

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kanker paru merupakan segala jenis penyakit keganasan yang terjadi di paru, melingkupi keganasan yang asalnya dari paru itu sendiri atau dikenal sebagai kanker paru primer. Berdasarkan definisi secara klinik, pengertian dari kanker paru primer sendiri merupakan tumor ganas yang sumbernya dari *epitel bronkus (karsinoma bronkus/bronchogenic carcinoma)*. Data WHO menunjukkan bahwa, kanker yang paling banyak diidap kaum laki-laki di Indonesia merupakan kanker paru, serta kanker ini merupakan yang terbanyak kelima yang terjadi pada kaum perempuan setelah kanker payudara, serviks-uteri, kolorektal, dan ovarium. Data hasil pengecekan dari laboratorium Patologi Anatomi RSUP Persahabatan menunjukkan bahwa dari semua jenis kanker yang berhasil didiagnosa, kanker paru menyumbang lebih dari 50 persen kasus. Departemen Pulmonologi dan Kedokteran Respirasi FKUI-RSUP Persahabatan juga menunjukkan data bahwa dalam 10 tahun belakangan ini telah terjadi peningkatan sebesar lebih dari 5 kali lipat angka kasus baru kanker paru dimana penderita kanker paru didominasi berada pada stadium lanjut. Selain itu, terdapat lebih dari 1000 kasus baru kanker paru per tahun yang melakukan pengobatan di RSUP Persahabatan (Hudoyo et al., 2017).

Penyebab paling utama terjadinya kanker paru ialah kebiasaan merokok oleh karena itu penderita kanker ini didominasi perokok aktif dan kaum laki-laki, tetapi tidak hanya perokok aktif, kanker paru juga berisiko tinggi terjadi pada perokok pasif. Selain kebiasaan merokok, penyebab kanker paru lainnya adalah karena kerap terpapar polusi udara, pasien atau keluarga pasien mempunyai riwayat penyakit paru atau kanker lainnya, ataupun pasien berdomisili di wilayah yang terkontaminasi zat kimia berbahaya. Kanker paru tidak mempunyai gejala klinis khusus yang hanya dimiliki jenis kanker ini namun pada umumnya batuk, sesak nafas, nyeri dada, suara serak, nafsu makan menurun, demam hilang timbul, sakit kepala, nyeri di tulang atau parese, dan pembengkakan atau ditemukannya benjolan di leher yang terjadi dalam kurun waktu yang lama atau tidak adanya tanda-tanda kesembuhan setelah dilakukan prosedur pengobatan biasa dapat dicurigai sebagai

kanker paru. Terdapat beberapa tahapan dalam mendiagnosis kanker paru mulai dari anamnesis, pemeriksaan fisik, pemeriksaan patologi anatomi, pemeriksaan laboratorium, pemeriksaan pencitraan, pemeriksaan khusus, dan pemeriksaan lainnya (Hudoyo et al., 2017).

Tahapan-tahapan dalam mendeteksi kanker paru tersebut bukanlah hal yang mudah untuk dilakukan sehingga hasil diagnosis tidak bisa didapat secara cepat. Tenaga medis perlu pengamatan dan pertimbangan yang mendalam supaya diagnosis yang dihasilkan tepat dan akurat. Rumitnya proses dalam mendiagnosis kanker paru memunculkan gagasan untuk menggunakan algoritma klasifikasi dalam melakukan deteksi dini kanker paru, dimana dalam penelitian akan diuji apakah algoritma klasifikasi yang dipilih dapat mengklasifikasikan suatu kasus termasuk kedalam kanker paru atau tidak secara benar dilihat dari hasil performa algoritma klasifikasi tersebut.

Berdasarkan informasi yang telah dijabarkan diatas, peneliti berniat menerapkan kombinasi algoritma klasifikasi *Adaboost* dan *Random Forest* dalam mendeteksi dini kanker paru. Kedua algoritma tersebut dipilih karena algoritma sangat cocok dalam masalah klasifikasi pada *machine learning* atau *data mining*, dimana kedua algoritma tersebut memiliki hasil prediksi dengan akurasi tinggi dan *error* yang rendah. Hal ini dibuktikan dengan berbagai penelitian terdahulu menggunakan algoritma ini. Pertama, penelitian yang dilakukan oleh Thongkam et al. (2008) dimana dilakukan penelitian dengan judul *AdaBoost Algorithm with Random Forests for Predicting Breast Cancer Survivability*. Pada penelitian ini diterapkan kombinasi antara algoritma *Adaboost* dan *Random Forest* dalam membentuk model prediksi kelangsungan hidup kanker payudara. Hasil akhir menunjukkan bahwa dibandingkan dengan beberapa algoritma lainnya, kombinasi *Adaboost* dan *Random Forest* berhasil memperoleh akurasi tertinggi yaitu sebesar 88.60%. Kedua, penelitian yang dilakukan oleh Byna et al. (2020) melakukan penelitian dengan judul Penerapan Metode *Adaboost* Untuk Mengoptimasi Prediksi

Penyakit Stroke Dengan Algoritma *Naïve Bayes*. Dalam penelitian ini algoritma *Adaboost* digabung dengan *Naïve Bayes* untuk mengoptimasi akurasi yang didapat. Terbukti pada awalnya akurasi yang dihasilkan algoritma *Naïve Bayes* saja sebesar 97,60% sedangkan setelah dioptimasi dengan *Adaboost* akurasi naik menjadi 98,10%.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka rumusan masalah dalam penelitian ini adalah bagaimana performa yang dihasilkan dari kombinasi algoritma klasifikasi *Adaboost* dan *Random Forest* dalam mendeteksi dini kanker paru dibandingkan dengan pengguna algoritma klasifikasi *Random Forest* tanpa penggabungan dengan *Adaboost*.

## 1.3 Tujuan Penelitian

Tujuan dari penelitian berdasarkan latar belakang yang telah disampaikan sebelumnya adalah untuk mengetahui tingkat performa yang didapat dari pengimplementasian kombinasi algoritma klasifikasi *Adaboost* dan *Random Forest* dalam membuat sistem deteksi dini terhadap kanker paru dibandingkan dengan pengguna algoritma klasifikasi *Random Forest* tanpa penggabungan dengan *Adaboost*.

## 1.4 Ruang Lingkup

Supaya pembahasan pada penelitian ini tidak meluas dan tidak melenceng maka ruang lingkup dari penelitian ditentukan sebagai berikut.

- a. Penelitian dilakukan dengan fokus utama untuk mengetahui seberapa baik dan seberapa besar performa yang dihasilkan dari gabungan algoritma *Adaboost* dan *Random Forest* dalam mendeteksi kanker paru.

- b. Deteksi dini kanker paru didasarkan pada 15 variabel sebagai indikator yang terdiri dari jenis kelamin, umur, gejala umum kanker paru, serta kebiasaan yang dapat menyebabkan kanker paru.
- c. Deteksi yang dilakukan hanya menggunakan dua algoritma klasifikasi yaitu *Adaboost* dan *Random Forest*.
- d. Data yang digunakan merupakan data sekunder, dimana data tersebut telah digunakan oleh peneliti sebelumnya yang diperoleh dari *website kaggle.com*.

### **1.5 Luaran yang Diharapkan**

Adapun luaran yang diharapkan dari penelitian ini yaitu dihasilkan performa yang sangat baik pada penggunaan kombinasi algoritma *Adaboost* dan *Random Forest* dalam mendeteksi kasus kanker paru.

### **1.6 Manfaat Penelitian**

Penelitian yang dilakukan ini diharapkan dapat memberi manfaat. Adapun beberapa manfaat yang diharapkan didapat dari penelitian ini adalah sebagai berikut.

- a. Dapat mengetahui seberapa baik performa yang didapat dari penerapan kombinasi algoritma klasifikasi *Adaboost* dan *Random Forest* dalam mendeteksi dini kanker paru.
- b. Dapat menambah ilmu baru terkait penggunaan metode klasifikasi yang dikombinasikan serta penerapannya pada data dibidang medis.
- c. Dapat dijadikan pembandingan dengan hasil klasifikasi menggunakan metode lain dengan dataset yang serupa.

## 1.7 Sistematika Penulisan

Sistematika penulisan penelitian ini dibagi kedalam bab dan sub bab. Berikut sistematika penulisannya:

### **BAB 1           PENDAHULUAN**

Bab ini berisikan penjelasan terkait latar belakang dilakukan penelitian, rumusan masalah dalam penelitian, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian, luaran yang diharapkan dari penelitian yang dilakukan, dan sistematika penulisan penelitian.

### **BAB 2           LANDASAN TEORI**

Bab ini berisikan uraian terkait teori-teori dasar yang digunakan sebagai acuan dan pendukung dalam penelitian ini. Tidak hanya berisikan kumpulan teori saja, bab 2 juga berisikan penelitian terdahulu mengenai metode yang digunakan penulis dalam hal ini penelitian terdahulu terkait *Adaboost* dan *Random Forest*.

### **BAB 3           METODOLOGI PENELITIAN**

Bab ini berisikan tahapan-tahapan yang harus dilakukan dalam penelitian sehingga tujuan dari penelitian yang telah ditetapkan dapat terwujud dan terlaksana dengan baik dan benar.

### **BAB 4           HASIL DAN PEMBAHASAN**

Bab ini berisikan penjelasan terkait implementasi dari tahapan-tahapan penelitian yang dirumuskan pada bab 3 mengenai pengumpulan data, representasi data, hingga proses pembuatan sistem deteksi serta hasil dari sistem yang telah selesai dibuat.

### **BAB 5           PENUTUP**

Bab ini berisikan kesimpulan dari hasil pemecahan masalah yang berhasil didapat berdasarkan pembahasan hasil di bab 4, serta saran-

saran berdasarkan penelitian yang telah dilakukan agar penelitian kedepannya lebih berkembang lagi.

## **DAFTAR PUSTAKA**

## **RIWAYAT HIDUP**