

BAB III METODOLOGI PENELITIAN

3.1 Arsitektur Sistem

Penelitian ini bertujuan menganalisis sistem klasifikasi efektivitas Program Makan Bergizi Gratis menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes*. Dataset penelitian mencakup variabel kehadiran siswa, semangat belajar, persepsi siswa, dan kualitas pelaksanaan program dengan label target berupa tingkat efektivitas yang diklasifikasikan ke dalam kategori “Efektif” dan “Tidak Efektif”. Data diproses melalui tahap *preprocessing*, dibagi menjadi 80% data latih dan 20% data uji, kemudian dibangun model klasifikasi menggunakan kedua algoritma tersebut. Evaluasi model dilakukan menggunakan *confusion matrix* dan metrik evaluasi berupa *accuracy*, *precision*, *recall*, dan *F1-score* untuk membandingkan kinerja klasifikasi secara objektif dan terukur.

3.1.1. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di SMA Kemala Bhayangkari Rantauprapat, Kabupaten Labuhanbatu, Provinsi Sumatera Utara, dengan pertimbangan bahwa sekolah tersebut merupakan penerima Program Makan Bergizi Gratis sehingga relevan untuk mengkaji efektivitas pelaksanaan program dari sudut pandang siswa sebagai penerima manfaat langsung. Pelaksanaan penelitian berlangsung pada bulan Januari 2026 dengan tujuan memastikan seluruh proses penelitian berjalan secara sistematis dan komprehensif.

3.1.2. Populasi dan Sampel

Populasi penelitian mencakup siswa Kelas X, XI-2 dan XII-4 di SMA Kemala Bhayangkari Rantauprapat yang mengikuti Program Makan Bergizi Gratis sebagai bagian dari program kesejahteraan peserta didik. Seluruh populasi tersebut memiliki karakteristik yang relevan dengan tujuan penelitian karena terlibat langsung sebagai penerima manfaat program yang dianalisis. Penentuan sampel dilakukan menggunakan teknik *purposive sampling*, dengan kriteria siswa yang menjadi peserta Program Makan Bergizi Gratis pada kelas yang diteliti dan mengisi kuesioner secara lengkap. Sampel penelitian ditetapkan sebanyak 57 siswa yang memenuhi kriteria tersebut. Jumlah sampel ini dinilai mampu merepresentasikan kondisi efektivitas program karena memiliki keterkaitan langsung dengan variabel penelitian, sehingga hasil analisis prediktif dengan pendekatan klasifikasi yang diperoleh dapat menggambarkan tingkat efektivitas Program Makan Bergizi Gratis secara objektif dan terukur.

3.1.3. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui penyebaran kuesioner tertutup kepada siswa penerima Program Makan Bergizi Gratis di SMA Kemala Bhayangkari Rantauprapat. Kuesioner disusun secara terstruktur menggunakan skala Likert lima tingkat untuk mengukur persepsi dan pengalaman siswa terhadap pelaksanaan program. Instrumen ini digunakan untuk memperoleh data kuantitatif yang menjadi dasar analisis prediktif dengan pendekatan klasifikasi efektivitas program.

Penyebaran kuesioner dilakukan secara langsung kepada responden dengan pendampingan pihak sekolah guna memastikan pemahaman terhadap setiap pernyataan. Pertanyaan kuesioner mencakup aspek kehadiran siswa, semangat belajar, persepsi siswa, serta kualitas pelaksanaan Program Makan Bergizi Gratis. Data yang terkumpul selanjutnya diperiksa kelengkapannya dan dilakukan tahap *preprocessing* dengan mengeluarkan data yang tidak memenuhi kriteria penelitian sebelum dianalisis lebih lanjut.

Tabel 3.1. Data Kuesioner Penelitian

No	Nama Variabel	Deskripsi
1	KEHADIRAN SISWA	Saya lebih jarang absen sejak mengikuti Program Makan Bergizi Gratis.
		Program Makan Bergizi Gratis membantu saya lebih rutin hadir di sekolah.
		Saya merasa kondisi fisik saya mendukung untuk mengikuti pelajaran di sekolah.
2	SEMANGAT BELAJAR	Saya merasa lebih bersemangat mengikuti kegiatan belajar setelah menerima MBG.
		Konsentrasi saya saat belajar meningkat sejak adanya program MBG.
		Saya lebih aktif mengikuti pembelajaran di kelas.
3	PERSEPSI SISWA	Program Makan Bergizi Gratis bermanfaat bagi saya sebagai siswa.
		Menu makanan yang diberikan sesuai dengan kebutuhan saya.
		Saya merasa nyaman mengikuti program MBG di sekolah.
4	KUALITAS PELAKSANAAN MBG	Pelaksanaan Program Makan Bergizi Gratis berjalan dengan tertib.
		Makanan diberikan tepat waktu sesuai jadwal sekolah.
		Kebersihan dan penyajian makanan sudah baik.

Tabel 3.1 menyajikan variabel penelitian beserta pernyataan kuesioner yang digunakan sebagai instrumen pengumpulan data. Variabel Kehadiran Siswa terdiri dari tiga pernyataan yang menggambarkan perubahan kehadiran siswa, kebiasaan

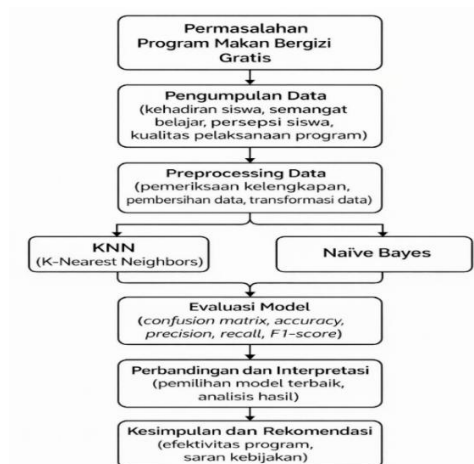
absen, serta kondisi fisik siswa dalam mengikuti pembelajaran setelah menerima Program Makan Bergizi Gratis.

Variabel Semangat Belajar diukur melalui tiga pernyataan yang mencerminkan tingkat motivasi, konsentrasi, dan keaktifan siswa dalam kegiatan pembelajaran. Selanjutnya, variabel Persepsi Siswa mencakup pernyataan yang menilai manfaat program, kesesuaian menu makanan, serta kenyamanan siswa dalam mengikuti Program Makan Bergizi Gratis. Variabel Kualitas Pelaksanaan Program Makan Bergizi Gratis diukur melalui pernyataan yang berkaitan dengan ketertiban pelaksanaan, ketepatan waktu pemberian makanan, serta kebersihan dan penyajian makanan.

3.2 Desain Aktifitas Sistem

Desain aktivitas sistem dalam penelitian ini menggambarkan alur kerja analisis prediktif dengan pendekatan klasifikasi efektivitas Program Makan Bergizi Gratis di SMA Kemala Bhayangkari Rantauprapat menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes*. Alur aktivitas dimulai dari pengumpulan data kuesioner siswa sebagai input sistem yang mencerminkan variabel kehadiran siswa, semangat belajar, persepsi siswa, dan kualitas pelaksanaan program. Data yang diperoleh selanjutnya diproses melalui tahap *preprocessing* yang meliputi pemeriksaan kelengkapan data, penghapusan data yang tidak valid, serta penyeragaman format. Setelah itu, dilakukan transformasi data untuk menyesuaikan bentuk data agar siap digunakan dalam proses klasifikasi.

Dataset kemudian dibagi menjadi data latih dan data uji sesuai dengan alur pada diagram aktivitas sistem. Data latih digunakan untuk membangun model klasifikasi menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes*, sedangkan data uji digunakan untuk mengevaluasi kinerja model. Evaluasi dilakukan menggunakan *confusion matrix* dengan perhitungan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil evaluasi tersebut menjadi keluaran sistem yang digunakan untuk membandingkan kinerja kedua algoritma serta menarik kesimpulan mengenai efektivitas pelaksanaan Program Makan Bergizi Gratis.



Gambar 3.1 Alur Penelitian Analisis Prediktif Program Makan Bergizi Gratis

Gambar 3.1 menunjukkan alur penelitian yang digunakan dalam mengklasifikasikan tingkat efektivitas Program Makan Bergizi Gratis. Proses penelitian diawali dengan identifikasi permasalahan, dilanjutkan dengan pengumpulan dan *preprocessing* data, kemudian pemodelan menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes*. Tahap evaluasi model dilakukan untuk menilai kinerja klasifikasi, yang selanjutnya digunakan sebagai dasar perbandingan, interpretasi hasil, serta penarikan kesimpulan dan rekomendasi.

3.3 Tahapan Pengolahan Data

Tahapan pengolahan data penelitian dilakukan untuk mendukung proses analisis prediktif dengan pendekatan klasifikasi dalam mengevaluasi efektivitas pelaksanaan Program Makan Bergizi Gratis di SMA Kemala Bhayangkari Rantauprapat. Data yang digunakan berasal dari hasil kuesioner siswa yang telah dikumpulkan dan direkapitulasi ke dalam bentuk dataset terstruktur. Proses pengolahan data dilaksanakan secara bertahap dan sistematis agar data yang dianalisis memiliki kualitas yang baik serta siap digunakan dalam pemodelan menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes*. Tahapan pengolahan data mencakup pengumpulan data, seleksi data, transformasi data, pembagian data, pemodelan, dan evaluasi model. Setiap tahapan dilaksanakan secara berurutan untuk memastikan hasil klasifikasi yang diperoleh akurat, konsisten, dan dapat dipertanggung jawabkan secara ilmiah.

3.3.1. Pengumpulan Data

Pengumpulan data dilakukan melalui penyebaran kuesioner kepada siswa SMA Kemala Bhayangkari Rantauprapat yang mengikuti Program Makan Bergizi Gratis. Kuesioner disusun untuk memperoleh data mengenai kondisi siswa setelah mengikuti program, dengan fokus pada variabel kehadiran siswa, semangat belajar, persepsi siswa, serta kualitas pelaksanaan Program Makan Bergizi Gratis. Data yang diperoleh dari kuesioner kemudian direkapitulasi ke dalam bentuk dataset. Jumlah data responden yang digunakan dalam penelitian ini sebanyak 57 data yang dinyatakan valid dan layak digunakan untuk proses analisis lebih lanjut.

Tabel 3.2 Data Hasil Kuesioner

No	Nama	Kelas	Jenis Kelamin	Kehadiran Siswa			Semangat Belajar			Persepsi Siswa			Kualitas Pelaksanaan MBG			Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Total	Kategori
				P1X1	P1X2	P1X3	P2X1	P2X2	P2X3	P3X1	P3X2	P3X3	P4X1	P4X2	P4X3						
3	Khairunnida	X	Perempuan	4	4	4	4	3	4	4	5	5	4	3	5	12	11	14	12	49	Efektif
4	Nadia Melani Tambunan	X	Perempuan	5	5	5	4	3	2	3	3	5	4	3	4	15	9	11	11	46	Efektif
5	Rahma Pane	X	Perempuan	5	4	3	5	3	3	3	3	4	5	3	5	12	11	10	13	46	Efektif
6	Richnes A Nadeak	X	Perempuan	5	3	4	5	3	3	5	4	5	5	3	4	12	11	14	12	49	Efektif
7	Raisa Andin Marpaung	X	Perempuan	5	4	3	3	3	3	4	3	4	5	4	5	12	9	11	14	46	Efektif
8	Zaskia Kimanda furi	X	Perempuan	4	4	5	4	3	3	5	4	4	5	5	5	13	10	13	15	51	Efektif
9	Gresia Indah Maretah	X	Perempuan	4	4	4	5	4	2	4	4	4	4	4	4	12	11	12	12	47	Efektif
10	Dinda Nur Syadah	X	Perempuan	4	4	3	5	3	3	4	3	5	5	3	4	11	11	12	12	46	Efektif
11	Hanny Octavia Nasution	X	Perempuan	4	5	5	5	4	3	3	4	5	5	3	5	14	12	12	13	51	Efektif
12	Suci Ramadhani	X	Perempuan	2	4	3	5	3	3	4	3	5	5	5	5	9	11	12	15	47	Efektif
13	Padilah	X	Perempuan	5	5	5	5	3	3	3	4	5	4	3	5	15	11	12	12	50	Efektif
14	Hasanah Imanensia	X	Perempuan	5	4	3	5	3	3	5	4	5	5	4	5	12	11	14	14	51	Efektif
15	Cesya Bintang	X	Perempuan	4	5	5	5	3	3	5	4	5	5	3	5	14	11	14	13	52	Efektif
16	Alsi Ananda	X	Perempuan	5	4	4	5	3	3	4	4	4	5	3	5	13	11	12	13	49	Efektif
17	Aamad Reffy	X	Laki-Laki	4	5	4	5	3	3	4	4	4	3	2	3	13	11	12	8	44	Tidak Efektif
18	Rihai May Yanti Pasaribu	X	Perempuan	2	5	3	5	3	4	4	4	5	4	3	5	10	12	13	12	47	Efektif
19	Anyaasya Zului	X	Laki-Laki	5	4	3	5	3	1	3	4	5	4	3	5	12	9	12	12	45	Tidak Efektif
20	Alvi Randi Ramadhani	X	Laki-Laki	3	3	4	4	3	3	4	4	5	4	3	5	10	10	13	12	45	Tidak Efektif

Tabel 3.2 menampilkan 20 data sampel sebagai cuplikan dari total 57 data responden. Cuplikan data tersebut disajikan untuk memberikan gambaran struktur dan karakteristik dataset penelitian. Penyajian dalam bentuk cuplikan dilakukan untuk menjaga keterbacaan dan efisiensi penyajian pada dokumen, sedangkan seluruh proses pengolahan data, pemodelan, dan evaluasi model tetap menggunakan keseluruhan dataset sebanyak 57 data responden.

3.3.2. Seleksi Data

Tahap seleksi data bertujuan menyaring dan mempertahankan atribut yang relevan dengan tujuan penelitian. Atribut yang digunakan sebagai variabel input mencakup kehadiran siswa, semangat belajar, persepsi siswa, dan kualitas pelaksanaan Program Makan Bergizi Gratis, sedangkan kategori efektivitas program digunakan sebagai label kelas dalam proses klasifikasi. Data identitas responden seperti nama, usia, jenis kelamin, dan kelas tidak disertakan dalam pemodelan karena tidak memiliki keterkaitan langsung dengan proses klasifikasi. Baris data yang hanya memuat kode indikator kuesioner juga

dikeluarkan dari analisis karena tidak merepresentasikan data responden yang dapat diolah.

Tabel 3.3 Hasil Seleksi Data

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R1	12	11	14	12	Efektif
R2	15	9	11	11	Efektif
R3	12	11	10	13	Efektif
R4	12	11	14	12	Efektif
R5	12	9	11	14	Efektif
R6	13	10	13	15	Efektif
R7	12	11	12	12	Efektif
R8	11	11	12	12	Efektif
R9	14	12	12	13	Efektif
R10	9	11	12	15	Efektif
R11	15	11	12	12	Efektif
R12	12	11	14	14	Efektif
R13	14	11	14	13	Efektif
R14	13	11	12	13	Efektif
R15	13	11	12	8	Tidak Efektif
R16	10	12	13	12	Efektif
R17	12	9	12	12	Tidak Efektif
R18	10	10	13	12	Tidak Efektif
R19	10	9	12	10	Tidak Efektif
R20	7	10	7	7	Tidak Efektif
R21	12	9	13	14	Efektif
R22	14	10	14	13	Efektif
R23	12	9	10	12	Tidak Efektif
R24	12	11	15	15	Efektif
R25	15	15	15	15	Efektif
R26	15	13	15	15	Efektif
R27	10	10	12	15	Efektif
R28	9	10	11	13	Tidak Efektif
R29	11	9	14	13	Efektif
R30	11	9	12	12	Tidak Efektif
R31	10	10	12	13	Tidak Efektif
R32	9	12	14	12	Efektif
R33	9	10	12	13	Tidak Efektif
R34	10	10	12	13	Tidak Efektif
R35	10	11	11	12	Tidak Efektif
R36	9	9	12	13	Tidak Efektif
R37	9	9	11	11	Tidak Efektif
R38	9	11	13	12	Tidak Efektif
R39	11	9	13	13	Efektif

R40	9	11	11	12	Tidak Efektif
R41	10	10	11	12	Tidak Efektif
R42	9	11	12	12	Tidak Efektif
R43	13	11	12	13	Efektif
R44	10	12	12	12	Efektif
R45	12	9	14	14	Efektif
R46	8	10	11	14	Tidak Efektif
R47	8	7	11	8	Tidak Efektif
R48	10	10	10	10	Tidak Efektif
R49	9	10	12	13	Tidak Efektif
R50	10	11	11	9	Tidak Efektif
R51	10	11	11	9	Tidak Efektif
R52	11	10	11	11	Tidak Efektif
R53	10	10	13	15	Efektif
R54	11	13	13	12	Efektif
R55	9	9	14	14	Efektif
R56	11	13	10	10	Tidak Efektif
R57	11	10	11	13	Tidak Efektif

Tabel 3.3 menunjukkan data hasil seleksi yang hanya memuat atribut-atribut yang relevan dengan proses klasifikasi. Proses seleksi ini dilakukan untuk meningkatkan kualitas data serta mengurangi noise yang dapat memengaruhi kinerja model klasifikasi.

3.3.3. Transformasi Data

Transformasi data dilakukan untuk menyesuaikan format data agar sesuai dengan kebutuhan proses klasifikasi. Data hasil kuesioner yang semula berupa nilai numerik merupakan hasil penjumlahan skor dari beberapa pernyataan pada setiap variabel. Nilai numerik tersebut dikonversi ke dalam bentuk kategori untuk memudahkan proses pengelompokan data dan perhitungan manual sebelum tahap pemodelan.

Tabel 3.4 Aturan Transformasi Data

Variabel	Rentang Nilai	Kategori
Kehadiran Siswa	12 - 15	Tinggi
	9 - 11	Sedang

	7 - 8	Rendah
Semangat Belajar	12 - 15	Tinggi
	9 - 11	Sedang
	7 - 8	Rendah
Persepsi Siswa	12 - 15	Baik
	9 - 11	Cukup
	7 - 8	Kurang
Kualitas Pelaksanaan MBG	12 - 15	Baik
	9 - 11	Cukup
	7 - 8	Kurang

Tabel 3.4 menyajikan aturan transformasi data yang digunakan untuk mengonversi nilai numerik hasil kuesioner ke dalam kategori tertentu. Variabel Kehadiran Siswa dan Semangat Belajar diklasifikasikan ke dalam kategori Tinggi, Sedang, dan Rendah, sedangkan variabel Persepsi Siswa, dan Kualitas Pelaksanaan Program Makan Bergizi Gratis diklasifikasikan ke dalam kategori Baik, Cukup, dan Kurang berdasarkan rentang nilai yang telah ditentukan. Aturan transformasi ini diterapkan secara konsisten pada seluruh dataset agar data memiliki format yang seragam sebelum dilakukan proses pemodelan.

Tabel 3.5 Data Hasil Transformasi

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R1	Tinggi	Sedang	Baik	Baik	Efektif
R2	Tinggi	Sedang	Cukup	Cukup	Efektif
R3	Tinggi	Sedang	Cukup	Baik	Efektif
R4	Tinggi	Sedang	Baik	Baik	Efektif
R5	Tinggi	Sedang	Cukup	Baik	Efektif
R6	Tinggi	Sedang	Baik	Baik	Efektif
R7	Tinggi	Sedang	Baik	Baik	Efektif
R8	Sedang	Sedang	Baik	Baik	Efektif
R9	Tinggi	Tinggi	Baik	Baik	Efektif
R10	Sedang	Sedang	Baik	Baik	Efektif
R11	Tinggi	Sedang	Baik	Baik	Efektif
R12	Tinggi	Sedang	Baik	Baik	Efektif
R13	Tinggi	Sedang	Baik	Baik	Efektif
R14	Tinggi	Sedang	Baik	Baik	Efektif

R15	Tinggi	Sedang	Baik	Kurang	Tidak Efektif
R16	Sedang	Tinggi	Baik	Baik	Efektif
R17	Tinggi	Sedang	Baik	Baik	Tidak Efektif
R18	Sedang	Sedang	Baik	Baik	Tidak Efektif
R19	Sedang	Sedang	Baik	Cukup	Tidak Efektif
R20	Rendah	Sedang	Kurang	Kurang	Tidak Efektif
R21	Tinggi	Sedang	Baik	Baik	Efektif
R22	Tinggi	Sedang	Baik	Baik	Efektif
R23	Tinggi	Sedang	Cukup	Baik	Tidak Efektif
R24	Tinggi	Sedang	Baik	Baik	Efektif
R25	Tinggi	Tinggi	Baik	Baik	Efektif
R26	Tinggi	Tinggi	Baik	Baik	Efektif
R27	Sedang	Sedang	Baik	Baik	Efektif
R28	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R29	Sedang	Sedang	Baik	Baik	Efektif
R30	Sedang	Sedang	Baik	Baik	Tidak Efektif
R31	Sedang	Sedang	Baik	Baik	Tidak Efektif
R32	Sedang	Tinggi	Baik	Baik	Efektif
R33	Sedang	Sedang	Baik	Baik	Tidak Efektif
R34	Sedang	Sedang	Baik	Baik	Tidak Efektif
R35	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R36	Sedang	Sedang	Baik	Baik	Tidak Efektif
R37	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R38	Sedang	Sedang	Baik	Baik	Tidak Efektif
R39	Sedang	Sedang	Baik	Baik	Efektif
R40	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R41	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R42	Sedang	Sedang	Baik	Baik	Tidak Efektif
R43	Tinggi	Sedang	Baik	Baik	Efektif
R44	Sedang	Tinggi	Baik	Baik	Efektif
R45	Tinggi	Sedang	Baik	Baik	Efektif
R46	Rendah	Sedang	Cukup	Baik	Tidak Efektif
R47	Rendah	Rendah	Cukup	Kurang	Tidak Efektif
R48	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R49	Sedang	Sedang	Baik	Baik	Tidak Efektif
R50	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R51	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R52	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R53	Sedang	Sedang	Baik	Baik	Efektif
R54	Sedang	Tinggi	Baik	Baik	Efektif
R55	Sedang	Sedang	Baik	Baik	Efektif
R56	Sedang	Tinggi	Cukup	Cukup	Tidak Efektif
R57	Sedang	Sedang	Cukup	Baik	Tidak Efektif

Tabel 3.5 menampilkan seluruh data hasil transformasi dari nilai numerik ke dalam bentuk kategorikal berdasarkan aturan pada Tabel 3.4. Data yang disajikan

mencakup 57 responden dan menjadi dataset akhir yang digunakan dalam proses pembagian data, pemodelan, serta evaluasi kinerja algoritma klasifikasi. Data hasil transformasi digunakan secara langsung pada proses klasifikasi menggunakan algoritma *Naive Bayes* karena algoritma tersebut mampu mengolah data kategorikal berbasis probabilitas, sedangkan pada metode *K-Nearest Neighbors* (KNN) dilakukan proses encoding numerik pada tahap pemodelan untuk memungkinkan perhitungan jarak antar data dalam proses klasifikasi.

3.3.4. Pembagian Data

Pembagian data dilakukan untuk memisahkan dataset menjadi data latih (*training data*) dan data uji (*testing data*) sebagai dasar dalam proses pemodelan dan evaluasi algoritma klasifikasi. Dataset yang digunakan merupakan data hasil transformasi yang telah dikategorikan pada tahap sebelumnya.

Pada penelitian ini, pembagian data dilakukan dengan rasio 80% data latih dan 20% data uji. Dari total 57 data responden, sebanyak 46 data digunakan sebagai data latih, sedangkan 11 data digunakan sebagai data uji. Pembagian ini bertujuan untuk memastikan model memiliki data yang cukup dalam proses pembelajaran serta menyediakan data uji yang representatif untuk mengukur kinerja model klasifikasi secara objektif.

Data latih digunakan dalam proses pemodelan menggunakan algoritma *K-Nearest Neighbors* (KNN) dan *Naive Bayes* untuk membangun model klasifikasi tingkat efektivitas Program Makan Bergizi Gratis ke dalam kategori “Efektif” dan “Tidak Efektif”. Data uji digunakan untuk mengevaluasi hasil klasifikasi yang dihasilkan oleh masing-masing algoritma.

Tabel 3.6 Pembagian Data

Jenis Data	Jumlah Data	Persentase
Data Latih	46	80%
Data Uji	11	20%
Total	57	100%

Pembagian data dilakukan secara acak (*random*) untuk menghindari bias dalam proses pemodelan dan memastikan bahwa data latih dan data uji memiliki karakteristik yang seimbang.

3.3.5. Pemodelan Menggunakan K-Nearest Neighbors (KNN)

Pemodelan *K-Nearest Neighbors (KNN)* dilakukan menggunakan data yang telah melalui tahap seleksi, transformasi, dan pembagian data. Tahapan yang dilakukan adalah penentuan nilai K , *encoding* data, perhitungan jarak *Euclidean*, pengurutan jarak, pemilihan tetangga terdekat, dan penentuan kelas berdasarkan mayoritas tetangga.

1. Penentuan Nilai K

Nilai K pada penelitian ini ditetapkan sebesar $k = 3$. Nilai ini digunakan secara konsisten pada seluruh proses klasifikasi untuk menghindari hasil yang bersifat imbang serta memastikan keputusan kelas ditentukan oleh mayoritas tetangga terdekat.

2. Encoding Data

Untuk keperluan perhitungan jarak, kategori pada setiap variabel dikonversi menjadi nilai numerik sebagai berikut:

Kehadiran Siswa : Rendah = 1, Sedang = 2, Tinggi = 3

Semangat Belajar : Rendah = 1, Sedang = 2, Tinggi = 3

Persepsi Siswa : Kurang = 1, Cukup = 2, Baik = 3

Kualitas Pelaksanaan MBG : Kurang = 1, Cukup = 2, Baik = 3

Hasil *encoding* menghasilkan dataset numerik seperti yang disajikan pada

Tabel 3.7 Data Latih KNN dan Tabel 3.8 Data Uji KNN.

Tabel 3.7 Data Latih KNN

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R1	3	2	3	3	Efektif
R17	3	2	3	3	Tidak Efektif
R3	3	2	2	3	Efektif
R23	3	2	2	3	Tidak Efektif
R38	2	2	3	3	Tidak Efektif
R6	3	2	3	3	Efektif
R29	2	2	3	3	Efektif
R51	2	2	2	2	Tidak Efektif
R9	3	3	3	3	Efektif
R57	2	2	2	3	Tidak Efektif
R11	3	2	3	3	Efektif
R12	3	2	3	3	Efektif
R4	3	2	3	3	Efektif
R14	3	2	3	3	Efektif
R55	2	2	3	3	Efektif
R16	2	3	3	3	Efektif
R56	2	3	2	2	Tidak Efektif
R18	2	2	3	3	Tidak Efektif
R54	2	3	3	3	Efektif
R47	1	1	2	1	Tidak Efektif
R21	3	2	3	3	Efektif
R22	3	2	3	3	Efektif
R15	3	2	3	1	Tidak Efektif
R24	3	2	3	3	Efektif
R25	3	3	3	3	Efektif
R7	3	2	3	3	Efektif
R27	2	2	3	3	Efektif
R28	2	2	2	3	Tidak Efektif
R53	2	2	3	3	Efektif
R30	2	2	3	3	Tidak Efektif
R31	2	2	3	3	Tidak Efektif
R10	2	2	3	3	Efektif
R33	2	2	3	3	Tidak Efektif
R34	2	2	3	3	Tidak Efektif
R52	2	2	2	2	Tidak Efektif

R36	2	2	3	3	Tidak Efektif
R37	2	2	2	2	Tidak Efektif
R48	2	2	2	2	Tidak Efektif
R39	2	2	3	3	Efektif
R40	2	2	2	3	Tidak Efektif
R41	2	2	2	3	Tidak Efektif
R42	2	2	3	3	Tidak Efektif
R43	3	2	3	3	Efektif
R50	2	2	2	2	Tidak Efektif
R45	3	2	3	3	Efektif
R46	1	2	2	3	Tidak Efektif

Tabel 3.8 Data Uji KNN

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R8	2	2	3	3	Efektif
R13	3	2	3	3	Efektif
R5	3	2	2	3	Efektif
R20	1	2	1	1	Tidak Efektif
R2	3	2	2	2	Efektif
R19	2	2	3	2	Tidak Efektif
R26	3	3	3	3	Efektif
R44	2	3	3	3	Efektif
R35	2	2	2	3	Tidak Efektif
R49	2	2	3	3	Tidak Efektif
R32	2	3	3	3	Efektif

3. Perhitungan Jarak Euclidean

Perhitungan jarak dilakukan menggunakan jarak *Euclidean* dengan empat variabel, yaitu Kehadiran Siswa (x_1), Semangat Belajar (x_2), Persepsi Siswa (x_3), dan Kualitas Pelaksanaan MBG (x_4). Rumus jarak *Euclidean* yang digunakan adalah:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

Sebagai contoh perhitungan manual digunakan pada satu data uji Responden 8 dengan nilai (2,2,3,3) seperti pada Tabel 3.9.

Tabel 3.9. Data Baru (KNN)

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG
R8	2	2	3	3

Contoh perhitungan jarak dari data uji R8 (2,2,3,3) terhadap salah satu data latih R1 (3,2,3,3) adalah:

$$d = \sqrt{(3 - 2)^2 + (2 - 2)^2 + (3 - 3)^2 + (3 - 3)^2}$$

$$d = \sqrt{1 + 0 + 0 + 0} = \sqrt{1} = 1$$

Perhitungan jarak dilakukan dengan cara yang sama terhadap seluruh data latih hingga diperoleh hasil jarak seperti pada Tabel 3.10 Hasil Perhitungan Jarak *Euclidean*.

Tabel 3.10 Hasil Perhitungan Jarak Euclidaen

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori	Jarak
R38	2	2	3	3	Tidak Efektif	0,000
R29	2	2	3	3	Efektif	0,000
R55	2	2	3	3	Efektif	0,000
R18	2	2	3	3	Tidak Efektif	0,000
R27	2	2	3	3	Efektif	0,000
R53	2	2	3	3	Efektif	0,000
R30	2	2	3	3	Tidak Efektif	0,000
R31	2	2	3	3	Tidak Efektif	0,000
R10	2	2	3	3	Efektif	0,000
R33	2	2	3	3	Tidak Efektif	0,000
R34	2	2	3	3	Tidak Efektif	0,000
R36	2	2	3	3	Tidak Efektif	0,000
R39	2	2	3	3	Efektif	0,000
R42	2	2	3	3	Tidak Efektif	0,000
R1	3	2	3	3	Efektif	1,000
R17	3	2	3	3	Tidak Efektif	1,000
R6	3	2	3	3	Efektif	1,000
R57	2	2	2	3	Tidak Efektif	1,000
R11	3	2	3	3	Efektif	1,000
R12	3	2	3	3	Efektif	1,000
R4	3	2	3	3	Efektif	1,000
R14	3	2	3	3	Efektif	1,000
R16	2	3	3	3	Efektif	1,000

R54	2	3	3	3	Efektif	1,000
R21	3	2	3	3	Efektif	1,000
R22	3	2	3	3	Efektif	1,000
R24	3	2	3	3	Efektif	1,000
R7	3	2	3	3	Efektif	1,000
R28	2	2	2	3	Tidak Efektif	1,000
R40	2	2	2	3	Tidak Efektif	1,000
R41	2	2	2	3	Tidak Efektif	1,000
R43	3	2	3	3	Efektif	1,000
R45	3	2	3	3	Efektif	1,000
R3	3	2	2	3	Efektif	1,414
R23	3	2	2	3	Tidak Efektif	1,414
R51	2	2	2	2	Tidak Efektif	1,414
R9	3	3	3	3	Efektif	1,414
R25	3	3	3	3	Efektif	1,414
R52	2	2	2	2	Tidak Efektif	1,414
R37	2	2	2	2	Tidak Efektif	1,414
R48	2	2	2	2	Tidak Efektif	1,414
R50	2	2	2	2	Tidak Efektif	1,414
R46	1	2	2	3	Tidak Efektif	1,414
R56	2	3	2	2	Tidak Efektif	1,732
R15	3	2	3	1	Tidak Efektif	2,236
R47	1	1	2	1	Tidak Efektif	2,646

4. Pemilihan k Tetangga Terdekat

Hasil perhitungan jarak pada Tabel 3.10 diurutkan dari nilai jarak terkecil hingga terbesar. Tiga data latih dengan jarak terdekat dipilih sebagai tetangga terdekat sesuai dengan nilai k yang telah ditetapkan. Daftar tiga tetangga terdekat disajikan pada Tabel 3.11.

Tabel 3.11 Nilai K = 3

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori	R8
R38	2	2	3	3	Tidak Efektif	0,000
R29	2	2	3	3	Efektif	0,000
R55	2	2	3	3	Efektif	0,000

5. Penentuan Kelas Mayoritas

Penentuan kelas dilakukan dengan melihat kategori yang paling banyak muncul pada tiga data latih terdekat sesuai dengan nilai $K = 3$. Berdasarkan Tabel 3.11, tiga data latih terdekat pada data uji Responden 8 (R8) terdiri dari dua kategori

Efektif dan satu kategori Tidak Efektif. Oleh karena itu, hasil klasifikasi untuk data uji Responden 8 adalah Efektif. Proses klasifikasi yang sama diterapkan pada seluruh data uji sebanyak 11 data. Hasil klasifikasi aktual dan prediksi disajikan pada Tabel 3.12.

Tabel 3.12 Hasil Klasifikasi Metode KNN

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Aktual	Prediksi
R8	2	2	3	3	Efektif	Efektif
R13	3	2	3	3	Efektif	Efektif
R5	3	2	2	3	Efektif	Efektif
R20	1	2	1	1	Tidak Efektif	Tidak Efektif
R2	3	2	2	2	Efektif	Tidak Efektif
R19	2	2	3	2	Tidak Efektif	Tidak Efektif
R26	3	3	3	3	Efektif	Efektif
R44	2	3	3	3	Efektif	Efektif
R35	2	2	2	3	Tidak Efektif	Tidak Efektif
R49	2	2	3	3	Tidak Efektif	Efektif
R32	2	3	3	3	Efektif	Efektif

3.3.6. Evaluasi Model K-Nearest Neighbors (KNN)

Berdasarkan hasil prediksi *K-Nearest Neighbors (KNN)* terhadap 11 data uji, dilakukan penyusunan *confusion matrix* untuk membandingkan hasil prediksi dengan kelas aktual. *Confusion matrix* digunakan untuk mengetahui jumlah prediksi yang sesuai dan tidak sesuai antara kelas aktual dan kelas hasil prediksi.

Tabel 3.13. Confusion Matrix KNN

Confusion Table	AKTUAL	PREDIKSI	
		EFEKTIF	TIDAK EFEKTIF
	EFEKTIF	6 (TP)	1 (FN)
	TIDAK EFEKTIF	1 (FP)	3 (TN)

Berdasarkan tabel *confusion matrix* di atas, diperoleh nilai *True Positive (TP)* sebesar 6, *True Negative (TN)* sebesar 3, *False Positive (FP)* sebesar 1, dan *False Negative (FN)* sebesar 1. Berdasarkan nilai tersebut, diperoleh nilai akurasi, presisi, *recall*, dan *F-Score* sebagai.

1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{6 + 3}{6 + 3 + 1 + 1} = \frac{9}{11} = 0,82 \times 100\% = 82\%$$

2. Precision

$$Precision = \frac{TP}{TP + FP} = \frac{6}{6 + 1} = \frac{6}{7} = 0,86 \times 100\% = 86\%$$

3. Recall

$$Recall = \frac{TP}{TP + FN} = \frac{6}{6 + 1} = \frac{6}{7} = 0,86 \times 100\% = 86\%$$

4. F1-Score

$$\begin{aligned} F1-Score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,86 \times 0,86}{0,86 + 0,86} = 2 \times \frac{0,74}{1,72} = 2 \times 0,43 \\ &= 0,86 \times 100\% = 86\% \end{aligned}$$

3.3.7. Pemodelan Menggunakan Naive Bayes

Pemodelan menggunakan metode *Naive Bayes* pada penelitian ini dilakukan dengan perhitungan manual terhadap data latih dan satu data uji. Data latih digunakan untuk menghitung nilai probabilitas awal dan probabilitas kondisional setiap atribut. Nilai-nilai probabilitas tersebut kemudian digunakan untuk menghitung probabilitas akhir masing-masing kategori pada data uji. Hasil klasifikasi ditentukan berdasarkan kategori dengan nilai probabilitas terbesar.

1. Penentuan Probabilitas Awal (Prior)

Penentuan probabilitas awal dilakukan berdasarkan distribusi kategori pada data latih. Data latih yang digunakan dalam perhitungan manual *Naive Bayes* disajikan pada Tabel 3.14, sedangkan distribusi jumlah kategori ditunjukkan pada Tabel 3.15.

Tabel 3.14 Data Latih Naive Bayes

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R1	Tinggi	Sedang	Baik	Baik	Efektif
R17	Tinggi	Sedang	Baik	Baik	Tidak Efektif
R3	Tinggi	Sedang	Cukup	Baik	Efektif
R23	Tinggi	Sedang	Cukup	Baik	Tidak Efektif
R38	Sedang	Sedang	Baik	Baik	Tidak Efektif
R6	Tinggi	Sedang	Baik	Baik	Efektif
R29	Sedang	Sedang	Baik	Baik	Efektif
R51	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R9	Tinggi	Tinggi	Baik	Baik	Efektif
R57	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R11	Tinggi	Sedang	Baik	Baik	Efektif
R12	Tinggi	Sedang	Baik	Baik	Efektif
R4	Tinggi	Sedang	Baik	Baik	Efektif
R14	Tinggi	Sedang	Baik	Baik	Efektif
R55	Sedang	Sedang	Baik	Baik	Efektif
R16	Sedang	Tinggi	Baik	Baik	Efektif
R56	Sedang	Tinggi	Cukup	Cukup	Tidak Efektif
R18	Sedang	Sedang	Baik	Baik	Tidak Efektif
R54	Sedang	Tinggi	Baik	Baik	Efektif
R47	Rendah	Rendah	Cukup	Kurang	Tidak Efektif
R21	Tinggi	Sedang	Baik	Baik	Efektif
R22	Tinggi	Sedang	Baik	Baik	Efektif
R15	Tinggi	Sedang	Baik	Kurang	Tidak Efektif
R24	Tinggi	Sedang	Baik	Baik	Efektif
R25	Tinggi	Tinggi	Baik	Baik	Efektif
R7	Tinggi	Sedang	Baik	Baik	Efektif
R27	Sedang	Sedang	Baik	Baik	Efektif
R28	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R53	Sedang	Sedang	Baik	Baik	Efektif
R30	Sedang	Sedang	Baik	Baik	Tidak Efektif
R31	Sedang	Sedang	Baik	Baik	Tidak Efektif
R10	Sedang	Sedang	Baik	Baik	Efektif
R33	Sedang	Sedang	Baik	Baik	Tidak Efektif

R34	Sedang	Sedang	Baik	Baik	Tidak Efektif
R52	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R36	Sedang	Sedang	Baik	Baik	Tidak Efektif
R37	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R48	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R39	Sedang	Sedang	Baik	Baik	Efektif
R40	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R41	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R42	Sedang	Sedang	Baik	Baik	Tidak Efektif
R43	Tinggi	Sedang	Baik	Baik	Efektif
R50	Sedang	Sedang	Cukup	Cukup	Tidak Efektif
R45	Tinggi	Sedang	Baik	Baik	Efektif
R46	Rendah	Sedang	Cukup	Baik	Tidak Efektif

Tabel 3.15 Distribusi Kategori Data Latih

Kategori	Jumlah Data
Efektif	23
Tidak Efektif	23
Total	46

Tabel 3.16 Probabilitas Awal (Prior)

Kategori	Jumlah Data
Efektif	$23/46 = 0,50$
Tidak Efektif	$23/46 = 0,50$
Total	100%

2. Perhitungan Probabilitas Kondisional

Perhitungan probabilitas kondisional dilakukan dengan menghitung frekuensi kemunculan setiap nilai atribut pada masing-masing kategori berdasarkan data latih. Hasil perhitungan probabilitas kondisional untuk setiap atribut disajikan pada tabel berikut.

Tabel 3.17 Probabilitas Kondisional Atribut Kehadiran Siswa

Kehadiran Siswa	Efektif	Tidak Efektif	P(Efektif)	P(Tidak Efektif)
Tinggi	15	3	$15/23 = 0,65$	$3/23 = 0,13$
Sedang	8	18	$8/23 = 0,35$	$18/23 = 0,78$
Rendah	0	2	$0/23 = 0,00$	$2/23 = 0,09$
Total	23	23	100%	100%

Tabel 3.18 Probabilitas Kondisional Atribut Semangat Belajar

Semangat Belajar	Efektif	Tidak Efektif	P(Efektif)	P(Tidak Efektif)
Tinggi	4	0	$4/23 = 0,17$	$0/23 = 0,00$
Sedang	19	22	$19/23 = 0,83$	$22/23 = 0,96$
Rendah	0	1	$0/23 = 0,00$	$1/23 = 0,04$
Total	23	23	100%	100%

Tabel 3.19 Probabilitas Kondisional Atribut Persepsi Siswa

Persepsi Siswa	Efektif	Tidak Efektif	P(Efektif)	P(Tidak Efektif)
Baik	22	10	$22/23 = 0,96$	$10/23 = 0,43$
Cukup	1	13	$1/23 = 0,04$	$13/23 = 0,57$
Kurang	0	0	$0/23 = 0,00$	$0/23 = 0,00$
Total	23	23	100%	100%

Tabel 3.20 Probabilitas Kondisional Atribut Kualitas Pelaksanaan MBG

Kualitas Pelaksanaan MBG	Efektif	Tidak Efektif	P(Efektif)	P(Tidak Efektif)
Baik	23	15	$23/23 = 1,00$	$15/23 = 0,65$
Cukup	0	6	$0/23 = 0,00$	$6/23 = 0,26$
Kurang	0	2	$0/23 = 0,00$	$2/23 = 0,09$
Total	23	23	100%	100%

3. Penentuan Kelas Berdasarkan Probabilitas Tertinggi

Data uji yang digunakan pada metode *Naive Bayes* disajikan secara keseluruhan pada Tabel 3.21. Dari seluruh data uji tersebut, satu data dipilih sebagai contoh untuk menjelaskan proses perhitungan manual secara rinci.

Tabel 3.21 Data Uji Naive Bayes

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Kategori
R8	Sedang	Sedang	Baik	Baik	Efektif
R13	Tinggi	Sedang	Baik	Baik	Efektif
R5	Tinggi	Sedang	Cukup	Baik	Efektif
R20	Rendah	Sedang	Kurang	Kurang	Tidak Efektif
R2	Tinggi	Sedang	Cukup	Cukup	Efektif
R19	Sedang	Sedang	Baik	Cukup	Tidak Efektif

R26	Tinggi	Tinggi	Baik	Baik	Efektif
R44	Sedang	Tinggi	Baik	Baik	Efektif
R35	Sedang	Sedang	Cukup	Baik	Tidak Efektif
R49	Sedang	Sedang	Baik	Baik	Tidak Efektif
R32	Sedang	Tinggi	Baik	Baik	Efektif

Berdasarkan data latih yang telah disusun sebelumnya, perhitungan dilakukan untuk menentukan kategori kualitas pelaksanaan Program Makan Bergizi Gratis menggunakan metode *Naive Bayes*. Perhitungan probabilitas akhir masing-masing kategori dilakukan menggunakan rumus *Naive Bayes* sebagai berikut:

$$P(A | B) = \frac{P(B | A) \times P(A)}{P(B)}$$

$P(A|B)$ = Probabilitas A bersyarat yang diberikan oleh B

$P(B|A)$ = Probabilitas B bersyarat yang diberikan oleh A

$P(A)$ = Probabilitas kejadian A

$P(B)$ = Probabilitas kejadian B

Contoh perhitungan manual dilakukan pada data uji Responden 8. Adapun perhitungannya sebagai berikut.

$$P(\text{Kategori}) = P(\text{Kehadiran Siswa}|\text{Sedang}) \times P(\text{Semangat Belajar}|\text{Sedang}) \times P(\text{Persepsi Siswa}|\text{Baik}) \times P(\text{Kualitas Pelaksanaan MBG}|\text{Baik}) \times P(\text{Kategori}|\text{Efektif})$$

$$\begin{aligned}
P(\text{Efektif}) &= P(\text{Sedang}|\text{Efektif}) \times P(\text{Sedang}|\text{Efektif}) \times P(\text{Baik}|\text{Efektif}) \times P(\text{Baik}|\text{Efektif}) \times P(\text{Kategori}|\text{Efektif}) \\
&= \left(\frac{8}{23}\right) \times \left(\frac{19}{23}\right) \times \left(\frac{22}{23}\right) \times \left(\frac{23}{23}\right) \times \left(\frac{23}{46}\right) \\
&= 0,1374(\text{Nilai Efektif})
\end{aligned}$$

$$\begin{aligned}
P(\text{Tidak Efektif}) &= P(\text{Sedang}|\text{Tidak Efektif}) \times P(\text{Sedang}|\text{Tidak Efektif}) \times \\
&P(\text{Baik}|\text{Tidak Efektif}) \times P(\text{Baik}|\text{Tidak Efektif}) \times \\
&P(\text{Kategori}|\text{Tidak Efektif}) \\
&= \left(\frac{18}{23}\right) \times \left(\frac{22}{23}\right) \times \left(\frac{10}{23}\right) \times \left(\frac{15}{23}\right) \times \left(\frac{23}{46}\right) \\
&= 0,1061 \text{ (Nilai Tidak Efektif)}
\end{aligned}$$

Berdasarkan hasil perhitungan manual menggunakan metode *Naive Bayes* pada salah satu data uji (R8), diperoleh nilai probabilitas untuk kategori Efektif sebesar 0,1374, sedangkan nilai probabilitas untuk kategori Tidak Efektif sebesar 0,1061. Dengan demikian, data uji tersebut diklasifikasikan ke dalam kategori Efektif karena memiliki nilai probabilitas tertinggi. Hasil klasifikasi data uji lainnya selanjutnya disajikan dalam bentuk tabel.

Tabel 3.22 . Hasil Klasifikasi Metode Naive Bayes

Responden	Kehadiran Siswa	Semangat Belajar	Persepsi Siswa	Kualitas Pelaksanaan MBG	Aktual	Prediksi	Efektif	Tidak Efektif
R8	Sedang	Sedang	Baik	Baik	Efektif	Efektif	0,1374	0,1061
R13	Tinggi	Sedang	Baik	Baik	Efektif	Efektif	0,2577	0,0177
R5	Tinggi	Sedang	Cukup	Baik	Efektif	Tidak Efektif	0,0117	0,0230
R20	Rendah	Sedang	Kurang	Kurang	Tidak Efektif	Tidak Efektif	0,0000	0,0000
R2	Tinggi	Sedang	Cukup	Cukup	Efektif	Tidak Efektif	0,0000	0,0092
R19	Sedang	Sedang	Baik	Cukup	Tidak Efektif	Tidak Efektif	0,0000	0,0425
R26	Tinggi	Tinggi	Baik	Baik	Efektif	Efektif	0,0542	0,0000
R44	Sedang	Tinggi	Baik	Baik	Efektif	Efektif	0,0289	0,0000
R35	Sedang	Sedang	Cukup	Baik	Tidak Efektif	Tidak Efektif	0,0062	0,1380
R49	Sedang	Sedang	Baik	Baik	Tidak Efektif	Efektif	0,1374	0,1061
R32	Sedang	Tinggi	Baik	Baik	Efektif	Efektif	0,0289	0,0000

Data pada Tabel 3.22 merupakan hasil perhitungan manual metode *Naive Bayes* yang dilakukan menggunakan Microsoft Excel dengan menampilkan nilai probabilitas untuk setiap kategori serta hasil prediksi pada masing-masing data uji. Berdasarkan hasil klasifikasi tersebut, model mampu mengelompokkan sebagian besar data uji ke dalam kategori Efektif dan Tidak Efektif berdasarkan nilai probabilitas tertinggi yang diperoleh. Meskipun demikian, masih terdapat beberapa kesalahan prediksi berupa *false positive* dan *false negative*, yaitu kondisi ketika hasil prediksi model tidak sesuai dengan kategori aktual. Adanya kesalahan klasifikasi tersebut menunjukkan bahwa model *Naive Bayes* belum sepenuhnya mampu mengklasifikasikan seluruh data uji secara akurat.

3.3.8. Evaluasi Model Naive Bayes

Berdasarkan hasil prediksi *Naive Bayes* terhadap 11 data uji, dilakukan penyusunan *confusion matrix* untuk membandingkan hasil prediksi dengan kelas aktual. *Confusion matrix* digunakan untuk mengetahui jumlah prediksi yang sesuai dan tidak sesuai antara kelas aktual dan kelas hasil prediksi.

Tabel 3.23. Confusion Matrix Naive Bayes

Confusion Table	AKTUAL	PREDIKSI	
		EFEKTIF	TIDAK EFEKTIF
	EFEKTIF	5 (TP)	2 (FN)
	TIDAK EFEKTIF	1 (FP)	3 (TN)

Berdasarkan tabel *confusion matrix* di atas, diperoleh nilai *True Positive (TP)* sebesar 5, *True Negative (TN)* sebesar 3, *False Positive (FP)* sebesar 1, dan *False Negative (FN)* sebesar 2. Berdasarkan nilai tersebut, diperoleh nilai akurasi, presisi, *recall*, dan *F-Score* sebagai berikut.

1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{5 + 3}{5 + 3 + 1 + 2} = \frac{8}{11} = 0,73 \times 100\% = 73\%$$

2. Precision

$$Precision = \frac{TP}{TP + FP} = \frac{5}{5 + 1} = \frac{5}{6} = 0,83 \times 100\% = 83\%$$

3. Recall

$$Recall = \frac{TP}{TP + FN} = \frac{5}{5 + 2} = \frac{5}{7} = 0,71 \times 100\% = 71\%$$

4. F1-Score

$$\begin{aligned} F1-Score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,83 \times 0,71}{0,83 + 0,71} = 2 \times \frac{0,59}{1,54} = 2 \times 0,38 \\ &= 0,76 \times 100\% = 76\% \end{aligned}$$