

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Penelitian ini merupakan penelitian terapan (applied research) dengan pendekatan kualitatif-kuantitatif campuran. Fokus utama adalah mengimplementasikan algoritma Naive Bayes untuk mengklasifikasikan ketepatan diagnosis penyakit lambung berdasarkan gejala pasien di Puskesmas Sei Berombang. Penelitian ini termasuk dalam kategori *predictive analytics* dengan tujuan membangun model klasifikasi yang dapat memprediksi akurasi diagnosis berdasarkan pola gejala.

3.2 Lokasi dan Waktu Penelitian

Lokasi penelitian di Puskesmas Sei Berombang, Kabupaten Labuhanbatu, Sumatera Utara. Penelitian berlangsung pada bulan April - Mei 2025, dengan fokus utama pada gejala penyakit lambung selama tahun 2024. Populasi dan Sampel Penelitian

Seluruh rekam medis pasien rawat jalan di Puskesmas Sei Berombang yang didiagnosis menderita penyakit lambung selama tahun 2024 ($n = 90$ pasien). Teknik purposive sampling dengan kriteria: Diagnosis meliputi 5 jenis penyakit lambung (Gastritis, Tukak Lambung, GERD, Dispepsia, Kanker Lambung), Kelengkapan data gejala minimal 80%, dan Pembagian data.

Populasi penelitian adalah seluruh pasien Puskesmas Sei Berombang yang telah mendapatkan diagnosis terkait gejala penyakit lambung selama tahun 2024.

Sampel penelitian dipilih secara *purposive* dengan kriteria pasien yang memiliki riwayat gejala spesifik penyakit lambung, seperti nyeri ulu hati, mual, atau muntah, dengan total sampel sebanyak 90 pasien.

3.3 Teknik Pengumpulan Data

Penulis menggunakan berbagai sumber data. Adapun data yang digunakan adalah data diperoleh langsung dari hasil wawancara dengan tenaga medis, rekam medis pasien, dan observasi langsung terhadap proses diagnosis di Puskesmas Sei Berombang[33]. Lalu, data sekunder berupa literatur, jurnal ilmiah, dan laporan terkait penyakit lambung dan algoritma Naive Bayes yang mendukung penelitian ini[34].

3.4 Alat dan Bahan

RapidMiner Studio merupakan salah satu aplikasi pendukung pengujian, dan library pendukung seperti untuk implementasi algoritma *Naive Bayes*[20]. Microsoft Excel untuk preprocessing data. Adapun hardwarenya seperti laptop dengan spesifikasi minimal prosesor Intel Core i5 dan RAM 8GB. Dan, jaringan internet untuk akses sumber data sekunder.

3.5 Menghitung Jumlah Class atau Label

Berikut jumlah class dalam label:

Tabel 3. 1 Hasil Evaluasi Status

No.	ID Pasien	Jenis Kelamin	Gejala	Diagnosis	Status
1	P1	Perempuan	G12,G18,G17,G14	Gastritis	Tidak Tepat
2	P2	Laki-Laki	G5	Gastritis	Tepat
3	P3	Perempuan	G18	GERD	Tidak Tepat
4	P4	Perempuan	G4,G15	Gastritis	Tepat
5	P5	Laki-Laki	G8,G9	Tukak Lambung	Tepat
6	P6	Perempuan	G3,G14	Gastritis	Tepat
7	P7	Laki-Laki	G17	GERD	Tidak Tepat

8	P8	Perempuan	G13,G18	GERD	Tidak Tepat
9	P9	Perempuan	G18,G15,G13,G8	Tukak Lambung	Tepat
10	P10	Laki-Laki	G17	GERD	Tidak Tepat
11	P11	Perempuan	G15,G9,G4,G14	Dispepsia	Tepat
12	P12	Perempuan	G8,G16	Tukak Lambung	Tepat
13	P13	Laki-Laki	G4,G17,G15,G14	Tukak Lambung	Tepat
14	P14	Perempuan	G9,G5	GERD	Tepat
15	P15	Perempuan	G16,G10,G13	GERD	Tidak Tepat
16	P16	Laki-Laki	G15,G3,G14,G12	Gastritis	Tepat
17	P17	Perempuan	G9,G18,G5,G6	GERD	Tidak Tepat
18	P18	Perempuan	G6,G18,G10,G13	Gastritis	Tidak Tepat
19	P19	Laki-Laki	G10,G4	Tukak Lambung	Tidak Tepat
20	P20	Perempuan	G14	Dispepsia	Tepat
21	P21	Perempuan	G11,G14,G8	GERD	Tidak Tepat
22	P22	Laki-Laki	G3,G10	Gastritis	Tepat
23	P23	Perempuan	G8	Tukak Lambung	Tepat
24	P24	Perempuan	G17,G10,G6	Tukak Lambung	Tepat
25	P25	Laki-Laki	G14,G16,G8	Gastritis	Tidak Tepat
26	P26	Perempuan	G17,G18,G17	Tukak Lambung	Tepat
27	P27	Perempuan	G16	Gastritis	Tidak Tepat
28	P28	Laki-Laki	G16,G12,G10	GERD	Tidak Tepat
29	P29	Perempuan	G15,G17	Dispepsia	Tepat
30	P30	Perempuan	G8,G4,G15,G11	Tukak Lambung	Tepat
31	P31	Laki-Laki	G17,G3,G17,G15	GERD	Tidak Tepat
32	P32	Perempuan	G3	Gastritis	Tepat
33	P33	Perempuan	G8,G17,G15	Tukak Lambung	Tepat
34	P34	Laki-Laki	G17,G13	GERD	Tidak Tepat
35	P35	Perempuan	G8,G9,G16,G10	Tukak Lambung	Tepat
36	P36	Perempuan	G6	Tukak Lambung	Tepat
37	P37	Laki-Laki	G17,G15,G3,G19	GERD	Tidak Tepat
38	P38	Perempuan	G6,G18	Tukak Lambung	Tepat
39	P39	Perempuan	G16	GERD	Tidak Tepat
40	P40	Laki-Laki	G15,G11,G5,G16	Gastritis	Tidak Tepat
41	P41	Perempuan	G15,G16,G17	Kanker Lambung	Tepat
42	P42	Perempuan	G6,G9	Tukak Lambung	Tepat
43	P43	Laki-Laki	G15,G18	Dispepsia	Tepat
44	P44	Perempuan	G17,G13	GERD	Tidak Tepat
45	P45	Perempuan	G6	Tukak Lambung	Tepat
46	P46	Laki-Laki	G11	Dispepsia	Tepat
47	P47	Perempuan	G18,G17,G11	GERD	Tidak Tepat
48	P48	Perempuan	G4,G8,G17	Gastritis	Tidak Tepat
49	P49	Laki-Laki	G9,G15,G15,G12	Tukak Lambung	Tepat
50	P50	Perempuan	G5,G10	Gastritis	Tepat
51	P51	Perempuan	G12,G19,G15,G10	Dispepsia	Tepat

52	P52	Laki-Laki	G15,G17	Dispepsia	Tepat
53	P53	Perempuan	G6,G15,G5	Dispepsia	Tepat
54	P54	Perempuan	G6	Tukak Lambung	Tepat
55	P55	Laki-Laki	G9,G14,G15,G5	GERD	Tepat
56	P56	Perempuan	G6,G11,G14	Tukak Lambung	Tepat
57	P57	Perempuan	G6	Tukak Lambung	Tepat
58	P58	Laki-Laki	G10,G13,G15	GERD	Tepat
59	P59	Perempuan	G17,G9,G19,G14	GERD	Tidak Tepat
60	P60	Perempuan	G12,G15	GERD	Tepat
61	P61	Laki-Laki	G10,G18,G14	GERD	Tidak Tepat
62	P62	Perempuan	G9,G5	GERD	Tepat
63	P63	Perempuan	G4,G11	Gastritis	Tepat
64	P64	Laki-Laki	G19	Gastritis	Tepat
65	P65	Perempuan	G10,G5,G12	GERD	Tepat
66	P66	Perempuan	G3,G16,G13,G12	Gastritis	Tidak Tepat
67	P67	Laki-Laki	G8	Gastritis	Tidak Tepat
68	P68	Perempuan	G15,G5,G4	Dispepsia	Tepat
69	P69	Perempuan	G13,G3,G15,G13	Gastritis	Tepat
70	P70	Laki-Laki	G4	Gastritis	Tepat
71	P71	Perempuan	G4	Gastritis	Tepat
72	P72	Perempuan	G4,G10	Gastritis	Tepat
73	P73	Laki-Laki	G10,G4,G5,G3	GERD	Tepat
74	P74	Perempuan	G6,G15,G13,G14	Tukak Lambung	Tepat
75	P75	Perempuan	G15,G3,G17	Dispepsia	Tepat
76	P76	Laki-Laki	G15,G9,G5,G13	Gastritis	Tidak Tepat
77	P77	Perempuan	G11,G10	GERD	Tepat
78	P78	Perempuan	G11,G12,G16	GERD	Tidak Tepat
79	P79	Laki-Laki	G16,G13,G14	Tukak Lambung	Tepat
80	P80	Perempuan	G3,G10	Gastritis	Tepat
81	P81	Perempuan	G4,G19,G16	Gastritis	Tidak Tepat
82	P82	Laki-Laki	G13,G11,G15,G13	Gastritis	Tepat
83	P83	Perempuan	G19	GERD	Tidak Tepat
84	P84	Perempuan	G8,G15,G6	Tukak Lambung	Tepat
85	P85	Laki-Laki	G10,G15,G16,G13	GERD	Tidak Tepat
86	P86	Perempuan	G17,G4	Tukak Lambung	Tepat
87	P87	Perempuan	G4,G19,G15	Gastritis	Tidak Tepat
88	P88	Laki-Laki	G10,G13,G8	GERD	Tidak Tepat
89	P89	Perempuan	G11	GERD	Tepat
90	P90	Perempuan	G12	Tukak Lambung	Tepat

Sumber: Penelitian (2025)

Data diatas merupakan data pasien yang akan digunakan untuk membantu proses klasifikasi ataupun perhitungan pada metode *naïve bayes*. Tetapi tidak

hanya itu saja, data diatas akan dibagi dan dipecah menjadi beberapa tabel. Tabel diatas merupakan data seluruh pasien digunakan untuk menguji penelitian.

3.5.1 Menghitung Jumlah Kelas (Label) Penyakit

Kelas penyakit yang diteliti adalah Gastritis (Radang Lambung), Tukak Lambung (Ulkus Peptikum), Gastroesophageal Reflux Disease (GERD), Dispepsia (Sakit Maag Fungsional), dan Kanker Lambung. Adapun rumus dari probabilitas prior adalah sebagai berikut:

$$P(C) = \frac{\text{Jumlah data dalam kelas } C}{\text{Jumlah total data}} \quad (\text{Sumber: Thomas Bayes})$$

Penjelasan Rumus:

- Jumlah data dalam kelas C ini adalah jumlah pasien dalam data kita yang diketahui menderita penyakit lambung (C).

Jumlah total data ini adalah total jumlah pasien dalam data kita, tanpa memandang jenis penyakit lambung yang mereka derita.

3.5.2 Menghitung Probabilitas (Peluang) dari Masing-Masing Gejala

Penghitungan probabilitas dari masing-masing gejala bertujuan untuk mendapatkan peluang kemunculan gejala tersebut pada setiap penyakit lambung.

$$P(X_i/C) = \frac{\text{Jumlah sampel dengan fitur } X_i \text{ pada kelas } C}{\text{Jumlah sampel dengan fitur } X_i \text{ pada kelas } C} \quad (\text{Sumber: Thomas Bayes})$$

Penjelasan Rumus:

- Jumlah sampel dengan fitur X pada kelas C ini adalah jumlah pasien dalam data kita yang menderita penyakit lambung (C) dan juga memiliki gejala (X).

- Jumlah sampel pada kelas C ini adalah total jumlah pasien dalam data kita yang menderita penyakit lambung (C), tanpa memandang apakah mereka memiliki gejala (X) atau tidak.

3.5.3 Menggunakan *Naïve Bayes* untuk Menentukan Prediksi Penyakit

Setelah probabilitas masing-masing gejala dihitung, proses selanjutnya yaitu menggunakan algoritma *Naïve Bayes* untuk melakukan prediksi penyakit lambung berdasarkan gejala yang dimiliki pasien.

Rumus:

$$P(C/X) = \frac{P(X|C)P(C)}{P(X)} \quad (\text{Sumber: Thomas Bayes})$$

Dimana:

- $P(X/C)$ adalah *likelihood probability*
- $P(C)$ adalah *prior probability*
- $P(X)$ adalah probabilitas total fitur yang dapat diabaikan dalam perbandingan

Algoritma *Naïve Bayes* yang diterapkan pada data pasien Puskesmas Sei Berombang dapat memberikan prediksi penyakit lambung berdasarkan gejala yang dialami pasien.

3.5.4 Pengujian Tingkat Akurasi dengan Excel

Pengujian tingkat akurasi bertujuan untuk melihat seberapa besar kesalahan dan kebenaran prediksi yang dihasilkan oleh algoritma *Naïve Bayes*. Pengujian akurasi yang diterapkan menggunakan Microsoft Excel. Langkahnya yaitu:

1. Mengumpulkan data prediksi dan data yang sesungguhnya.

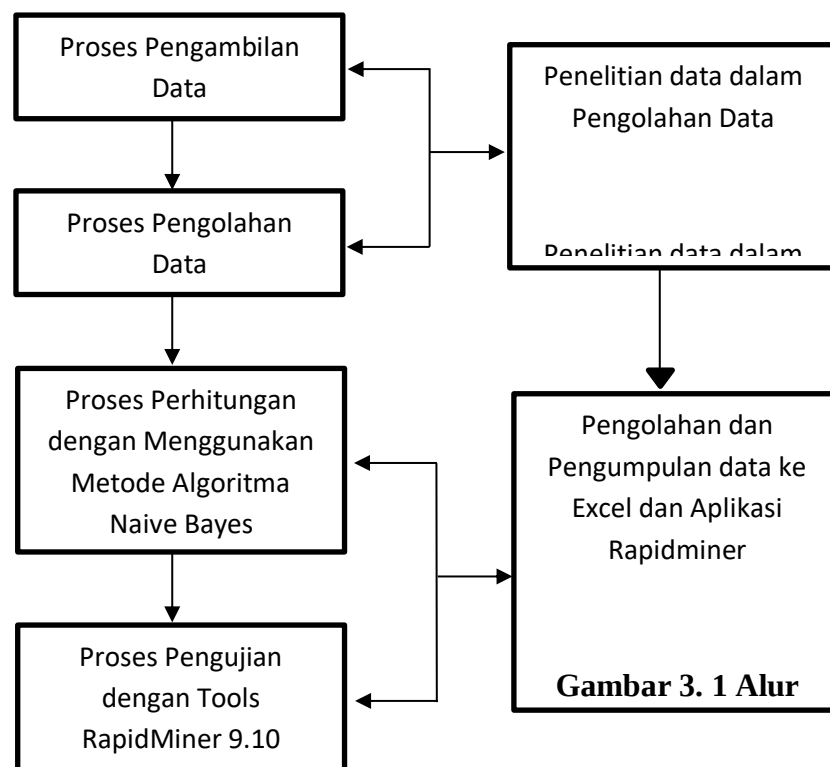
2. Menghitung jumlah prediksi yang sesuai (TRUE) dan yang tidak sesuai (FALSE).
3. Menghitung akurasi dengan formula:

3.5.5 Pengujian Tingkat Akurasi dengan RapidMiner

Selain menggunakan Microsoft Excel, pengujian juga diterapkan pada aplikasi RapidMiner untuk mendapatkan hasil yang lebih rinci mengenai performa algoritma.

RapidMiner juga menampilkan matriks konfusi, yang berguna untuk melihat rincian prediksi yang sesuai dan yang tidak sesuai, sehingga dapat dianalisa lebih luas mengenai kinerja algoritma yang diterapkan.

3.6 Diagram Penelitian



Gambar 3. 1 Alur