

# **BAB I**

## **PENDAHULUAN**

### **1.1 LATAR BELAKANG MASALAH**

Kemajuan teknologi yang berkembang sangat pesat mengakibatkan pola kehidupan masyarakat menjadi serba instan, hal ini mengakibatkan perilaku masyarakat yang tradisional beralih kepada kehidupan yang serba cepat. Mulai dari komunikasi, transportasi hingga makanan semua berubah menjadi instan. Banyak orang yang tidak sadar akan dampak dari perubahan tersebut, pola hidup yang tidak sehat, asap kendaraan bermotor yang mempengaruhi kesehatan pernafasan, radiasi yang ditimbulkan oleh alat-alat komunikasi juga makanan siap saji yang kurang di perhatikan gizi dan nutrisinya. Perubahan tersebut membawa dampak yang signifikan bagi kehidupan masyarakat, dampak yang paling signifikan dari mengkonsumsi makanan instan dan pola hidup tidak sehat adalah munculnya berbagai penyakit salah satunya adalah penyakit kanker.

Salah satu penyebab kematian utama di seluruh dunia adalah penyakit kanker. Sekitar 8,2 juta orang meninggal disebabkan kanker tahun 2012 [1]. Berdasarkan data Globocan Tahun 2018, Indonesia menduduki peringkat 8 di Asia Tenggara dan 23 di Asia dalam hal jumlah penyakit kanker. Kanker payudara merupakan salah satu jenis kanker yang banyak diderita oleh kaum wanita di Indonesia dengan jumlah kejadian 42,1 per 100.000 orang dan jumlah kematian rata-rata 17 per 100.000 orang [2].

Faktor risiko yang dapat mempengaruhi kanker payudara diantaranya ialah usia, genetik, gender, riwayat kanker payudara, etnis, jaringan payudara yang padat, hormon, paparan radiasi, dan pola hidup tidak sehat. Penyebaran dan perkembangan kanker payudara tergolong sangat cepat. Kanker payudara dapat dibedakan menjadi dua macam yaitu jinak (*benign*) dan ganas (*malignant*). Tumor yang termasuk kategori jinak, dapat terdeteksi namun kondisi dan perkembangannya belum bersifat kanker. Tumor yang sudah memasuki kategori ganas telah bersifat kanker dimana waktu penyebarannya cepat dan dapat merusak jaringan serta organ disekitarnya [3].

Penanganan kanker yang terlambat akan membahayakan keselamatan penderita kanker payudara sehingga diperlukan adanya tindakan lanjut untuk mengatasi penyakit kanker payudara. Ada beberapa penanganan kanker payudara tergantung pada stadium klinik penyakitnya, salah satunya ialah melakukan operasi pengangkatan tumor payudara. Operasi kanker payudara sangat berpengaruh terhadap fisik dan mental kesehatan seorang wanita seperti keterkejutan mental, takut bahkan sampai depresi berat. Menurut organisasi kesehatan dunia, 1 juta wanita didiagnosis menderita kanker payudara setiap tahun dan separuh dari mereka akhirnya meninggal dunia [4].

Hingga saat ini, kanker payudara menempati urutan pertama terkait jumlah kanker terbanyak di Indonesia serta menjadi salah satu penyumbang kematian terbanyak pertama akibat kanker. Kementerian Kesehatan (Kemenkes) menyatakan, angka kanker payudara di Indonesia mencapai 42,1 orang per 100 ribu penduduk. Rata-rata kematian akibat kanker ini mencapai 17 orang per 100

ribu penduduk [5]. Dengan banyaknya data kematian akibat kanker payudara, perlu dilakukan pengolahan terhadap data pasien kanker payudara agar rekam medis tersebut berguna bagi masa yang akan datang serta menjadi informasi dan pengetahuan baru terkait kanker payudara. Mengacu pada permasalahan tersebut maka diperlukan sebuah metode yang digunakan untuk melakukan pengolahan data status pasien yaitu dengan menerapkan teknik *data mining*.

Penerapan *data mining* dinilai tepat karena memudahkan dalam melakukan penggalian informasi terhadap data status pasien pasca operasi, sehingga data yang telah diperoleh dapat diolah untuk mengetahui informasi-informasi yang tersembunyi dari data yang ada. Salah satu algoritma dalam *data mining* yang dapat digunakan untuk pengklasifikasian adalah *Naïve Bayes*. Terdapat beberapa penelitian sebelumnya yang menerapkan metode *data mining* yang terkait tentang penyakit antara lain : penelitian [6] dengan tujuan untuk memprediksi keberadaan kanker payudara pada tubuh seseorang menggunakan metode *Naïve Bayes* memperoleh hasil sebesar 80%, dimana dari 35 data uji 28 diantaranya diklasifikasikan dengan tepat.

Berikutnya penelitian [7] dengan tujuan untuk mengklasifikasikan dan diagnosa penderita penyakit Kanker Payudara. Untuk perhitungan manual sendiri, hasil akurasi untuk *data training* sebesar 81.00% sama dengan hasil pengklasifikasian menggunakan *RapidMiner* sementara hasil akurasi data testing sebesar 62.50% sama persis dengan hasil perhitungan menggunakan *RapidMiner*.

Berdasarkan uraian permasalahan dari penelitian-penelitian sebelumnya, maka penulis tertarik untuk mengklasifikasikan status pasien pasca operasi kanker

payudara. Untuk klasifikasi tersebut dibutuhkan sebuah algoritma yang tepat. Penulis menggunakan algoritma *Naïve Bayes* dikarenakan algoritma *Naïve Bayes* termasuk salah satu algoritma yang sering digunakan dan dinilai berpotensi baik dalam mengklasifikasikan data. Penerapan algoritma *Naïve Bayes* dapat melakukan perhitungan secara efisien, cepat, algoritma yang sederhana dan berakurasi tinggi [8].

Berdasarkan permasalahan di atas, penulis tertarik melakukan penelitian pengklasifikasian status pasien pasca operasi kanker payudara dengan mengangkat judul **“Penerapan Algoritma *Naïve Bayes* Untuk Klasifikasi Status Pasien Pasca Operasi Kanker Payudara”**.

## **1.2 RUMUSAN MASALAH**

Berdasarkan latar belakang masalah yang telah dipaparkan, maka dapat dirumuskan masalah yang akan diteliti yaitu : “Bagaimana menerapkan algoritma *Naïve Bayes* untuk mengklasifikasi dan menganalisis status pasien pasca operasi kanker payudara?”

## **1.3 BATASAN MASALAH**

Batasan masalah digunakan dalam sebuah penelitian dengan tujuan agar pembahasan lebih terarah dan sesuai dengan tujuan yang akan dicapai. Maka penulis memberikan batasan masalah pada penelitian kali ini sebagai berikut:

1. Objek penelitian bersumber dari [www.kaggle.com](http://www.kaggle.com) dengan data sebanyak 334 *record*.
2. Data yang digunakan merupakan dataset *public* dari pasien kanker payudara yang telah menjalani operasi pengangkatan kanker payudara.

3. Dalam melakukan pengolahan data, *software* yang digunakan adalah aplikasi *Microsoft Excel*, *RapidMiner*, *Google Collaboratory* dan *Anaconda Navigator*.
4. Metode yang digunakan adalah metode *Naive Bayes*.
5. Atribut yang digunakan yaitu *Age*, *Gender*, *Tumor Stage*, *Histology*, *ER Status*, *PR Status*, *HER2 Status*, *Surgery Type*, *Duration* dan *Patient Status*.

## **1.4 TUJUAN DAN MANFAAT PENELITIAN**

### **1.4.1 Tujuan Penelitian**

Tujuan yang ingin dicapai dari penelitian ini adalah :

1. Menghasilkan data klasifikasi status pasien pasca operasi kanker payudara.
2. Menerapkan keakurasian klasifikasi menggunakan algoritma *Naive Bayes* pada status pasien pasca operasi kanker payudara.

### **1.4.2 Manfaat Penelitian**

Diharapkan dengan adanya tugas akhir ini dapat memberikan manfaat sebagai berikut :

1. Bagi Masyarakat dapat memberikan informasi untuk mengetahui rata-rata status pasien pasca menjalani operasi kanker payudara.
2. Bagi Tenaga Medis dapat mengolah data pasien pasca operasi sehingga memberikan informasi tentang prediksi keselamatan pasien, serta tenaga medis dapat melakukan penanganan lebih lanjut secara cepat dan efisien.

3. Bagi Penulis yaitu dapat menambah wawasan tentang metode *Data mining* untuk klasifikasi data pada data pasien pasca operasi kanker payudara.
4. Memberikan informasi mengenai status pasien setelah melakukan operasi kanker payudara.
5. Dapat dijadikan acuan untuk penelitian berikutnya.
6. Dapat memberikan info penting yang tersembunyi dalam suatu data menggunakan metode *Naïve Bayes*.

## **1.5 SISTEMATIKA PENULISAN**

Laporan ini terdiri dari 5 bab dan setiap bab terdapat sub bab. Untuk mendapatkan gambaran yang lebih jelas mengenai susunan penulisan penelitian ini, penulis akan menguraikan secara singkat hal-hal yang akan dibahas dalam laporan penelitian ini.

### **BAB I : PENDAHULUAN**

Bab ini berisi pendahuluan yang membahas latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

### **BAB II : LANDASAN TEORI**

Bab ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang dianalisis. Teori-teori yang digunakan antara lain mengenai klasifikasi, *Naive Bayes*, *data mining*, alat bantu aplikasi seperti *Microsoft Excel*, *RapidMiner* serta tentang bahasa pemrograman *Python*.

### **BAB III : METODOLOGI PENELITIAN**

Pada bab ini menjelaskan mengenai kerangka kerja penelitian, metode ataupun teknik pengumpulan data serta alat-alat dan pembahasan pendukung selama melakukan penelitian.

### **BAB IV : ANALISIS**

Pada bab ini menjelaskan tentang analisis dan seleksi data pasien dengan melakukan *preprocessing* data menjadi *preprocessed* data. Transformasi data tersebut kemudian diolah dengan teknik *data mining* untuk mengetahui keakurasian data menggunakan metode klasifikasi algoritma *Naive Bayes*. Hasil dari analisis yang telah dilakukan di interpretasikan menjadi sebuah informasi mengenai evaluasi status pasien pasca operasi kanker payudara.

### **BAB V : KESIMPULAN**

Bab ini merupakan penutup yang berisi kesimpulan yang merupakan jawaban dari masalah-masalah yang diperoleh selama penelitian dan saran-saran yang ditunjukkan kepada pihak terkait dengan hasil penelitian.