

BAB I PENDAHULUAN

1.1 Latar Belakang

Asma adalah penyakit yang menyerang saluran pernapasan dan dapat dialami oleh siapa saja, namun jumlah penderita terbanyak adalah anak-anak. Berdasarkan data dari KEMENKES (2008), sekitar 100 hingga 150 juta orang di seluruh dunia menderita asma, dan jumlah ini diperkirakan akan meningkat sebanyak 18.000 kasus setiap tahunnya (Tiwi & Maryuni, 2021). Asma tidak menular, namun dapat menjadi salah satu penyebab utama kematian di seluruh dunia. Penyakit ini mengganggu saluran pernapasan dan tidak hanya menyerang orang dewasa, tetapi juga dapat mempengaruhi anak-anak, yang mana sistem kekebalan tubuh mereka masih belum cukup kuat untuk melawan penyakit ini (Putri *et al.*, 2022). Menurut data dari Organisasi Kesehatan Dunia (*WHO*) pada tahun 2005, jumlah penderita asma di seluruh dunia diperkirakan akan meningkat sebanyak 180.000 orang setiap tahunnya. Setiap tahun, sekitar 250.000 orang meninggal akibat asma. Di Indonesia, berdasarkan hasil survei Riset Kesehatan Dasar (RISKESDAS) pada tahun 2007, prevalensi asma mencapai 3,5% dari total jumlah penduduk. Asma termasuk dalam sepuluh besar penyebab kesakitan dan kematian di Indonesia. Data dari Survei Kesehatan Rumah Tangga (SKRT) di berbagai provinsi menunjukkan bahwa ada 9 provinsi dengan prevalensi asma tertinggi (Wijaya & Toyib, 2018).

Penyebab terjadinya asma dapat dibagi menjadi dua faktor utama. Faktor pertama adalah faktor keturunan atau genetik, di mana individu dapat mewarisi kecenderungan untuk menderita asma melalui gen yang diturunkan. Beberapa gen yang terlibat dalam proses terjadinya asma telah diidentifikasi, dan sejumlah kromosom diketahui berperan dalam meningkatkan risiko terjadinya asma. Namun, faktor keturunan ini saja tidak cukup untuk menyebabkan timbulnya asma. Faktor kedua adalah faktor pencetus, yang terbagi menjadi dua kategori: faktor pencetus internal (dalam tubuh) dan eksternal (dari lingkungan). Faktor pencetus internal mencakup infeksi saluran pernapasan, stres, aktivitas fisik berat, olahraga, serta emosi yang berlebihan. Sementara itu, faktor pencetus eksternal meliputi debu, paparan asap rokok (baik aktif maupun pasif), konsumsi makanan dan minuman

tertentu, obat-obatan, zat pewarna, bau-bauan, bahan kimia, polusi udara, serta perubahan suhu atau cuaca.

Peningkatan kasus asma disebabkan karena adanya keterbatasan kemampuan kerja dan banyaknya masyarakat cenderung berinisiatif mengkonsumsi obat yang dijual bebas (Muhdi *et al.*, 2017). Akibatnya, penyakit yang seharusnya tidak terlalu berbahaya menjadi lebih berat dan bahkan dapat menyebabkan munculnya penyakit baru yang sulit untuk diatasi. Oleh karena itu, diperlukan klasifikasi penderita asma untuk menilai risiko terpapar penyakit tersebut berdasarkan gejala yang dialami. Pada data pasien dikelompokkan menjadi 2, yaitu asma dan tidak asma. Keterangan pasien asma menunjukkan bahwa pasien didiagnosis mengalami penyakit asma, sedangkan pasien tidak asma menunjukkan bahwa pasien didiagnosis tidak mengalami penyakit asma.

Klasifikasi dalam *machine learning* adalah metode pembelajaran yang digunakan untuk mengelompokkan nilai dari sejumlah atribut dalam menggambarkan kelas, dengan tujuan untuk memprediksi kelas dari objek yang kelasnya sudah diketahui. Pemanfaatan pembelajaran mesin semakin berkembang dalam bidang perawatan kesehatan untuk mendukung para praktisi kesehatan dalam proses pengambilan keputusan. Dalam pengelolaan asma, *machine learning* terbukti efektif dalam melaksanakan tugas-tugas yang jelas, seperti diagnosis, prediksi, pengobatan, dan manajemen (Darsha Jayamini *et al.*, 2024).

Algoritma *Random Forest* adalah salah satu algoritma klasifikasi yang sangat efektif untuk digunakan dalam mendiagnosis kesalahan. Algoritma ini bekerja dengan cepat, bahkan pada dataset yang besar, meskipun memiliki kelemahan dalam menghasilkan *overfitting* pada data yang tidak seimbang (Sebayang *et al.*, 2023). Berdasarkan hasil uji yang dilakukan sebelumnya dalam mengukur akurasi klasifikasi menggunakan *Confusion Matrix*, metode *Random Forest* menunjukkan hasil terbaik dengan akurasi mencapai 70,60%, sedangkan *Naïve Bayes* hanya memperoleh akurasi 55,80%. Dengan demikian, *Random Forest* memiliki kinerja yang lebih baik dalam menghasilkan akurasi yang lebih tinggi dalam klasifikasi penyakit liver (Lubis *et al.*, 2022).

Dalam penelitian serupa pada tahun 2019, penelitian yang menerapkan metode *Random Forest* untuk memprediksi harga ponsel, dengan tingkat *accuracy*, *precision*, *recall* dan *F1-score* mencapai nilai 81% (Siburian & Mulyana, 2018). Metode *Random Forest* juga diterapkan dalam penelitian deteksi dan prediksi cerdas penyakit paru-paru yang dilakukan oleh Deny Kurniawan dkk. Dalam penelitian ini algoritma *Random Forest* diimplementasikan dengan 10 pohon keputusan dan enam atribut pada setiap pembagian datanya dengan hasil pengujian menunjukkan nilai AUC sebesar 0,993 yang mengindikasikan bahwa tingkat akurasi yang sangat tinggi (Kurniawan *et al.*, 2024).

Random Forest merupakan merupakan algoritma model yang sudah teruji dapat menghasilkan nilai akurasi yang tinggi (Suryanegara *et al.*, 2021). Untuk membantu para profesional medis mengurangi kemungkinan kesalahan dalam proses diagnosis, dibutuhkan sebuah teknologi yang dapat mengklasifikasikan dan menganalisis potensi terjadinya penyakit asma. Dengan adanya sistem klasifikasi penyakit asma yang efektif, para dokter dapat lebih mudah dan cepat dalam memperoleh informasi yang akurat terkait kondisi pasien. Hal ini tentunya akan sangat mendukung dalam pengambilan keputusan medis yang lebih tepat, sehingga diagnosis yang diberikan menjadi lebih valid dan dapat meningkatkan kualitas penanganan penyakit tersebut. Maka dari itu, penelitian ini dilakukan untuk membantu tenaga medis mengurangi kesalahan diagnosis dan mengukur kinerja metode *Random Forest* pada dataset penyakit asma yang diperoleh dari RSUD Sumberrejo, Bojonegoro dan. Penghitungan kinerja yang dilakukan mencakup *accuracy*, *precision*, *recall*, dan *f1-score*. Dengan menggunakan metode tersebut, diharapkan dapat mendukung para dokter di RSUD Sumberrejo, Bojonegoro dalam melakukan diagnosis asma dengan lebih akurat dan cepat.

1.2 Rumusan Masalah

Dengan merujuk dari informasi pada bagian latar belakang, dideskripsikan rumusan masalah dalam penelitian ini adalah sebagai berikut:

1.2.1 Permasalahan

Penyakit asma sudah menjadi salah satu masalah kesehatan yang serius dengan kasus yang semakin meningkat tiap tahunnya menjadikan penyakit ini

adalah salah satu penyebab kematian di Indonesia yang dapat menyerang anak-anak maupun dewasa (Wijaya & Toyib, 2018). Peningkatan kasus asma disebabkan karena adanya keterbatasan informasi yang masyarakat peroleh mengenai gejala penyakit ini dan kurangnya kemampuan kerja yang mengakibatkan banyaknya masyarakat cenderung berinisiatif mengkonsumsi obat yang dijual bebas. Maka dari itu, perlu adanya model klasifikasi yang dapat membantu tenaga medis dalam proses mendiagnosis pasien penyakit asma dengan cepat dan akurat.

1.2.2 Metode Usulan

Dari permasalahan itu, maka didapat solusi dengan melakukan klasifikasi penyakit asma menggunakan metode *Random Forest* dengan perhitungan *K-Fold Cross Validation*, sehingga mampu menghasilkan model klasifikasi yang akurat

1.2.3 Pertanyaan Penelitian

Berapa nilai *accuracy*, *precision*, *recall* dan *f1-score* pada klasifikasi penyakit asma dengan menggunakan metode *Random Forest*?

1.3 Batasan Masalah

Agar tercapainya tujuan dalam pengerjaan penelitian ini, maka terdapat batasan masalah sebagai berikut:

1. Dataset yang digunakan merupakan dataset yang diperoleh dari RSUD Sumberrejo, Bojonegoro.
2. *Dataset* yang digunakan berjumlah data yang digunakan sebanyak 800 data, 9 fitur dengan 2 kelas yaitu asma dan tidak asma.
3. Metode yang digunakan adalah *Random Forest*.
4. Penggunaan proses *drop missing value* pada tahap *preprocessing*.
5. Penggunaan proses normalisasi data dengan metode *min-max scaler*.
6. Kinerja model evaluasi dengan menggunakan *Confusion Matrix* meliputi *accuracy*, *precision*, *recall* dan *f1-score*.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk membangun model klasifikasi penyakit asma dan membantu dalam mendiagnosis penyakit secara cepat dan akurat.
2. Untuk mengetahui serta mengevaluasi kinerja model *Random Forest* dalam klasifikasi penyakit asma.

1.4.2 Manfaat Penelitian

Dari tujuan di atas, penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Sistem klasifikasi dapat membantu tenaga medis untuk lebih cepat dalam proses mendiagnosis pasien penyakit asma.
2. Mengetahui kinerja algoritma model *Random Forest* dalam mengklasifikasi penyakit asma.

1.5 Sistematika Laporan

Sistematika laporan skripsi merupakan tahapan laporan skripsi dari bab 1, 2, 3, 4, dan 5 yang akan dijelaskan sebagai berikut:

a. BAB I PENDAHULUAN

Bab I memuat penjelasan mengenai latar belakang penyakit asma, guna mendapatkan informasi mengenai masalah yang dialami. Selain itu, bab ini juga menjelaskan beberapa metode yang dilakukan pada penelitian sebelumnya dan menjabarkan juga terkait alasan dipilihnya metode *Random Forest*. Lalu terdapat juga tujuan dan manfaat penelitian, batasan penelitian dan sistematika dalam laporan ini.

b. BAB II KAJIAN PUSTAKA

Bab II berisi mengenai metode dan teori yang dijadikan landasan dalam penelitian ini terkait tahap *preprocessing*, perhitungan *K-Fold Cross Validation*, implementasi metode *Random Forest* dan *Confusion Matrix* serta penelitian terkait dari metode yang sudah digunakan pada penelitian terdahulu.

Diharapkan hal ini membantu dalam mempermudah dasar teori yang digunakan dalam penelitian ini.

c. BAB III METODE USULAN

Bab III memuat penjesalan mengenai dataset yang akan digunakan, arsitektur sistem yang akan diterapkan, serta tahapan-tahapan penerapan algoritma *Random Forest* dan juga skenario uji coba. Tahapan penelitian diantaranya terkait arsitektur sistem serta dataset yang digunakan

d. BAB IV HASIL DAN PEMBAHASAN

Bab IV berisi mengenai pembahasan atau penjelasan yang lebih rinci mengenai proses implementasi model klasifikasi penyakit asma menggunakan metode *Random Forest*. Hasil implementasi terdiri dari lingkungan uji coba dan hasil uji coba pengujian *dataset*. Pengujian ini dilakukan untuk mengetahui hasil akurasi yang didapat dan menilai seberapa tepat model dalam menghasilkan output.

e. BAB V PENUTUP

Bab V berisi kesimpulan dari penelitian yang disesuaikan dengan pertanyaan penelitian serta saran berdasarkan penelitian yang sudah dilakukan untuk penelitian selanjutnya.