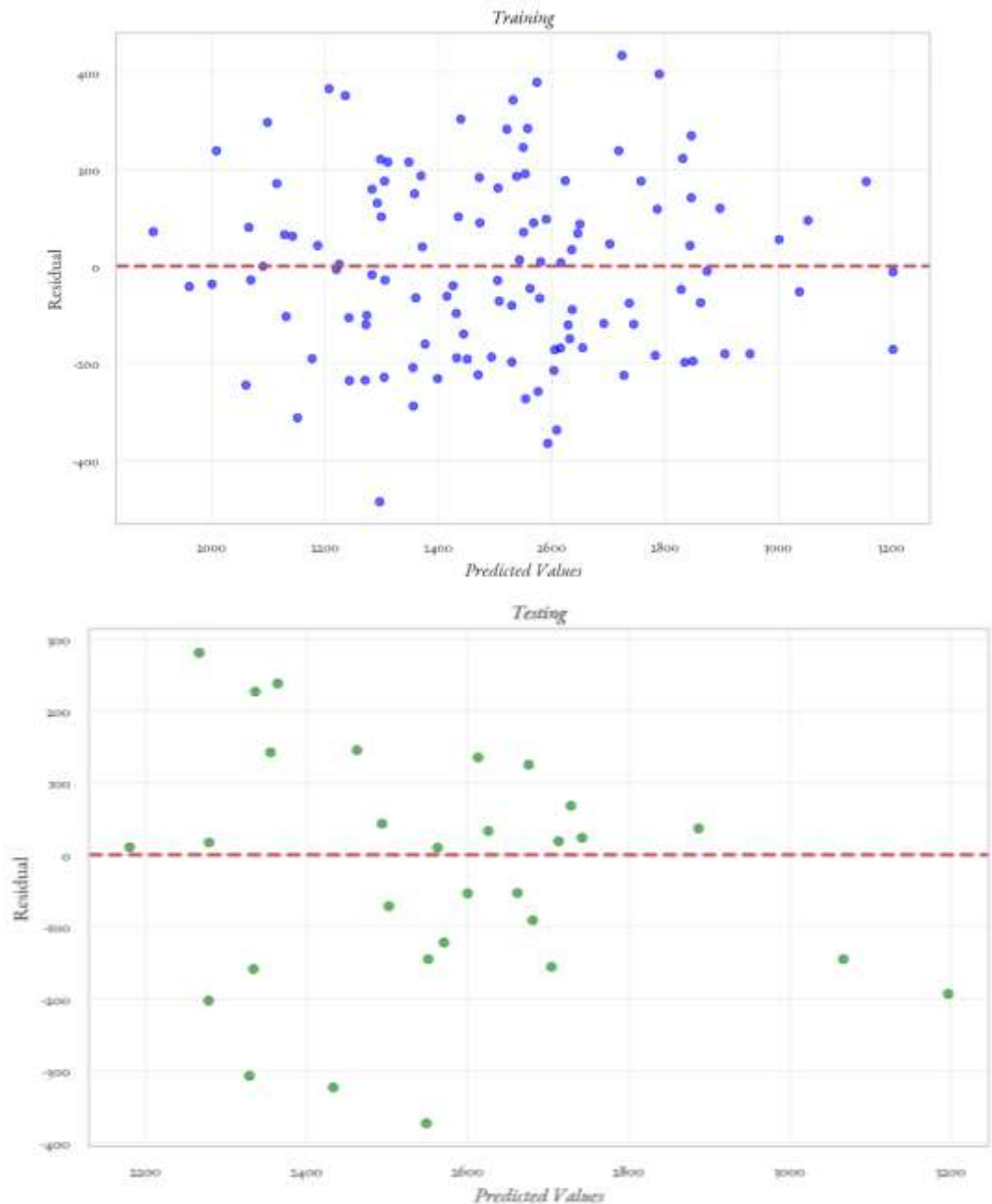
Metric	Training	Testing
R ² Score	0.68	0.58
Adjusted R ²	0.67	0.48
RMSE (gram)	185.94	164.72
MSE (gram ²)	34575.03	27132.69
MAE (gram)	154.48	131.64
MAPE (%)	-	5.44

Hasil dari Gambar 4 menunjukkan scatter plot antara nilai aktual dan nilai prediksi pada *data training* dan *data testing*, terlihat bahwa sebagian besar titik data mengikuti garis diagonal (*perfect prediction*), yang menunjukkan adanya hubungan linear yang cukup baik antara nilai aktual dan hasil prediksi model. Pada *data training*, sebaran titik relatif lebih rapat di sekitar garis prediksi sempurna, yang mengindikasikan bahwa model mampu menjelaskan variasi data dengan cukup baik pada *data training*. Sementara itu pada *data testing* sebaran titik terlihat menyebar dan beberapa titik berada cukup jauh dari garis diagonal, tercermin dari nilai R^2 yang menurun. Kondisi ini menunjukkan adanya penurunan kemampuan generalisasi model, ketika diaplikasikan pada data yang belum pernah dilihat sebelumnya. Meskipun demikian, hubungan linier masih dapat dideteksi, menunjukkan bahwa model regresi menunjukkan kemampuan prediksi yang dapat diterima. Namun, adanya kesalahan prediksi menandakan bahwa ada potensi untuk meningkatkan model, baik dengan memasukkan variabel yang relevan, meningkatkan kualitas data, atau menggunakan teknik pemodelan yang lebih canggih. Nilai R^2 berada dalam kisaran sedang, menunjukkan bahwa model tersebut dapat menjelaskan sebagian besar variasi dalam data, namun masih ada faktor tambahan di luar model yang dapat memengaruhi BBBL.



Gambar 4. Scatter Plot *Data Training* dan *Data Testing*

Plot residual untuk *data training* menunjukkan bahwa residual terdistribusi secara acak di sekitar garis nol tanpa pola yang jelas, yang menunjukkan bahwa asumsi linearitas dan homoskedastisitas model regresi relatif terpenuhi. Model tidak menunjukkan bias sistemik pada *data training* sebagaimana dibuktikan oleh distribusi residual yang sama di atas dan di bawah garis nol. Residual pada *data testing* cenderung lebih tersebar, dengan beberapa nilai ekstrim, menunjukkan bahwa ketika model diterapkan pada data baru, kesalahan prediksi meningkat. Meskipun demikian, tidak terlihat pola tertentu seperti tren naik atau turun yang jelas, sehingga model masih dapat dikatakan cukup stabil, walaupun kemampuan generalisasi tersebut lebih rendah dibandingkan pada *data training* yang ditunjukkan oleh Gambar 5.



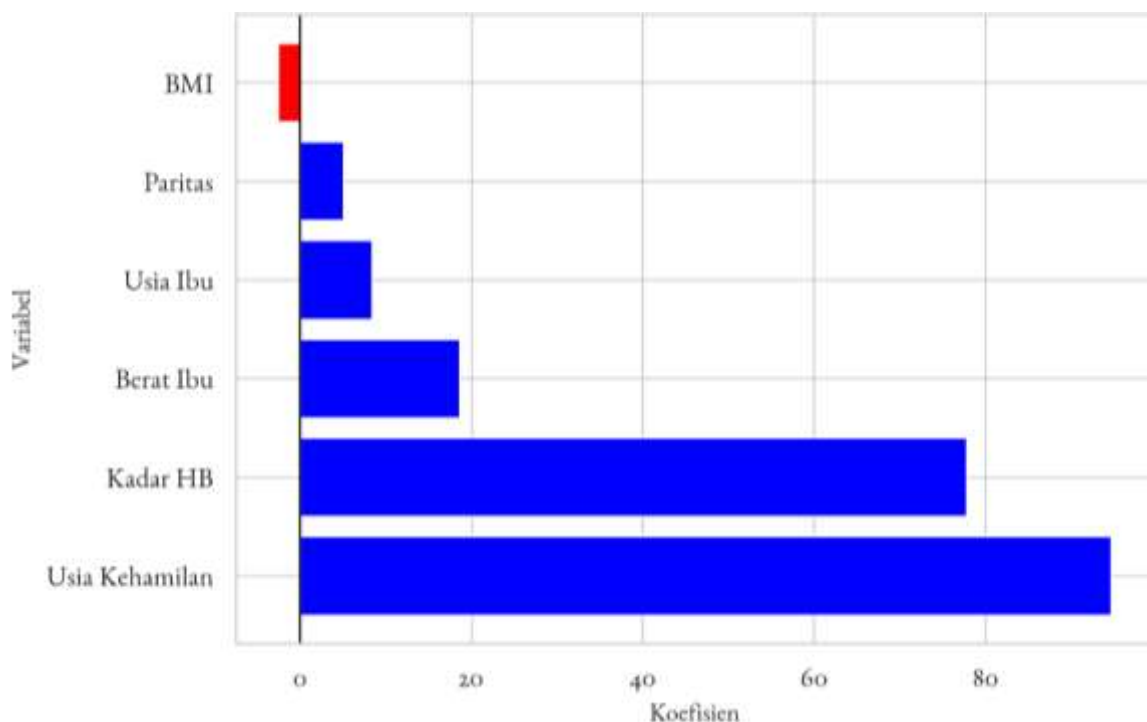
Gambar 5. Plot Residual *Data Training* dan *Data Testing*

Menurut temuan dari analisis regresi, faktor yang memiliki dampak terbesar pada BBBL adalah usia kehamilan, yang ditunjukkan oleh koefisien tertinggi. Ini berarti bahwa seiring bertambahnya usia kehamilan, BBBL juga meningkat. Kadar HB juga memainkan peran positif yang signifikan, dan berat ibu mengikuti dengan ketat, yang menyoroti peran penting kondisi nutrisi dan kesehatan ibu selama kehamilan berlangsung. Sebaliknya, usia ibu dan paritas memiliki efek relatif kecil namun tetap positif. Di sisi lain, BMI menunjukkan koefisien negatif dengan dampak paling kecil, menunjukkan korelasi negatif terhadap BBBL. Singkatnya, temuan ini mendukung gagasan bahwa unsur biologis dan kesehatan ibu selama kehamilan lebih berpengaruh daripada faktor lain dalam menentukan BBBL. Faktor-faktor yang mempengaruhi BBBL diurutkan berdasarkan nilai absolut koefisiennya, seperti yang disajikan pada Tabel 6.

Tabel 6. Koefisien Model Multiple Linear Regression

Variabel	Koefisien	Nilai Absolut Koefisien
Usia Kehamilan (minggu)	94.70	94.70
Kadar HB (g/dL)	77.80	77.80
Berat Ibu (kg)	18.61	18.61
Usia Ibu (tahun)	8.42	8.42
Paritas	5.07	5.07
BMI	-2.61	2.61

Koefisien positif berwarna biru menunjukkan hubungan searah, sedangkan koefisien negatif berwarna merah menunjukkan hubungan berlawanan arah dengan BBBL. Usia kehamilan menunjukkan pengaruh paling dominan terhadap BBBL. Kadar HB dan berat ibu memberikan pengaruh positif pada tingkat menengah. Usia ibu dan paritas menunjukkan pengaruh positif yang relatif kecil. Sebaliknya, BMI ibu menunjukkan hubungan negatif dengan pengaruh paling rendah dalam model. Visualisasi koefisien model ditampilkan pada Gambar 6.



Gambar 6. Bar Plot Koefisien Model

Singkatnya, temuan penelitian ini menunjukkan bahwa model Multiple Linear Regression yang dibuat dapat berfungsi sebagai alat bantu prediksi dini yang efektif untuk BBBL. Meskipun demikian, peningkatan presisi model masih dapat dioptimalkan dengan memasukkan variabel yang lebih lengkap atau menggunakan teknik pemodelan yang lebih canggih dalam penelitian selanjutnya.

KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi Berat Badan Bayi Lahir (BBBL) kontinu berbasis Multiple Linear Regression dengan performa prediktif moderat ($R^2 = 0.58$, Adjusted $R^2 = 0.48$ pada data testing, RMSE = 164.72 gram, dan MAPE = 5.44%). Usia kehamilan (koefisien = +94.70 g/minggu) dan kadar hemoglobin maternal (koefisien = +77.80 g/(g/dL)) terbukti sebagai prediktor utama yang paling

berpengaruh secara signifikan terhadap luaran berat lahir bayi di Puskesmas Kadudampit, Kabupaten Sukabumi.

Keterbatasan utama penelitian ini adalah ukuran sampel yang relatif terbatas (N=150) dan berasal dari satu fasilitas pelayanan primer (single-center). Untuk pengembangan selanjutnya, disarankan: (1) melakukan validasi eksternal multisentrik lintas puskesmas dan rumah sakit rujukan; (2) mengintegrasikan variabel prediktor klinis tambahan seperti lingkaran lengan atas (LILA), tinggi fundus uteri (TFU), dan status glukosa darah; serta (3) mengimplementasikan model ini ke dalam Sistem Informasi Manajemen Puskesmas (SIMPUS) atau aplikasi e-Kohort KIA sebagai sistem pendukung keputusan klinis (clinical decision support) guna memperkuat deteksi dini risiko BBLR pada layanan antenatal primer [19], [33], [34].

Pernyataan Mengenai Penggunaan Kecerdasan Buatan (Statement on AI)

Penulis menyatakan bahwa tidak ada perangkat lunak kecerdasan buatan generatif (generative artificial intelligence/GenAI) maupun model bahasa besar (large language models/LLMs) yang digunakan untuk menghasilkan ide konseptual, analisis data statistik, sintesis temuan, atau perumusan interpretasi ilmiah dalam naskah ini. Seluruh analisis statistik, perancangan model, dan penulisan artikel diselesaikan secara orisinal dan independen oleh para penulis, yang bertanggung jawab penuh atas keaslian, integritas, dan validitas seluruh isi publikasi.

Pernyataan Konflik Kepentingan (Conflicts of Interest)

Seluruh penulis menyatakan tidak memiliki konflik kepentingan (conflict of interest) baik yang bersifat finansial, hubungan personal, afiliasi komersial, maupun kepentingan non-finansial dengan pihak mana pun yang berpotensi memengaruhi objektivitas riset, metodologi, interpretasi hasil, atau kesimpulan yang disajikan dalam naskah ini.

Kontribusi Penulis (Author Contributions)

Asep Surahman Sulaeman: Konseptualisasi penelitian, formulasi metodologi pemodelan, komputasi dan pengujian Multiple Linear Regression, penulisan naskah draf awal (original draft preparation). Hilman Yunus: Supervisi penelitian, validasi metodologis dan analitis, kurasi hasil statistik, serta telaah kritis dan penyuntingan naskah (review & editing). Tasya Yustira Agustian: Kurasi dan pra-pemrosesan data rekam medis Puskesmas, verifikasi pengujian statistik deskriptif dan diagnostik asumsi klasik, serta penelaahan naskah akhir. Seluruh penulis telah membaca, meninjau secara seksama, dan menyetujui naskah akhir untuk diajukan.

Ucapan Terima Kasih (Acknowledgments)

Penulis menyampaikan apresiasi dan ucapan terima kasih yang sebesar-besarnya kepada Pimpinan dan seluruh staf tenaga medis Kesehatan Ibu dan Anak (KIA) di Puskesmas Kadudampit, Kabupaten Sukabumi, atas izin penelitian, kerja sama, dan keterbukaan akses terhadap data rekam medis antenatal. Penulis juga menyampaikan terima kasih kepada Program Studi Magister Informatika, Universitas Nusa Putra atas penyediaan lingkungan akademik, fasilitas komputasi, dan bimbingan yang memungkinkan terlaksananya penelitian ini hingga selesai.

DAFTAR PUSTAKA

- [1] R. F. Goldstein, S. K. Abell, S. Ranasinha, M. Misso, J. A. Boyle, M. H. Black, et al., "Association of Gestational Weight Gain With Maternal and Infant Outcomes: A Systematic Review and Meta-analysis," *JAMA*, vol. 317, no. 21, pp. 2207–2225, Jun. 2017, doi: 10.1001/jama.2017.3635.

- [2] World Health Organization, *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience*, Geneva: World Health Organization, 2018. [Online]. Available: <https://www.who.int/publications/i/item/9789241549912>.
- [3] F. Asefa, A. Cummins, Y. Dessie, F. Hayen, and M. Foureur, "Gestational Weight Gain and its Effect on Birth Outcomes: Systematic Review and Meta-Analysis," *PLOS ONE*, vol. 15, no. 4, p. e0231889, Apr. 2020, doi: 10.1371/journal.pone.0231889.
- [4] A. C. M. G. Figueiredo, I. S. Gomes-Filho, R. B. Silva, P. P. S. Pereira, F. A. S. Mata, M. C. Lyrio, et al., "Maternal Anemia and Low Birth Weight: A Systematic Review and Meta-Analysis," *Nutrients*, vol. 10, no. 5, p. 601, May 2018, doi: 10.3390/nu10050601.
- [5] M. S. Kramer, "Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis," *Bulletin of the World Health Organization*, vol. 65, no. 5, pp. 663–737, 1987.
- [6] E. A. Nohr, M. Vaeth, E. B. Baker, et al., "Combined Associations of Prepregnancy Body Mass Index and Gestational Weight Gain with the Risk of Small-for-Gestational-Age and Large-for-Gestational-Age Infants," *The American Journal of Clinical Nutrition*, vol. 87, no. 6, pp. 1750–1759, Jun. 2008, doi: 10.1093/ajcn/87.6.1750.
- [7] G. Shi, Z. Zhang, L. Ma, et al., "Association Between Maternal Iron Supplementation and Newborn Birth Weight: A Quantile Regression Analysis," *Italian Journal of Pediatrics*, vol. 47, no. 1, p. 133, Jun. 2021, doi: 10.1186/s13052-021-01084-7.
- [8] M. M. Rahman, S. K. Abe, M. S. Rahman, et al., "Maternal Anemia and Risk of Adverse Birth and Health Outcomes in Low- and Middle-Income Countries: Systematic Review and Meta-Analysis," *The American Journal of Clinical Nutrition*, vol. 103, no. 2, pp. 495–504, Feb. 2016, doi: 10.3945/ajcn.115.107896.
- [9] D. Lestari and A. Adisasmita, "Berat Badan Lahir Rendah (BBLR) sebagai Determinan Terjadinya ISPA pada Balita Analisis SDKI Tahun 2017," *Jurnal Epidemiologi Kesehatan Indonesia*, vol. 5, no. 1, pp. 27–34, Jun. 2021, doi: 10.7454/epidkes.v5i1.4083.
- [10] S. C. Jwa, T. Fujiwara, Y. Yamanobe, et al., "Changes in Maternal Hemoglobin During Pregnancy and Birth Outcomes," *BMC Pregnancy and Childbirth*, vol. 15, no. 1, p. 80, Mar. 2015, doi: 10.1186/s12884-015-0516-1.
- [11] World Health Organization, *Global Nutrition Targets 2025: Low Birth Weight Policy Brief*, Geneva: World Health Organization, 2014. [Online]. Available: <https://www.who.int/publications/i/item/WHO-NMH-NHD-14.5>.
- [12] Global Nutrition Target Collaborators, "Global, Regional, and National Progress Towards the 2030 Global Nutrition Targets and Forecasts to 2050: A Systematic Analysis for the Global Burden of Disease Study 2021," *The Lancet*, vol. 404, no. 10471, pp. 2543–2583, Jan. 2025, doi: 10.1016/S0140-6736(24)01821-X.
- [13] Kementerian Kesehatan Republik Indonesia, *Profil Kesehatan Indonesia 2023*, Jakarta: Kementerian Kesehatan RI, 2024.
- [14] Badan Pusat Statistik and United Nations Population Fund (UNFPA), *Mortalitas di Indonesia: Hasil Long Form Sensus Penduduk 2020*, Jakarta: BPS RI, 2023. [Online]. Available: <https://www.bps.go.id/publication/2024/02/02/mortalitas-di-indonesia.html>.
- [15] R. D. Wulandari, A. D. Laksono, and R. Matahari, "Policy to Decrease Low Birth Weight in Indonesia: Who Should Be the Target?," *Nutrients*, vol. 15, no. 2, p. 465, Jan. 2023, doi: 10.3390/nu15020465.

- [16] E. D. Utami and T. R. Anasthasia, "Faktor Kontekstual dan Individual terhadap Jumlah Anak Lahir Hidup: Sebuah Analisis Multilevel," *Jurnal Kependudukan Indonesia*, vol. 15, no. 1, pp. 1–16, Jun. 2020, doi: 10.14203/jki.v15i1.433.
- [17] S. V. Doubova, R. Pérez-Cuevas, E. Ortiz-Panozo, et al., "Evaluation of the Quality of Antenatal Care Using Electronic Health Record Information in Family Medicine Clinics of Mexico City," *BMC Pregnancy and Childbirth*, vol. 14, no. 1, p. 168, May 2014, doi: 10.1186/1471-2393-14-168.
- [18] J. S. Ancker, A. Edwards, S. Nosal, et al., "Effects of Workload, Work Complexity, and Repeated Alerts on Alert Fatigue in a Clinical Decision Support System," *BMC Medical Informatics and Decision Making*, vol. 17, no. 1, p. 36, Apr. 2017, doi: 10.1186/s12911-017-0430-8.
- [19] J. Desai, D. Saman, J. M. Sperl-Hillen, et al., "Implementing a Prediabetes Clinical Decision Support System in a Large Primary Care System: Design, Methods, and Pre-implementation Results," *Contemporary Clinical Trials*, vol. 114, p. 106686, Mar. 2022, doi: 10.1016/j.cct.2022.106686.
- [20] J. Song, X. Wang, B. Wang, et al., "Learning Implementation of a Guideline Based Decision Support System to Improve Hypertension Treatment in Primary Care: Pragmatic Cluster Randomised Controlled Trial," *BMJ*, vol. 386, p. e079143, Jul. 2024, doi: 10.1136/bmj-2023-079143.
- [21] I. Oktavianti, E. Ermatita, and D. Rini, "Analisis Pola Prediksi Data Time Series menggunakan Support Vector Regression, Multilayer Perceptron, dan Regresi Linear Sederhana," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 3, no. 2, pp. 282–287, Aug. 2019, doi: 10.29207/resti.v3i2.1013.
- [22] A. Ranjbar, F. Montazeri, M. V. Farashah, et al., "Machine Learning-Based Approach for Predicting Low Birth Weight," *BMC Pregnancy and Childbirth*, vol. 23, no. 1, p. 803, Nov. 2023, doi: 10.1186/s12884-023-06128-w.
- [23] W. T. Bekele, "Machine Learning Algorithms for Predicting Low Birth Weight in Ethiopia," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, p. 232, Sep. 2022, doi: 10.1186/s12911-022-01981-9.
- [24] T. B. Reza and N. Salma, "Prediction and Feature Selection of Low Birth Weight Using Machine Learning Algorithms," *Journal of Health, Population and Nutrition*, vol. 43, no. 1, p. 157, Oct. 2024, doi: 10.1186/s41043-024-00647-8.
- [25] E. W. Steyerberg and Y. Vergouwe, "Towards Better Clinical Prediction Models: Seven Steps for Development and an ABCD for Validation," *European Heart Journal*, vol. 35, no. 29, pp. 1925–1931, Aug. 2014, doi: 10.1093/eurheartj/ehu207.
- [26] D. G. Kleinbaum, L. L. Kupper, K. E. Muller, and A. Nizam, *Applied Regression Analysis and Other Multivariable Methods*, 5th ed. Boston: Cengage Learning, 2013.
- [27] D. Berrar, "Cross-Validation and Resampling Strategies for Predictive Modeling in Healthcare Informatics," in *Encyclopedia of Bioinformatics and Computational Biology*, vol. 1, Oxford: Elsevier, 2019, pp. 542–545, doi: 10.1016/B978-0-12-809633-8.20349-X.
- [28] M. Vigeh, E. Nishioka, K. Ohtani, et al., "Prenatal Mercury Exposure and Birth Weight," *Reproductive Toxicology*, vol. 76, pp. 78–83, Mar. 2018, doi: 10.1016/j.reprotox.2018.01.002.
- [29] C. Ogbutor, S. M. Mishaw, and Z. D. Mulla, "Mediation Analysis of Maternal Factors, Gestational Age, and Birth Weight on the Texas-Mexico Border," *Southern Medical Journal*, vol. 116, no. 6, pp. 478–481, Jun. 2023, doi: 10.14423/SMJ.0000000000001566.

- [30] S. Zhao, X. Yang, Q. Xu, et al., "Association of Maternal Exposures, Metabolites and Birth Outcomes in Newborns: A Prospective Cohort Study," *Environment International*, vol. 179, p. 108183, Sep. 2023, doi: 10.1016/j.envint.2023.108183.
- [31] G. S. Collins, J. B. Reitsma, D. G. Altman, and K. G. Moons, "Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement," *Annals of Internal Medicine*, vol. 162, no. 1, pp. 55–63, Jan. 2015, doi: 10.7326/M14-0697.
- [32] L. Zinsstag, S. Sonephet, J. Rigutto-Farebrother, et al., "Maternal and Infant Outcomes of Pregnancy Anemia at One Month Postpartum: A Cohort Analysis," *BMC Pregnancy and Childbirth*, vol. 25, no. 1, p. 1039, Dec. 2025, doi: 10.1186/s12884-025-08137-3.
- [33] E. W. Steyerberg, *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 2nd ed. Cham: Springer International Publishing, 2019, doi: 10.1007/978-3-030-16399-0.
- [34] F. E. Harrell, *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2nd ed. New York: Springer-Verlag, 2015, doi: 10.1007/978-3-319-19425-7.