

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kanker Payudara menempati urutan Pertama dengan jumlah kasus kanker terbanyak di Indonesia, dan menjadi salah satu penyumbang kematian pertama akibat kanker payudara (kemkes.go.id, 2022). Menurut data *statistic Globocan* Tahun 2020, tercatat jumlah kasus baru kanker payudara mencapai 68.858 (16,6%) dari total 396.914 di Indonesia. Kemudian, untuk jumlah kematiannya mencapai lebih dari 22 ribu jiwa kasus. Tingginya angka kematian karena kanker payudara di Indonesia perlu dicermati dengan tindakan pencegahan dan deteksi dini. Kanker payudara pada umumnya dibedakan menjadi dua jenis yaitu kanker jinak dan kanker ganas. Kanker jinak adalah kanker yang kondisinya masih dalam tahap awal sehingga jenis kanker ini masih bisa ditangani oleh tenaga medis dan bahkan sangat besar potensi penyembuhannya. Sedangkan kanker ganas adalah jenis kanker yang sangat berbahaya apabila tidak ditangani dengan baik akan berpotensi pada kematian. (Farahdiba and Nugroho, 2016)

Teknologi machine learning memungkinkan untuk dimanfaatkan dalam melakukan deteksi penyakit kanker payudara secara dini. Terdapat 9 variabel kuantitatif, dan variabel dependen biner, yang menunjukkan ada atau tidaknya kanker payudara, berdasarkan kumpulan dataset *Coimbra UCI Machine Learning Repository* tahun 2018 yang diperoleh dari kaggle seperti usia, BMI, Glukosa, Insulin, HOMA, leptin, Adiponectin, Resistin, MCP-1, dan 2 Classification yaitu 2

2(pasien kanker payudara), 1(kesehatan yang terkontrol). Pendeteksian tersebut berguna untuk menentukan perencanaan pengobatan yang akan diambil.

Naïve Bayes adalah salah satu algoritma *classification* yang sederhana tetapi cukup efektif dalam melakukan prediksi terhadap *dataset* baru. Algoritma ini digunakan dalam kasus deteksi kanker payudara karena mampu mengklasifikasikan data tersebut. Berdasarkan pokok permasalahan diatas, maka dilakukan penelitian penerapan metode *naïve bayes* pada klasifikasi kanker payudara untuk mengidentifikasi dan mengklasifikasi penyakit kanker payudara.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat diambil rumusan masalah yang akan dikembangkan dalam penelitian, yaitu: Bagaimana penerapan metode klasifikasi *Naïve Bayes* untuk mengidentifikasi penyakit kanker payudara? dan menganalisis seberapa Akurasi dari Model *Naive Bayes*.

1.3 Ruang Lingkup

Agar dapat mencapai sasaran dan tujuan yang diharapkan maka diberikan ruang lingkup sebagai berikut:

1. Metode yang digunakan yaitu *Naïve Bayes*.
2. Bahasa pemrograman yang digunakan adalah *python*.
3. Data yang digunakan adalah dataset dari *kaggle* [kaggle.com/datasets/yasserhessein/breast-cancer-coimbra-data-set/data](https://www.kaggle.com/yasserhessein/breast-cancer-coimbra-data-set/data) yang terdiri dari 116 data 80% untuk *data training*, dan 20% untuk *data testing*.
4. *Variabel* yang digunakan untuk klasifikasi ini adalah 9 atribut yaitu : Usia,

BMI, Glucose, Insulin, HOMA, Leptin, Adiponectin, Resistin, dan MCP.1 dan 2 kelas yaitu : 2 (pasien kanker payudara) dan 1 (kesehatan yang terkontrol).

5. Penelitian ini berfokus pada akurasi penerapan klasifikasi *Naïve Bayes* dalam mengidentifikasi penyakit kanker payudara.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengidentifikasi dan mengklasifikasi penyakit kanker payudara berdasarkan *dataset* yang diperoleh dari *kangle* dengan menggunakan metode *Naïve Bayes*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah: Membantu Pembaca dalam meningkatkan pemahaman tentang penyakit kanker Payudara, dan membantu atau memudahkan petugas rumah sakit dalam mengklasifikasi penyakit kanker payudara secara dini, agar bisa melakukan pengobatan secepatnya menggunakan metode *naive bayes*.

1.6 Sistematika Penulisan

Sistem penulisan skripsi yang digunakan dalam penelitian ini sebagai berikut:

Bab I Pendahuluan

Bab 1 membahas tentang latar belakang masalah, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

Bab II Dasar Teori

Bab II membahas tentang tinjauan pustaka dan dasar teori yang mendasari penelitian dalam melakukan klasifikasi pada dataset kanker payudara atau tidak kanker payudara menggunakan metode *Naive Bayes*.

Bab III Metode Penelitian

Bab III membahas tentang kebutuhan yang digunakan serta gambaran umum mengenai metode dan prosedur yang digunakan dalam klasifikasi pada penyakit kanker payudara menggunakan metode *naive bayes*.

Bab IV Implementasi dan Pembahasan

Bab IV membahas tentang implementasi dan pembahasan hasil dari klasifikasi pada penyakit kanker payudara menggunakan Metode *naive bayes*.

Bab V Kesimpulan

Bab V membahas tentang kesimpulan yang diperoleh dari hasil klasifikasi pengujian yang dilakukan dalam penelitian, dan saran yang diberikan agar penelitian selanjutnya bisa mencapai hasil yang lebih baik.