

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Kanker merupakan penyakit yang ditandai oleh pertumbuhan sel ganas pada jaringan tubuh akibat adanya sel abnormal yang berkembang secara tidak terkendali dan tidak mengikuti mekanisme pengaturan pertumbuhan sel yang normal. Kondisi tersebut menjadikan kanker sebagai salah satu permasalahan kesehatan utama yang berkontribusi terhadap tingginya angka kematian di berbagai negara di dunia. (Agustin, 2020). Penyakit ini termasuk ke dalam kelompok penyakit tidak menular (PTM) yang menunjukkan kecenderungan peningkatan jumlah kasus dari tahun ke tahun. Secara global, jumlah kasus baru kanker mencapai sekitar 20 juta orang pada tahun 2022. Sementara itu, jumlah kematian yang disebabkan kanker mencapai 9,7 juta jiwa atau sekitar 18,7% dari seluruh kematian di seluruh dunia per tahun. Prediksi menunjukkan bahwa jumlah kasus kanker berpotensi meningkat hingga 35 juta kasus per tahun pada 2050, menjadikan kanker sebagai salah satu ancaman utama bagi kesehatan global (Permenkes RI, 2024).

Berdasarkan laporan Global Cancer Observatory (GLOBOCAN) tahun 2022, jenis kanker dengan prevalensi tertinggi di Indonesia adalah kanker payudara, yang mendominasi angka kejadian dibandingkan jenis kanker lainnya. Pada tahun 2022, tercatat sebanyak 66.271 kasus baru kanker payudara di Indonesia, yang mewakili sekitar 16,2% dari total 408.661 kasus kanker baru pada periode tersebut. Selain menunjukkan insidensi yang tinggi, kanker payudara juga termasuk salah satu penyebab utama kematian akibat kanker di Indonesia. Tahun 2022 mencatat jumlah kematian yang disebabkan oleh penyakit ini tercatat mencapai lebih dari 22 ribu jiwa (Dewi *et al.*, 2025).

Menurut Permenkes Nomor 34 Tahun 2015, *carcinoma mammae* atau kanker payudara diidentifikasi sebagai manifestasi keganasan biologis yang menginvasi jaringan penunjang, duktus (saluran), serta lobulus (kelenjar) pada payudara akibat proliferasi seluler yang abnormal, dengan tidak melibatkan jaringan kulit sebagai asal pertumbuhan kanker. Sebagai jenis kanker dengan

insidensi tertinggi di Indonesia, kanker payudara memiliki karakteristik perkembangan penyakit yang perlu mendapat perhatian khusus. Pada tahap awal, pertumbuhan kanker payudara biasanya berkembang secara perlahan dan sering kali tanpa gejala yang jelas, sehingga banyak kasus baru diketahui setelah memasuki Stadium lanjut (Rosyidah & Supriani, 2023). Proses perkembangan kanker payudara dari Stadium awal hingga mencapai tahap yang tidak dapat disembuhkan dapat berlangsung dalam kurun waktu sekitar satu tahun. Pada fase awal, sel kanker biasanya mengalami pertumbuhan secara perlahan hingga membentuk tumor berdiameter ± 1 cm dalam rentang waktu 8–12 tahun. Sel-sel yang berpotensi menyebabkan kanker tersebut dapat berada dalam keadaan dorman di dalam tubuh penderita sebelum akhirnya menjadi aktif dan melanjutkan proses pertumbuhannya akan lebih cepat dan menyebar melalui aliran darah. Penanganan yang tidak dilakukan secara tepat waktu berpotensi menyebabkan penyebaran sel kanker tidak terdeteksi secara dini. Sel-sel kanker tersebut dapat terus bertumbuh hingga bersifat invasif dan tetap berada dalam keadaan dorman dalam jangka waktu yang panjang. Ketika kembali aktif, sel-sel tersebut dapat berkembang menjadi neoplasma ganas yang selanjutnya diidentifikasi sebagai kanker payudara (Savitri *et al.*, 2019).

Berdasarkan Profil Kesehatan Jawa Timur tahun 2024, cakupan pelaksanaan program deteksi dini kanker payudara di Provinsi Jawa Timur pada kelompok perempuan berusia 30–50 tahun berhasil menjangkau sebanyak 2.995.992 orang, atau sekitar 46,1% dari keseluruhan wanita usia subur dalam rentang usia tersebut. Dari hasil pemeriksaan deteksi dini, ditemukan sebanyak 6.638 orang atau 0,2% dengan kasus mengalami benjolan pada payudara dari keseluruhan perempuan yang telah diperiksa (Dinas Kesehatan Provinsi Jawa timur, 2023). Pada tingkat kabupaten, Profil Kesehatan Kabupaten Jember Tahun 2024 melaporkan bahwa dari 269.286 perempuan dalam rentang usia 30–50 tahun yang mengikuti pelaksanaan skrining dini terhadap kanker serviks dan kanker payudara, tercatat sebanyak 47 orang atau setara dengan 0,03% ditemukan memiliki tumor atau benjolan payudara dan 4 orang dikategorikan sebagai suspek kanker payudara, dengan total 48 kasus dirujuk ke fasilitas pelayanan kesehatan lanjutan

untuk pemeriksaan dan penegakan diagnosis lebih lanjut. Kondisi ini menunjukkan bahwa meskipun proporsi temuan relatif kecil, hasil deteksi dini tetap mengindikasikan adanya potensi kasus kanker payudara yang memerlukan penanganan lanjutan di rumah sakit rujukan.

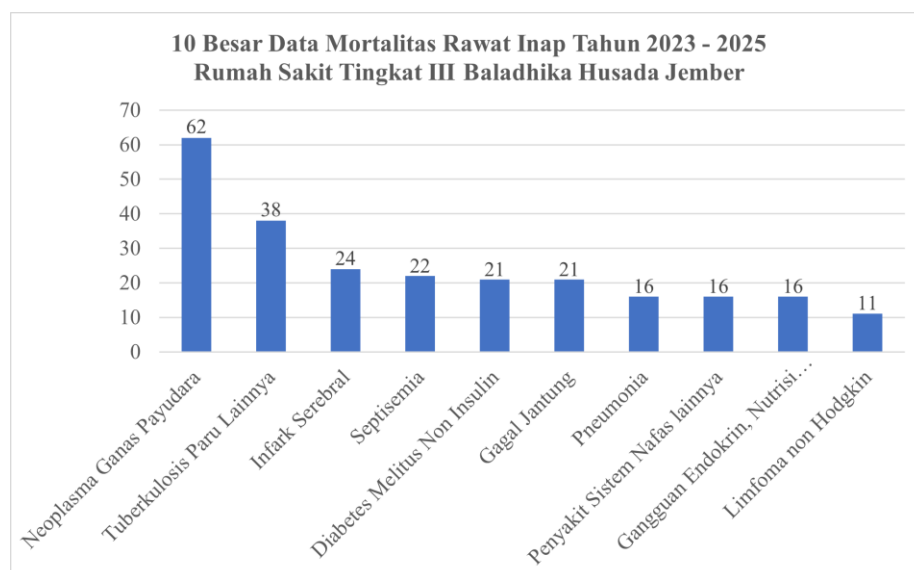
Rumah Sakit Tingkat III Baladhika Husada Jember merupakan rumah sakit umum tipe C yang berada di bawah naungan Tentara Nasional Indonesia Angkatan Darat (TNI AD). Fasilitas kesehatan ini dilengkapi dengan layanan medik dasar dan spesialis, termasuk fasilitas pelayanan lanjutan yang mendukung diagnosis dan penanganan penyakit. Sebagai salah satu rumah sakit rujukan di Kabupaten Jember dengan jumlah kunjungan pasien kanker payudara yang relatif tinggi. Rumah Sakit Tingkat III Baladhika Husada memiliki layanan unggulan yaitu poli onkologi dan kemoterapi yang menjadi pusat penanganan pasien dengan berbagai jenis kanker, termasuk kanker payudara. Hasil studi pendahuluan pada Instalasi Rekam Medis menunjukkan bahwa kanker payudara masuk dalam kategori sepuluh besar penyakit rawat inap pada tahun 2025, sebagaimana tersaji pada tabel berikut..

Tabel 1.1 10 Besar Penyakit Rawat Inap 2025

No.	Kode ICD-10	Nama Diagnosis	Jumlah	%
1	C50	Neoplasma ganas payudara	1.105	23,10%
2	A01.0	Demam tifoid dan paratifoid	1.064	22,24%
3	N25-N29, N31-N39	Penyakit sistem kemih lainnya	423	8,84%
4	J12-J18	Pneumonia	394	8,24%
5	D51-D58, D60, D62	Anemia lainnya	373	7,80%
6	P00-P04	Janin dan neonatus yang dipengaruhi faktor maternal serta komplikasi obstetri	326	6,81%
7	D10-D12, D50-D79	Neoplasma jinak lainnya	312	6,52%
8	A09	Diare dan gastroenteritis oleh penyebab infeksius yang tidak spesifik	282	5,89%
9	P08, P29, P50-P54	Gangguan lain yang muncul sejak periode perinatal	273	5,71%
10	C18	Neoplasma ganas colon	232	4,85%
Jumlah			4.784	100%

Sumber Data : Data Sekunder Tahun 2025

Berdasarkan data pelaporan 10 besar penyakit rawat inap di Rumah Sakit Tingkat III Baladhika Husada Jember pada tahun 2025, diketahui penyakit neoplasma ganas payudara, yang secara medis dikenal sebagai *carcinoma mammae* atau kanker payudara berada di urutan pertama dengan total kasus mencapai 1.105 kasus. Angka tersebut tergolong tinggi dan menunjukkan bahwa sekitar 23,10% pasien rawat inap di Rumah Sakit Tingkat III Baladhika Husada Jember mengidap penyakit neoplasma ganas payudara atau kanker payudara. Tingginya jumlah kasus kanker payudara tidak hanya berdampak pada tingginya morbiditas, tetapi juga berkaitan dengan kematian pasien. Data mortalitas kanker payudara pada pasien rawat inap tahun 2025 dapat dilihat pada gambar di bawah ini.



Gambar 1.1 10 Besar Mortalitas Rawat Inap Kanker Payudara 2023–2025

Berdasarkan gambar 1.1 angka kematian akibat kanker payudara berada di urutan pertama, hal ini menunjukkan adanya kebutuhan untuk memberikan perhatian khusus pada upaya deteksi dini, sehingga langkah pencegahan dan penanganan dapat dilakukan secara lebih cepat dan tepat waktu. Hingga saat ini, penyebab pasti terjadinya kanker payudara belum sepenuhnya dapat dipastikan. Namun demikian, terdapat sejumlah faktor risiko utama yang memiliki keterkaitan erat dengan munculnya kanker payudara pada wanita. Menurut Keputusan Menteri Kesehatan Tahun 2018 tentang Pedoman Nasional Pelayanan Kedokteran Tata Laksana Kanker Payudara, faktor risiko yang berhubungan dengan kanker payudara

mencakup jenis kelamin perempuan, usia lebih dari 50 tahun, riwayat keluarga dengan kanker payudara, riwayat penyakit payudara sebelumnya, menarche dini (kurang dari 12 tahun) atau menopause yang terlambat (lebih dari 55 tahun), aspek reproduksi dan hormonal, obesitas, konsumsi alkohol, riwayat paparan radiasi pada dinding dada, serta pengaruh lingkungan.

Berdasarkan *literature review*, faktor risiko kanker payudara berkaitan dengan usia, faktor genetik, serta adanya riwayat kanker payudara dalam keluarga, riwayat reproduksi, obesitas, aktivitas fisik (Ketut & Kartika, 2022). Wanita yang memiliki ibu dengan riwayat kanker payudara diketahui memiliki risiko sekitar dua kali lebih besar dibandingkan dengan wanita tanpa riwayat keluarga kanker payudara (Dewi *et al.*, 2025). Secara umum, penderita kanker payudara menampilkan gejala klinis berupa nyeri pada payudara, adanya benjolan di payudara maupun ketiak, perubahan bentuk dan ukuran payudara, keluarnya cairan dari puting, perubahan pada puting susu, serta timbulnya kerutan pada kulit payudara (Savitri *et al.*, 2019).

Tingginya angka insidensi dan mortalitas akibat kanker payudara menunjukkan perlunya dukungan metode analisis seperti *data mining* yang mampu membantu pengambilan keputusan klinis. Penerapan teknologi *data mining* dapat berperan dalam meminimalkan kesalahan serta keterlambatan dalam proses deteksi kanker payudara (Setiani & Trisanti, 2025). *Data mining* merupakan salah satu pendekatan dalam pengolahan data yang bertujuan untuk menemukan pola serta hubungan tertentu yang sebelumnya belum teridentifikasi oleh pengguna, kemudian menyajikannya dalam bentuk yang lebih mudah dipahami sehingga dapat dimanfaatkan sebagai dasar pengambilan keputusan (Putro *et al.*, 2020). Salah satu metode yang digunakan dalam *data mining* adalah klasifikasi, yaitu teknik untuk mengelompokkan data ke dalam kelas tertentu berdasarkan atribut atau karakteristik yang dimilikinya (Meilani & Nurdiawan, 2023). Terdapat beberapa algoritma klasifikasi yang dapat digunakan, salah satunya adalah Algoritma *Naïve Bayes*.

Algoritma *Naïve Bayes* merupakan salah satu metode klasifikasi yang berlandaskan pada *Teorema Bayes*. Metode ini memanfaatkan pendekatan

probabilistik dan statistik untuk melakukan prediksi terhadap kelas suatu objek yang belum diketahui, dengan memanfaatkan informasi dari setiap atribut yang dimiliki objek tersebut (Watratan *et al.*, 2020). Keunggulan utama algoritma ini terletak pada kesederhanaannya, efisiensi komputasi, serta kemampuan dalam memproses *Dataset* skala besar dengan cepat dan menghasilkan tingkat ketepatan yang tinggi.

Dalam bidang kesehatan, algoritma *Naïve Bayes* banyak dimanfaatkan untuk melakukan klasifikasi data medis, termasuk analisis laporan diagnosis maupun rekam medis pasien, sehingga pola-pola relevan pada data kanker payudara dapat teridentifikasi secara lebih efektif (Tanti *et al.*, 2024). Penelitian sebelumnya yang dilakukan oleh Muntiari & Hanif (2022) dengan judul “Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma *Machine Learning*” menunjukkan bahwa algoritma *Naïve Bayes* mampu mencapai tingkat akurasi sebesar 95%, lebih tinggi dibandingkan Algoritma *Neural Network* sebesar 94,25%, serta *Random Forest* dan *Support Vector Machine* akurasi sebesar 94,75%. Selain itu, algoritma klasifikasi lain seperti *Decision Tree* penerapannya cenderung lebih rentan terhadap *overfitting*, terutama jika tidak dikelola dengan baik dan pada data dengan jumlah atribut yang banyak (Munir *et al.*, 2024). Jika dibandingkan dengan algoritma *Support Vector Machine* (SVM), algoritma *Naïve Bayes* secara konsisten menunjukkan waktu pemrosesan yang lebih singkat, sedangkan *Support Vector Machine* (SVM) membutuhkan waktu pemrosesan yang lebih lama (Riyanto, 2018).

Berdasarkan keunggulan tersebut, Algoritma *Naïve Bayes* dinilai sesuai untuk klasifikasi kanker payudara karena karakteristik data yang digunakan berupa data kategorik yang bersifat diskret, yang berasal dari data rekam medis, seperti faktor risiko dan gejala klinis. Selain itu, algoritma *Naïve Bayes* menggunakan pendekatan berbasis probabilitas untuk menentukan kemungkinan suatu data diklasifikasikan ke dalam kelas tertentu berdasarkan atribut yang dimilikinya, sehingga efektif dalam mengklasifikasikan data rekam medis pasien kanker payudara dengan banyak atribut.

Tools yang digunakan untuk melakukan klasifikasi, salah satunya adalah Python. Python merupakan bahasa pemrograman berorientasi objek yang dapat dimanfaatkan dalam pengembangan perangkat lunak serta mendukung eksekusi pada berbagai sistem operasi. Saat ini, Python banyak dimanfaatkan dalam bidang data *science* dan analisis data karena dilengkapi dengan beragam *library* yang menyediakan fungsi analisis, *machine learning*, prapemrosesan data, serta visualisasi yang komprehensif dan bervariasi (Edwardo, 2018 dalam Guna *et al.*, 2023).

Berdasarkan uraian latar belakang di atas, maka penulis tertarik melakukan penelitian berjudul “Klasifikasi Penyakit Kanker Payudara (C50) Menggunakan Algoritma *Naïve Bayes* di Rumah Sakit Tingkat III Baladhika Husada Jember”. Penelitian ini diharapkan dapat membantu memprediksi penyakit kanker payudara dengan tujuan dapat mengurangi jumlah morbiditas dan mortalitas. Pada penelitian ini menggunakan variabel yaitu usia, riwayat kanker payudara dari keluarga, riwayat payudara sebelumnya, status IMT, nyeri payudara, benjolan di payudara, benjolan di ketiak, perubahan bentuk dan ukuran payudara, keluarnya cairan dari payudara, perubahan puting susu, kelainan pada kulit payudara, status lokalis kanker payudara, ukuran tumor payudara.

1.2 Rumusan Masalah

Berdasarkan penjelasan dari latar belakang di atas, maka rumusan masalah pada penelitian ini yaitu sebagai berikut: “Bagaimana penerapan algoritma *Naïve Bayes* dalam mengklasifikasikan penyakit kanker payudara (C50) di Rumah Sakit Tingkat III Baladhika Husada Jember”.

1.3 Tujuan Penelitian

1.3.1 Tujuan Umum

Mengklasifikasikan penyakit kanker payudara (C50) menggunakan Algoritma *Naïve Bayes* di Rumah Sakit Tingkat III Baladhika Husada Jember.

1.3.2 Tujuan Khusus

- a) Melakukan tahapan *preprocessing* data sesuai prosedur Algoritma *Naïve Bayes*.

- b) Mengidentifikasi karakteristik penyakit kanker payudara (C50) yang digunakan sebagai variabel untuk mengklasifikasikan penyakit kanker payudara di Rumah Sakit Tingkat III Baladhika Husada Jember.
- c) Mengklasifikasikan penyakit kanker payudara (C50) menggunakan Algoritma *Naïve Bayes* di Rumah Sakit Tingkat III Baladhika Husada Jember.
- d) Melakukan evaluasi *Confusion Matrix* dari klasifikasi penyakit kanker payudara (C50) menggunakan metode Algoritma *Naïve Bayes* di Rumah Sakit Tingkat III Baladhika Husada Jember.
- e) Menganalisis hasil klasifikasi Algoritma *Naïve Bayes* penyakit kanker payudara (C50) di Rumah Sakit Tingkat III Baladhika Husada Jember.

1.4 Manfaat Penelitian

1.4.1 Bagi Rumah Sakit

Sebagai bahan pertimbangan dan informasi mengenai klasifikasi penyakit kanker payudara serta mendukung akurasi diagnosis dan pengelolaan rekam medis di Rumah Sakit Tingkat III Baladhika Husada Jember.

1.4.2 Bagi Politeknik Negeri Jember

Menambah keilmuan dan bahan referensi terkait klasifikasi penyakit kanker payudara menggunakan Algoritma *Naïve Bayes* di Politeknik Negeri Jember khususnya jurusan kesehatan Program Studi Manajemen Informasi Kesehatan.

1.4.3 Bagi Peneliti

Meningkatkan kompetensi peneliti sebagai PMIK di bidang aplikasi statistik kesehatan, epidemiologi dasar, dan biomedik. Menambah pengetahuan dan pemahaman peneliti mengenai penyakit kanker payudara, serta meningkatkan kemampuan peneliti dalam pengelolaan dan analisis data rekam medis melalui penerapan Algoritma *Naïve Bayes*, khususnya dalam bidang kesehatan.