

BAB 1. PENDAHULUAN

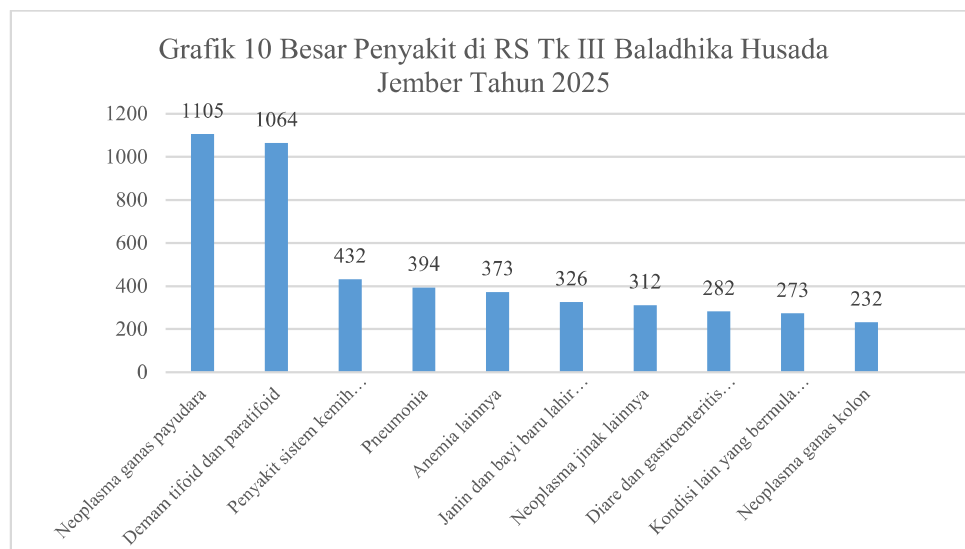
1.1 Latar Belakang

Kanker adalah penyakit ganas yang terjadi pada jaringan tubuh dan ditandai dengan pertumbuhan sel abnormal yang tidak normal dan tidak terkendali. (Agustin, 2020). Jenis kanker sangat bervariasi bergantung pada lokasi di mana pertumbuhannya dimulai, karena kanker dapat berkembang hampir di setiap bagian tubuh manusia. Kanker merupakan salah satu penyakit tidak menular yang dapat menyerang berbagai organ tubuh serta memiliki risiko tinggi menyebabkan kematian. Pada tahun 2022 tercatat bahwa kanker menyebabkan kematian sekitar 9,7 juta orang atau 18,7% dari seluruh kematian di seluruh dunia per tahun. Kanker merupakan penyebab kematian ketiga terbesar di Indonesia. Berdasarkan data *Global Cancer Observatory* pada tahun 2022 lebih dari 408.661 kasus baru dan hampir 242.099 kematian di Indonesia, dengan jumlah kematian tertinggi diakibatkan oleh kanker payudara, kanker leher rahim, kanker paru, dan kanker kolorektal (Kemenkes RI, 2024).

Kanker payudara adalah penyakit keganasan yang terjadi pada jaringan payudara, yang dapat berkembang dari sel epitel pada saluran susu (duktus) maupun dari lobulus payudara (Kemenkes RI, 2018). Kanker payudara termasuk salah satu jenis kanker yang paling banyak ditemukan di Indonesia dan merupakan penyakit yang sering terjadi pada perempuan. Pada tahun 2022 kanker payudara merupakan jenis kanker dengan angka kejadian tertinggi pada perempuan, yaitu sebesar 41,8 kasus per 100.000 penduduk. Sementara itu, jumlah kematian tertinggi juga disebabkan oleh kanker payudara, dengan angka kematian sebesar 14,4 kasus per 100.000 penduduk (Kemenkes RI, 2024).

Berdasarkan Profil Kesehatan Kabupaten Jember Tahun 2024, pemeriksaan deteksi dini kanker payudara terhadap 269.286 perempuan usia 30–50 tahun menunjukkan adanya temuan tumor atau benjolan pada payudara sebanyak 47 orang. Temuan tersebut memerlukan pemeriksaan lanjutan guna memastikan diagnosis dan menentukan penanganan yang tepat. Salah satu rumah sakit yang berperan dalam pelayanan rujukan serta penanganan lanjutan kasus kanker di wilayah Jember adalah Rumah Sakit Tingkat III Baladhika Husada Jember.

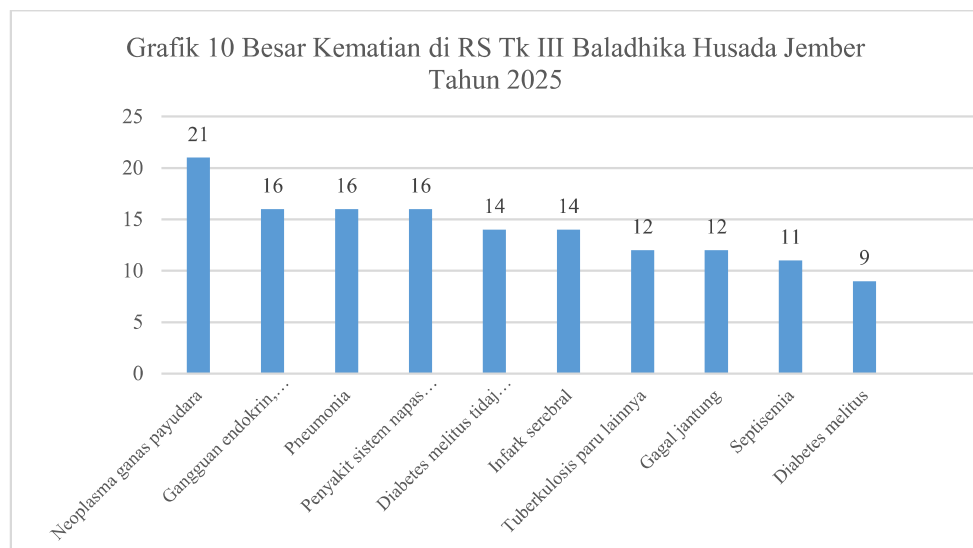
Rumah Sakit Tingkat III Baladhika Husada Jember merupakan Rumah Sakit Umum kelas C yang memberikan pelayanan kesehatan kepada masyarakat umum, peserta BPJS Kesehatan, serta anggota dan keluarga TNI/POLRI. Rumah sakit ini menyelenggarakan berbagai layanan kesehatan, meliputi pelayanan Rawat Jalan, Rawat Inap, serta Instalasi Gawat Darurat yang beroperasi selama 24 jam. Salah satu layanan unggulan yang tersedia adalah pelayanan kemoterapi, yang ditujukan bagi pasien kanker yang memerlukan terapi lanjutan. Ketersediaan layanan tersebut menjadikan Rumah Sakit Tingkat III Baladhika Husada Jember sebagai rumah sakit rujukan bagi pasien kanker. Berdasarkan hasil studi pendahuluan, kanker payudara tercatat sebagai salah satu dari sepuluh penyakit dengan jumlah kasus tertinggi di rumah sakit tersebut. Berikut merupakan data sepuluh besar penyakit di Rumah Sakit Tingkat III Baladhika Husada Jember tahun 2025.



Gambar 1.1 10 Besar Penyakit di RS Tk III Baladhika Husada Jember Tahun 2025

Gambar 1.1 menunjukkan grafik sepuluh besar penyakit yang diperoleh dari laporan rekapitulasi morbiditas pasien di Rumah Sakit Tingkat III Baladhika Husada Jember Tahun 2025. Pada grafik diatas dapat disimpulkan bahwa penyakit kanker payudara berada di peringkat pertama dengan jumlah kasus tertinggi, yaitu sebanyak 1105 kasus dan seluruh pasien yang terdiagnosis merupakan perempuan. Berdasarkan data tiga tahun terakhir tahun 2023-2025, kasus kanker payudara masih menempati peringkat pertama dibandingkan penyakit lainnya. Kondisi ini

menunjukkan bahwa kanker payudara masih menjadi masalah kesehatan yang perlu mendapat perhatian. Selain itu, kanker payudara juga berpotensi menyebabkan kematian. Hal ini dapat diketahui dari data mortalitas yang menunjukkan bahwa kanker payudara termasuk dalam sepuluh besar penyebab kematian. Berikut merupakan sepuluh besar kematian di Rumah Sakit Tingkat III Baladhika Husada Jember tahun 2025.



Gambar 1.2 10 Besar Kematian di RS Tk III Baladhika Husada Jember Tahun 2025

Berdasarkan grafik tersebut dapat diketahui bahwa kanker payudara merupakan salah satu penyebab kematian tertinggi dengan jumlah kasus sebanyak 21 kasus. Hal ini menunjukkan bahwa kanker payudara tidak hanya memiliki angka morbiditas yang tinggi, tetapi juga memiliki angka mortalitas yang tinggi. Kanker payudara ditandai oleh adanya berbagai faktor risiko dan gejala. Beberapa faktor risiko yang paling umum untuk kanker payudara adalah jenis kelamin wanita, usia di atas 50 tahun, riwayat keluarga dan penyakit payudara sebelumnya, menstruasi dini, riwayat reproduksi, hormon, obesitas, konsumsi alkohol, radiasi dinding dada, dan faktor lingkungan adalah (Kemenkes RI, 2018). Adanya faktor-faktor risiko tersebut dapat disertai dengan munculnya gejala klinis. Gejala kanker payudara umumnya meliputi benjolan pada payudara, perubahan ukuran atau bentuk payudara, perubahan pada puting payudara, keluarnya cairan yang tidak normal dari

puting, perubahan pada kulit payudara, serta benjolan di area ketiak (Rizka et al., 2022).

Wanita yang mengalami obesitas lebih banyak ditemukan pada kanker payudara stadium lanjut. Hal ini karena jaringan lemak menghasilkan enzim aromatase yang meningkatkan kadar hormon estrogen, sehingga berisiko kanker payudara (Farlikhatun, 2018). Wanita yang berusia > 50 tahun cenderung mengalami kanker payudara pada stadium lanjut, sementara wanita berusia < 50 tahun lebih sering ditemukan pada stadium awal (Soviyati et al., 2020). Tanda dan gejala kanker payudara juga bervariasi sesuai tingkat keparahan penyakit. Pada tahap awal, umumnya ditandai dengan benjolan atau penebalan pada payudara. Seiring perkembangan penyakit, gejala menjadi lebih kompleks seperti perubahan kulit payudara, perubahan puting, keluarnya cairan, hingga ulserasi. Pada kondisi metastasis, gejala dapat meluas ke bagian tubuh lain seperti nyeri pada bahu, pinggang, punggung bawah atau panggul (Rahmi & Andika, 2022).

Perbedaan karakteristik faktor risiko dan gejala tersebut memiliki keterkaitan dengan stadium penyakit yang diderita, sehingga stadium kanker payudara dapat digunakan untuk menggambarkan tingkat perkembangan dan keparahan penyakit. Oleh karena itu, diperlukan pemanfaatan data pasien secara optimal untuk mengklasifikasikan stadium penyakit berdasarkan karakteristik yang dimiliki pasien. Data rekam medis yang tersimpan di rumah sakit mengandung informasi tentang faktor risiko, dan gejala klinis yang dapat diolah untuk proses klasifikasi stadium kanker payudara. Hasil klasifikasi ini dapat mempermudah dalam mengelompokkan karakteristik pasien berdasarkan stadium, dan juga membantu tenaga medis dalam menentukan penanganan yang tepat, karena setiap stadium memiliki pendekatan terapi yang berbeda. Salah satu pendekatan yang dapat digunakan untuk mengolah data rekam medis tersebut adalah melalui *data mining*.

Data mining merupakan proses pengolahan data untuk menemukan pola atau informasi yang bernilai dengan menggunakan metode atau teknik tertentu (Yuli Mardi, 2019). Salah satu metode *data mining* adalah klasifikasi, yaitu teknik pengelompokan data berdasarkan karakteristik yang serupa untuk menentukan atau memprediksi kelas pada data yang belum memiliki label. (Amna et al., 2023). Salah

satu algoritma yang dapat digunakan dalam proses klasifikasi tersebut adalah *K-Nearest Neighbor (KNN)*. *K-Nearest Neighbor (KNN)* adalah metode klasifikasi yang mengelompokkan data berdasarkan kedekatan dengan data latih terdekat (Amna et al., 2023). Metode *K-Nearest Neighbor (KNN)* memiliki konsep yang sederhana, namun mampu menghasilkan tingkat akurasi yang cukup baik dalam proses prediksi (Sakti & Daulay, 2024). Variabel yang digunakan terdiri dari numerik maupun kategorik, yang merepresentasikan faktor risiko, gejala, dan karakteristik tumor. Kombinasi variabel tersebut memungkinkan KNN untuk menilai kemiripan antar pasien, sehingga pasien baru dapat diklasifikasikan ke dalam stadium kanker payudara yang paling sesuai berdasarkan pola pasien-pasien sebelumnya.

Menurut penelitian (Rizky et al., 2025) yang berjudul “Analisis Perbandingan Algoritma KNN dan SVM untuk Prediksi Risiko Kesehatan Ibu Hamil”, algoritma *K-Nearest Neighbor (KNN)* menunjukkan performa yang lebih unggul dengan nilai akurasi sebesar 81%, lebih tinggi dibandingkan algoritma *Support Vector Machine (SVM)* yang memiliki akurasi sebesar 75,50%. Pada penelitian (Kurniadi et al., 2021) yang berjudul “Perbandingan Akurasi Algoritma *K-Nearest Neighbor* dan *Logistic Regression* untuk Klasifikasi Penyakit Diabetes” diperoleh hasil bahwa algoritma *K-Nearest Neighbor (KNN)* juga menghasilkan nilai akurasi sebesar 85,06%, lebih tinggi dibandingkan dengan algoritma *Logistic Regression* yang memiliki nilai akurasi sebesar 77,92%. Hasil tersebut menunjukkan bahwa algoritma KNN memiliki performa yang lebih baik dalam melakukan klasifikasi penyakit diabetes.

Berdasarkan uraian latar belakang tersebut, peneliti tertarik untuk melakukan penelitian yang berjudul “Analisis Klasifikasi Penyakit Kanker Payudara (C50) Menggunakan Metode *K-Nearest Neighbor (KNN)* di Rumah Sakit Tingkat III Baladhika Husada Jember“. Tools yang digunakan dalam proses klasifikasi yaitu *Google Colaboratory*. *Google Colaboratory* adalah *platform* komputasi berbasis *cloud* yang memfasilitasi pengembangan model *machine learning* menggunakan *Python*. Platform ini mempermudah pengelolaan *library*, pemrosesan data, dan visualisasi hasil, tanpa memerlukan konfigurasi perangkat keras yang kompleks

(Wiguna et al., 2025). Penelitian ini menggunakan variabel usia, riwayat keluarga, riwayat payudara sebelumnya, IMT, nyeri payudara, benjolan pada payudara, benjolan pada ketiak, perubahan bentuk dan ukuran payudara, keluarnya cairan dari puting, perubahan pada puting, kelainan pada kulit payudara, status lokalis, dan ukuran tumor payudara. Kinerja metode *K-Nearest Neighbor* (KNN) pada penelitian ini selanjutnya dievaluasi menggunakan *Confusion Matrix* untuk memperoleh nilai *accuracy*, *precision*, dan *recall* guna mengetahui tingkat ketepatan metode yang digunakan. Hasil penelitian diharapkan dapat menjadi referensi bagi penelitian selanjutnya dalam mengembangkan sistem klasifikasi kanker payudara.

1.2 Rumusan Masalah

Berdasarkan latar belakang dapat dirumuskan permasalahan “Bagaimana Analisis Klasifikasi Penyakit Kanker Payudara (C50) Menggunakan Metode *K-Nearest Neighbor* (KNN) di Rumah Sakit Tingkat III Baladhika Husada Jember”.

1.3 Tujuan Penelitian

1.3.1 Tujuan Umum

Melakukan klasifikasi penyakit kanker payudara menggunakan algoritma *K-Nearest Neighbor* (KNN) di Rumah Sakit Tingkat III Baladhika Husada Jember.

1.3.2 Tujuan Khusus

- a. Melakukan *preprocessing data* penyakit kanker payudara di Rumah Sakit Tingkat III Baladhika Husada Jember.
- b. Mengklasifikasikan penyakit kanker payudara menggunakan *K-Nearest Neighbor* (KNN) dengan *Google Colab*.
- c. Menganalisis hasil klasifikasi penyakit kanker payudara menggunakan *Confusion Matrix* untuk mendapatkan nilai *accuracy*, *precision*, dan *recall* dari metode *K-Nearest Neighbor* (KNN).
- d. Menganalisis karakteristik penyakit kanker payudara di Rumah Sakit Tingkat III Baladhika Husada Jember
- e. Mengimplementasikan algoritma *K-Nearest Neighbor* (KNN) ke dalam sistem sederhana berbasis *Google Colab* untuk melakukan prediksi stadium kanker payudara.

1.4 Manfaat Penelitian

1.4.1 Bagi Peneliti

Penelitian ini bermanfaat untuk menambah pengetahuan serta mengimplementasikan ilmu yang diperoleh selama perkuliahan, khususnya dalam pengolahan data kesehatan menggunakan metode *data mining*, serta meningkatkan kompetensi PMIK dalam manajemen data dan informasi kesehatan melalui pemanfaatan data rekam medis untuk menghasilkan informasi yang mendukung pelayanan kesehatan.

1.4.2 Bagi Politeknik Negeri Jember

Sebagai referensi dalam pengembangan penelitian di bidang rekam medis dan informasi kesehatan, khususnya terkait penerapan metode *data mining* menggunakan algoritma *K-Nearest Neighbor* (KNN) dalam pengolahan data kesehatan untuk klasifikasi kanker payudara.

1.4.3 Bagi Rumah Sakit

Sebagai bahan masukan dan informasi dalam pengolahan dan analisis data mengenai klasifikasi kanker payudara di Rumah Sakit Tingkat III Baladhika Husada Jember melalui pemanfaatan data rekam medis menggunakan metode *data mining* dengan algoritma *K-Nearest Neighbor* (KNN).