BAB III
METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini termasuk dalam penelitian kuantitatif yang menerapkan pendekatan deskriptif. Penggunaan pendekatan kuantitatif didasarkan pada karakteristik data penelitian yang terdiri atas data numerik dan kategorikal dari hasil pengukuran serta pencatatan data balita. Variabel input yang digunakan dalam penelitian ini meliputi jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U. Data tersebut selanjutnya diolah menggunakan perhitungan matematis dan analisis statistik untuk menghasilkan klasifikasi status gizi yang dapat dievaluasi secara objektif.

Penelitian ini bersifat deskriptif karena bertujuan memberikan gambaran mengenai hasil klasifikasi status gizi balita berdasarkan data yang diperoleh dari Puskesmas Tanjung Gadang tanpa adanya perlakuan atau manipulasi terhadap data. Hasil klasifikasi status gizi balita yang diperoleh selanjutnya diimplementasikan ke dalam suatu aplikasi sederhana berbasis web yang dibangun dengan memanfaatkan bahasa pemrograman *Python* dan *framework* Streamlit sebagai tahap *deployment* dari model *Gaussian Naïve Bayes*. Melalui aplikasi tersebut, hasil klasifikasi dapat ditampilkan dan divisualisasikan sehingga lebih mudah dipahami dan digunakan dalam proses pemantauan status gizi balita.

Penelitian ini menerapkan metode *data mining* dengan algoritma *Gaussian Naïve Bayes* sebagai metode klasifikasi untuk menentukan status gizi balita berdasarkan karakteristik data numerik kontinu.

Tahapan analisis data dalam penelitian ini mengikuti kerangka kerja *CRISP-DM* (*Cross Industry Standard Process for Data Mining*), yang terdiri atas enam tahap, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*.

UIN IMAM BONJOL PADANG

B. Setting Penelitian

1. Lokasi Penelitian

Penelitian ini berlokasi di Puskesmas Tanjung Gadang, Kecamatan Tanjung Gadang, Kabupaten Sijunjung, Provinsi Sumatera Barat. Sebagai fasilitas pelayanan kesehatan tingkat pertama, Puskesmas Tanjung Gadang berada dalam naungan Dinas Kesehatan Kabupaten Sijunjung serta memiliki peran dalam memberikan pelayanan kesehatan kepada masyarakat.

Salah satu kegiatan yang dilakukan secara berkala di Puskesmas Tanjung Gadang adalah pemantauan pertumbuhan dan perkembangan balita. Kegiatan tersebut mencakup pengumpulan dan pencatatan data antropometri, yang menjadi dasar dalam menentukan status gizi balita. Karena data dikumpulkan secara rutin dan tersimpan secara terstruktur, data tersebut digunakan sebagai sumber dalam pelaksanaan penelitian ini.

Pemilihan Puskesmas Tanjung Gadang sebagai lokasi penelitian didasarkan pada ketersediaan data balita yang relevan dan terstruktur, serta kesesuaiannya dengan tujuan penelitian, yaitu melakukan klasifikasi status gizi balita menggunakan algoritma *Gaussian Naïve Bayes.* Selain itu, jumlah data balita yang cukup dan beragam menjadikan lokasi penelitian ini dinilai representatif untuk mendukung proses analisis dan pencapaian tujuan penelitian yang telah ditetapkan.

UIN IMAM BONJOL
PADANG

2. Jadwal Penelitian

Pelaksanaan penelitian berlangsung selama kurang lebih empat bulan. Seluruh kegiatan penelitian dijadwalkan secara sistematis untuk memastikan setiap tahapan dapat dilaksanakan secara terarah dan sesuai dengan rencana yang telah ditetapkan, mulai dari identifikasi dan perumusan masalah, pengumpulan data balita di Puskesmas Tanjung Gadang, pengolahan dan analisis data menggunakan algoritma *Gaussian Naïve Bayes*, hingga penyusunan laporan skripsi.

Rincian jadwal pelaksanaan penelitian disajikan pada Tabel 3.1 Jadwal Penelitian, yang memuat tahapan kegiatan penelitian beserta alokasi waktu pelaksanaannya.

Tabel 3. 1 Jadwal Penelitian

Kegiatan	Bulan 1				Bulan 2				Bulan 3				Bulan 4			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Identifikasi Masalah																
Studi Literatur																
Pengumpulan Data																
Data *Understanding*																
Data *Preparation* (*Preprocessing*)																
Penerapan Algoritma *Naïve Bayes*																
Evaluasi Model (*Confusion Matrix*)																
Analisis dan Interpretasi Hasil																
Penyusunan Laporan																

UIN IMAM BONJOL PADANG

C. Bahan atau Materi Penelitian

Bahan atau materi penelitian merupakan data dan informasi yang digunakan sebagai dasar dalam pelaksanaan penelitian guna mencapai tujuan yang telah ditetapkan. Pada penelitian ini, bahan penelitian diperoleh dari Puskesmas Tanjung Gadang melalui studi dokumentasi, berupa arsip dan laporan hasil pencatatan serta pengukuran status gizi balita yang telah terdokumentasi.

Adapun bahan atau materi yang digunakan dalam penelitian ini meliputi:

1. Data antropometri balita, yaitu data balita yang digunakan sebagai variabel *input* dalam proses klasifikasi status gizi, yang terdiri atas jenis kelamin (JK),

berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U. Data tersebut digunakan sebagai atribut *input* dalam proses klasifikasi menggunakan algoritma *Gaussian Naïve Bayes*.

2. Data status gizi balita, yaitu data BB/TB yang digunakan sebagai variabel target dalam proses klasifikasi. Data BB/TB merepresentasikan kategori status gizi balita yang terdiri atas Gizi Baik, Gizi Buruk, Gizi Kurang, Gizi Lebih, Obesitas, dan Risiko Gizi Lebih.
3. Data pendukung, yaitu data tambahan yang berkaitan dengan pelayanan kesehatan balita di Puskesmas Tanjung Gadang, yang digunakan untuk melengkapi informasi dalam penelitian.
4. Literatur pendukung, berupa buku dan jurnal ilmiah yang membahas sistem informasi, *data mining*, teknik klasifikasi, serta algoritma *Gaussian Naïve Bayes*, yang digunakan sebagai acuan dalam penyusunan penelitian.

D. Alat/Instrumen Penelitian

Alat penelitian membantu penelitian, terutama dalam proses pengolahan dan analisis data. Alat ini membantu setiap tahapan penelitian dilakukan secara sistematis dan sesuai dengan tujuan penelitian. Penelitian ini menggunakan perangkat keras (*hardware*) dan perangkat lunak (*software*) yang mendukung penggunaan teknik data mining.

UIN IMAM BONJOL PADANG

Adapun alat yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Perangkat Keras (Hardware)

a. Laptop Acer Aspire A514-54G
b. Processor minimal Intel Core i3 Generasi ke-11
c. RAM 8 GB
d. Penyimpanan HDD / SSD ≥ 256 GB
e. Mouse

2. Perangkat Lunak (Software)

a. Sistem Operasi Windows 11 Home Single Language (64-bit)
b. Microsoft Excel, untuk proses pra-pemrosesan data (preprocessing) dan rekapitulasi data

c. Google Colab, digunakan untuk menjalankan kode *Python* dalam proses klasifikasi status gizi balita menggunakan algoritma *Gaussian Naïve Bayes* dan evaluasi hasil klasifikasi.
d. Microsoft Word, untuk penyusunan laporan skripsi

E. Teknik Pengumpulan Data/Prosedur Penelitian

1. Teknik Pengumpulan Data

Teknik pengumpulan data merupakan cara yang digunakan oleh peneliti untuk memperoleh data dan informasi yang diperlukan sebagai dasar dalam proses analisis, sehingga tujuan penelitian dapat dicapai secara optimal. Dalam penelitian ini, metode pengumpulan data digunakan untuk mengumpulkan, mempelajari, dan memeriksa dokumen atau arsip di Puskesmas Tanjung Gadang tanpa melakukan pengumpulan data secara langsung di lapangan.

Dengan teknik tersebut, peneliti memanfaatkan data yang telah tersedia serta terdokumentasi secara sistematis. Kondisi tersebut memungkinkan proses memperoleh data dilakukan dengan lebih efektif dan efisien sesuai dengan kebutuhan penelitian. Data yang diperoleh dari Puskesmas Tanjung Gadang memuat berbagai atribut terkait data balita. Selanjutnya, dilakukan pemilihan atribut yang disesuaikan dengan tujuan penelitian. Atribut yang digunakan sebagai variable input dalam proses klasifikasi meliputi:

a. Jenis kelamin (JK)
b. Berat badan lahir (BB Lahir)
c. Tinggi badan (TB Lahir)
d. Usia saat ukur
e. Berat badan
f. Tinggi badan
g. ZS BB/U
h. BB/U
i. ZS TB/U
j. TB/U

Adapun BB/TB digunakan sebagai variabel target atau label klasifikasi, yaitu kategori yang akan diprediksi oleh algoritma *Gaussian Naïve Bayes*. Data yang

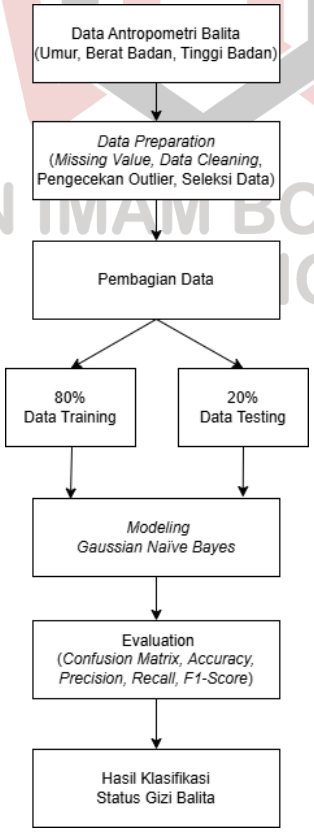
UIN IMAM BONJOL PADANG

telah melalui proses pemilihan atribut selanjutnya digunakan sebagai dasar dalam proses pengolahan dan analisis menggunakan algoritma *Gaussian Naïve Bayes* untuk mengklasifikasikan status gizi balita.

2. Prosedur Penelitian

Prosedur penelitian mencakup serangkaian tahapan yang dilaksanakan secara sistematis untuk mendukung proses penelitian. Dalam penelitian ini, pelaksanaan setiap tahapan mengacu pada kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*). Kerangka kerja tersebut dipilih karena menyediakan tahapan data mining yang terstruktur, dimulai dari pemahaman terhadap permasalahan hingga tahap evaluasi, sehingga pelaksanaan penelitian dapat berlangsung secara terarah dan sistematis.

Untuk memperjelas alur proses penelitian, digunakan skema proses pengolahan data yang meliputi tahap *data preparation*, pembagian data menjadi *data training* dan *data testing*, proses *modeling* menggunakan algoritma *Gaussian Naïve Bayes*, serta tahap *evaluation* menggunakan *confusion matrix*.



Gambar 3. 1 Skema Proses Klasifikasi Status Gizi Balita

UIN IMAM BONJOL

Adapun tahapan prosedur penelitian yang dilakukan adalah sebagai berikut:

a. Business Understanding

Tahap *business understanding* merupakan langkah awal dalam penelitian yang dilakukan untuk memperoleh pemahaman terhadap permasalahan yang menjadi fokus penelitian. Pada tahap ini, peneliti mengidentifikasi permasalahan yang berkaitan dengan pemantauan status gizi balita di Puskesmas Tanjung Gadang. Permasalahan yang dihadapi adalah data antropometri balita yang telah tersedia belum dimanfaatkan secara optimal untuk menghasilkan informasi klasifikasi status gizi yang dapat mendukung proses pemantauan. Oleh karena itu, penelitian ini bertujuan untuk melakukan klasifikasi status gizi balita dengan memanfaatkan data antropometri melalui penerapan algoritma *Gaussian Naïve Bayes.*

b. Data Understanding

Pada tahap *data understanding*, peneliti mengumpulkan dan menelaah data balita yang diperoleh dari Puskesmas Tanjung Gadang. Data tersebut dianalisis untuk mengetahui struktur, isi, dan karakteristiknya. Dataset yang diperoleh memuat beberapa atribut terkait data balita. Selanjutnya, dilakukan identifikasi terhadap atribut yang relevan dengan tujuan penelitian. Atribut yang digunakan sebagai variabel *input* dalam proses klasifikasi meliputi jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U, sedangkan BB/TB digunakan sebagai variabel target atau kelas yang akan diprediksi. Selain itu, dilakukan pengecekan awal terhadap kualitas data untuk mengetahui kelengkapan, konsistensi, dan kondisi data sebelum memasuki tahap pengolahan lebih lanjut.

c. Data Preparation

Tahap data preparation merupakan tahap persiapan data sebelum dilakukan proses pemodelan. Pada tahap ini dilakukan pemilihan atribut yang sesuai dengan tujuan penelitian, yaitu jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U, dengan BB/TB sebagai variabel target.

Selanjutnya, dilakukan pembentukan data dalam format observasi berdasarkan periode pengukuran agar setiap data balita pada setiap bulan dapat digunakan

UIN IMAM BONJOL PADANG

sebagai satu observasi. Setelah itu, dilakukan proses cleaning untuk menangani data yang tidak lengkap pada atribut yang digunakan serta memastikan data yang digunakan sesuai dengan kebutuhan analisis. Data kategorikal, yaitu JK, BB/U, dan TB/U, kemudian ditransformasikan ke dalam bentuk numerik. Atribut usia saat ukur juga dikonversi ke dalam satuan bulan agar dapat digunakan dalam proses pemodelan.

Setelah proses transformasi dilakukan, data dibagi menjadi data latih dan data uji dengan rasio 80:20 berdasarkan identitas balita. Data latih dan data uji kemudian dinormalisasi menggunakan metode *Min-Max Scaling*, dengan scaler dibentuk berdasarkan data latih dan digunakan pula untuk mentransformasikan data uji. Selanjutnya, data latih yang telah dinormalisasi diseimbangkan menggunakan metode SMOTE (*Synthetic Minority Over-sampling Technique*) untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas status gizi.

d. Modeling

Tahap modeling merupakan tahap penerapan metode klasifikasi yang digunakan dalam penelitian. Pada tahap ini, algoritma Gaussian Naïve Bayes diterapkan untuk mengklasifikasikan status gizi balita berdasarkan data yang telah dipersiapkan pada tahap data preparation. Atribut yang digunakan sebagai variabel input dalam proses klasifikasi terdiri atas jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U, sedangkan BB/TB digunakan sebagai variabel target atau kelas status gizi.

Sebelum proses pemodelan dilakukan, data dibagi menjadi data latih dan data uji menggunakan metode split data dengan rasio 80:20. Sebanyak 80% data digunakan sebagai data latih untuk membangun model klasifikasi, sedangkan 20% data digunakan sebagai data uji untuk mengukur performa model. Pada data latih, atribut kategorikal terlebih dahulu ditransformasikan ke dalam bentuk numerik, sedangkan atribut usia saat ukur dikonversi ke dalam satuan bulan. Selanjutnya, data dinormalisasi menggunakan Min-Max Scaling agar setiap atribut berada pada skala yang seragam.

UIN IMAM BONJOL PADANG

Data latih yang telah melalui proses normalisasi kemudian diseimbangkan menggunakan metode SMOTE (Synthetic Minority Over-sampling Technique) untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas status gizi. Data hasil SMOTE selanjutnya digunakan untuk melatih model Gaussian Naïve Bayes. Algoritma Gaussian Naïve Bayes menghitung probabilitas setiap data balita terhadap masing-masing kelas status gizi dengan menggunakan distribusi Gaussian berdasarkan nilai mean dan variance dari setiap atribut pada masing-masing kelas.

Setelah model terbentuk, data uji yang telah melalui tahapan preprocessing yang sama digunakan untuk melakukan prediksi. Proses klasifikasi dilakukan dengan menghitung probabilitas data balita terhadap masing-masing kelas status gizi berdasarkan sepuluh atribut input, kemudian kelas dengan probabilitas posterior tertinggi ditetapkan sebagai hasil klasifikasi status gizi balita.

e. Evaluation

Tujuan dari tahap evaluasi adalah untuk menentukan tingkat kinerja model klasifikasi yang dibuat. Proses evaluasi dilakukan dengan menggunakan *Confusion Matrix* melalui perbandingan antara hasil prediksi model dan kelas sebenarnya pada data uji. Selanjutnya, nilai *accuracy*, *precision*, *recall*, *dan F1-score* dihitung berdasarkan hasil tersebut untuk mengukur kemampuan model dalam mengklasifikasikan status gizi balita.

UIN IMAM BONJOL PADANG

f. Deployment

Tahap *deployment* merupakan tahap akhir dalam kerangka kerja *CRISP-DM* yang bertujuan untuk mengimplementasikan dan menyajikan hasil model yang telah diperoleh. Penelitian ini menggunakan *Streamlit* untuk mengintegrasikan model *Gaussian Naïve Bayes* yang telah dilatih dan dievaluasi ke dalam aplikasi berbasis web sederhana. Pengguna dapat menggunakan aplikasi tersebut dengan memasukkan sepuluh atribut input, yaitu jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U. Selanjutnya, data input diproses menggunakan tahapan *preprocessing* yang sama dengan proses pemodelan untuk memperoleh hasil prediksi status gizi balita. Hasil prediksi tersebut ditampilkan

melalui antarmuka *Streamlit* sebagai informasi pendukung dalam kegiatan pemantauan status gizi balita di Puskesmas Tanjung Gadang.

F. Teknik Analisis Data

Analisis data adalah fase yang dilakukan untuk mengolah data yang telah dikumpulkan untuk menghasilkan informasi yang relevan dengan tujuan penelitian. Untuk menentukan status gizi balita berdasarkan data antropometri, penelitian ini menggunakan analisis data *Gaussian Naïve Bayes* sebagai metode klasifikasi. Proses analisis berfokus pada perhitungan probabilitas dan penetapan kelas status gizi menggunakan algoritma ini.

Adapun tahapan analisis data menggunakan algoritma *Gaussian Naïve Bayes* adalah sebagai berikut:

1. Analisis Menggunakan *Gaussian Naïve Bayes*

Pada penelitian ini, proses klasifikasi status gizi balita dilakukan menggunakan algoritma *Gaussian Naïve Bayes*. Metode ini merupakan salah satu metode klasifikasi berbasis probabilitas yang dapat digunakan untuk mengklasifikasikan data yang telah ditransformasikan ke dalam bentuk numerik. *Gaussian Naïve Bayes* dipilih karena sesuai dengan karakteristik data yang digunakan dalam penelitian setelah melalui proses transformasi dan normalisasi. Atribut yang digunakan sebagai variabel *input* terdiri atas jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U. Data tersebut selanjutnya digunakan untuk menghitung probabilitas setiap kelas status gizi berdasarkan distribusi Gaussian, sehingga diperoleh prediksi status gizi balita.

Kelas status gizi yang digunakan dalam penelitian ini terdiri atas enam kelas, yaitu Gizi Baik, Gizi Buruk, Gizi Kurang, Gizi Lebih, Obesitas, dan Risiko Gizi Lebih. Proses klasifikasi dilakukan dengan menghitung probabilitas setiap kelas berdasarkan nilai atribut yang dimiliki oleh data balita, kemudian menentukan kelas dengan nilai probabilitas posterior terbesar sebagai hasil klasifikasi Adapun tahapan penerapan algoritma *Gaussian Naïve Bayes* dalam proses klasifikasi adalah sebagai berikut:

UIN IMAM BONJOL PADANG

a. Menentukan Peluang Prior

Menentukan peluang prior merupakan tahap awal dalam proses klasifikasi menggunakan algoritma *Gaussian Naïve Bayes*. Peluang prior digunakan untuk mengetahui kemungkinan awal setiap kelas status gizi sebelum mempertimbangkan nilai atribut data antropometri balita. Pada penelitian ini, peluang prior dihitung berdasarkan jumlah data pada masing-masing kelas status gizi yang terdapat dalam data latih.

Perhitungan peluang prior dilakukan dengan membandingkan jumlah data pada setiap kelas status gizi dengan jumlah keseluruhan data latih. Semakin besar jumlah data pada suatu kelas, maka semakin besar pula nilai peluang prior kelas tersebut. Perhitungan peluang prior dilakukan menggunakan Persamaan (2.1).

b. Menghitung Parameter Gaussian (Mean dan Standar Deviasi)

Pada tahap ini dilakukan perhitungan nilai rata-rata (*mean*) dan standar deviasi untuk setiap atribut *input* pada masing-masing kelas status gizi. Atribut yang digunakan terdiri atas jenis kelamin (JK), berat badan lahir (BB Lahir), tinggi badan lahir (TB Lahir), usia saat ukur, berat badan, tinggi badan, ZS BB/U, BB/U, ZS TB/U, dan TB/U yang telah melalui proses transformasi dan normalisasi. Nilai *mean* dan standar deviasi digunakan sebagai parameter dalam distribusi Gaussian untuk menghitung probabilitas atribut numerik pada setiap kelas status gizi. Perhitungan nilai *mean* dan standar deviasi dilakukan menggunakan persamaan (2.2) dan persamaan (2.3).

UIN IMAM BONJOL PADANG

c. Menghitung Peluang Kondisional (Likelihood)

Peluang kondisional digunakan untuk mengetahui kemungkinan suatu nilai atribut muncul pada kelas tertentu. Karena seluruh atribut *input* telah melalui proses transformasi ke dalam bentuk numerik, maka perhitungan peluang kondisional dilakukan menggunakan distribusi Gaussian berdasarkan parameter *mean* dan *variance* pada masing-masing kelas (2.4).

d. Menghitung Probabilitas Posterior

Setelah diperoleh nilai peluang prior dan peluang kondisional, langkah selanjutnya adalah menghitung probabilitas posterior untuk setiap kelas status gizi. Probabilitas posterior digunakan untuk menentukan kemungkinan suatu data

termasuk ke dalam kelas tertentu berdasarkan atribut yang dimilikinya. Perhitungan probabilitas posterior dilakukan menggunakan persamaan (2.6).

e. Menentukan Hasil Klasifikasi

Tahap terakhir dalam proses klasifikasi *Gaussian Naïve Bayes* adalah menentukan hasil klasifikasi. Data balita diklasifikasikan ke dalam kelas yang memiliki nilai probabilitas *posterior* terbesar. Dengan demikian, hasil klasifikasi ditentukan berdasarkan perbandingan nilai probabilitas *posterior* dari seluruh kelas status gizi.

2. Evaluasi Hasil Klasifikasi

Setelah model melakukan proses klasifikasi menggunakan algoritma *Gaussian Naïve Bayes*, tahap berikutnya adalah mengevaluasi hasil yang diperoleh. Evaluasi dilakukan untuk mengukur kemampuan model dalam menentukan status gizi balita secara tepat. Beberapa metrik yang digunakan dalam evaluasi meliputi akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score*.

a. Akurasi (Accuracy)

Akurasi merupakan metrik yang digunakan untuk mengetahui tingkat ketepatan prediksi model terhadap seluruh data yang diuji. Nilai tersebut diperoleh dengan membandingkan jumlah prediksi yang benar dengan keseluruhan data yang dievaluasi. Perhitungan akurasi mengacu pada Persamaan (2.7). Semakin besar nilai akurasi, semakin tinggi pula tingkat ketepatan model dalam melakukan klasifikasi status gizi balita.

b. Presisi (Precision)

Presisi (*precision*) merupakan metrik yang digunakan untuk mengukur ketepatan hasil prediksi model pada suatu kelas. Presisi menunjukkan seberapa banyak data yang diprediksi oleh model sebagai suatu kelas yang benar-benar termasuk dalam kelas tersebut. Nilai presisi dihitung berdasarkan jumlah True Positive (TP) dan False Positive (FP) menggunakan Persamaan (2.8). Semakin tinggi nilai presisi, semakin banyak data yang diprediksi sebagai suatu kelas yang sesuai dengan kelas sebenarnya.

UIN IMAM BONJOL PADANG

c. Recall (Sensitivity)

Recall digunakan untuk menilai kemampuan model dalam menemukan data yang benar-benar termasuk ke dalam suatu kelas tertentu. Nilai recall menunjukkan jumlah data dalam kelas yang berhasil dikenali dibandingkan dengan jumlah total data dalam kelas tersebut. Persamaan (2.9) digunakan untuk perhitungan recall, yang melibatkan True Positive (TP) dan False Negative (FN). Nilai recall yang lebih tinggi menunjukkan kemampuan model untuk mengidentifikasi data pada masing-masing kelas status gizi.

d. F1-Score

F1-score digunakan untuk mengevaluasi seberapa seimbang precision dan recall. Untuk memberikan gambaran yang lebih mendalam tentang kinerja model, metrik ini menggabungkan kedua nilai tersebut ke dalam satu ukuran evaluasi. Berdasarkan Persamaan (2.10), perhitungan nilai F1 dilakukan. Nilai F1 yang tinggi menunjukkan adanya keseimbangan yang baik antara ketepatan model dalam memberikan prediksi dan kemampuan untuk mengenali data untuk setiap kelas.

3. Confusion Matrix

Confusion matrix merupakan metode evaluasi yang digunakan untuk memperlihatkan hubungan antara kelas aktual dan kelas hasil prediksi dari suatu model. Penggunaan *confusion matrix* memungkinkan identifikasi jumlah data yang berhasil diklasifikasikan dengan tepat maupun data yang mengalami kesalahan prediksi pada setiap kelas status gizi.

Pada klasifikasi multikelas, komponen *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) ditentukan pada masing-masing kelas. Dalam proses tersebut, satu kelas diperlakukan sebagai kelas yang sedang dianalisis, sedangkan seluruh kelas lainnya dikelompokkan sebagai kelas selain kelas yang sedang dievaluasi.

Confusion matrix memiliki empat komponen utama, yaitu:

a. *True Positive* (TP), merupakan jumlah data yang pada kenyataannya berada dalam suatu kelas dan berhasil diprediksi secara tepat ke kelas tersebut.

UIN IMAM BONJOL PADANG

b. *True Negative* (TN), merupakan jumlah data yang sebenarnya tidak termasuk dalam suatu kelas dan berhasil diprediksi bukan sebagai bagian dari kelas tersebut.
c. *False Positive* (FP), merupakan jumlah data yang sebenarnya berasal dari kelas lain, tetapi oleh model diprediksi sebagai bagian dari kelas yang sedang dievaluasi.
d. *False Negative* (FN), merupakan jumlah data yang sebenarnya termasuk dalam suatu kelas, tetapi model memberikan prediksi bahwa data tersebut berada pada kelas lainnya.

UIN IMAM BONJOL
PADANG