

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ketersediaan air bersih yang memadai dan bermutu tinggi memiliki kaitan erat dengan kelangsungan hidup dan derajat kesehatan seluruh makhluk hidup. Secara nasional, hak untuk memperoleh akses air bersih merupakan hak dasar setiap manusia yang harus terpenuhi, mengingat kekurangan air bersih dapat menimbulkan dampak serius, termasuk memicu berbagai penyakit bahkan kematian [1].

Kenyataan di lapangan menunjukkan bahwa masyarakat cenderung memanfaatkan air dari berbagai sumber alternatif seperti air daur ulang, sumur bor, maupun sumur gali dibandingkan air perpipaan, tidak hanya sebagai air minum tetapi juga untuk kebutuhan sehari-hari seperti memasak dan mencuci. Di Sumatera Utara, sekitar 40% penduduk mengandalkan air galon isi ulang resmi, sementara 30% lainnya bergantung pada air daur ulang tidak resmi sebagai sumber utama [2]. Kondisi ini umumnya terjadi akibat terbatasnya infrastruktur air bersih di lingkungan sekitar, serta pertimbangan efisiensi dalam pemenuhan kebutuhan air sehari-hari [3]. Standar kelayakan air minum mensyaratkan nilai pH berada pada kisaran 6,5 hingga 8,5; sementara air yang mengandung zat berbahaya seperti arsenik atau bakteri patogen harus dihindari. Meski sudah ada berbagai metode pemeriksaan, sebagian besar masih membutuhkan waktu dan sumber daya yang signifikan, sehingga dibutuhkan pendekatan deteksi yang lebih cepat dan efisien guna memastikan keamanan air yang dikonsumsi [4].

Penetapan kualitas air secara tepat dan efisien sangat krusial untuk memastikan bahwa sumber air yang dipakai oleh masyarakat benar-benar aman dikonsumsi [5]. Pendekatan ini tidak hanya mampu meningkatkan akurasi pendeteksian, tetapi juga memungkinkan respons yang lebih cepat terhadap permasalahan kualitas air. Pemanfaatan kecerdasan buatan dalam analisis kualitas air menawarkan kemampuan untuk memproses dan menganalisis data berskala besar. Salah satu algoritma yang telah terbukti efektif adalah Random Forest, yang dapat digunakan untuk mengklasifikasikan kualitas air berdasarkan berbagai parameter terukur, sehingga memberikan penilaian yang lebih mendalam mengenai kelayakan konsumsi air [6].

Sejumlah penelitian terdahulu telah mengeksplorasi penggunaan algoritma Random Forest dalam identifikasi dan pengelompokan kualitas air. Salah satunya menggunakan Random Forest untuk klasifikasi kelayakan air minum, dan menghasilkan tingkat akurasi tertinggi sebesar 78% [7]. Penelitian tersebut juga mengungkapkan bahwa performa algoritma dapat ditingkatkan melalui penyesuaian parameter atau hyperparameter tuning [8].

Penelitian lainnya yang memadukan algoritma Random Forest dengan teknik optimasi seperti Grid Search dan Random Search difokuskan pada peningkatan akurasi model klasifikasi keamanan air minum berdasarkan data yang ada. Faktor kualitas air yang dianggap berisiko dapat mendukung pembentukan dataset yang lebih komprehensif [9]. Berdasarkan berbagai temuan tersebut, riset ini bertujuan untuk menyempurnakan penerapan algoritma Random Forest dalam kategorisasi kualitas air bersih, dengan harapan dapat memberikan kontribusi ilmiah dalam pengembangan dan validasi akurasi algoritma ini [10].

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini mencakup:

1. Bagaimana merancang dan membangun sistem berbasis IoT untuk memonitor parameter kualitas air minum, khususnya nilai pH dan kadar TDS secara real-time?
2. Parameter input apa yang paling berpengaruh dan relevan dalam proses klasifikasi tingkat keamanan konsumsi air minum?
3. Bagaimana implementasi algoritma Random Forest dalam pengolahan data sensor untuk mengklasifikasikan tingkat keamanan air minum secara presisi?
4. Bagaimana perbandingan unjuk kerja sistem deteksi kualitas air minum berbasis IoT dan Random Forest dengan metode konvensional dalam mengevaluasi kelayakan konsumsi air?

1.3 Batasan Masalah

Agar penelitian ini tetap terarah dan tidak melebar, ditetapkan batasan-batasan sebagai berikut:

1. Ruang lingkup penelitian dibatasi pada klasifikasi data kelayakan air jernih untuk kategori dapat atau tidak dapat dikonsumsi.
2. Dataset yang digunakan adalah "Water Quality and Water Potability" yang bersumber dari platform Kaggle, dengan total 3.276 data.
3. Dataset memuat 9 atribut, meliputi: pH, hardness, solids, chloramines, sulfate, conductivity, organic carbon, trihalomethanes, dan turbidity.
4. Pada implementasi perangkat keras, sensor yang digunakan hanya mencakup pengukuran pH dan turbidity.
5. Algoritma yang diterapkan adalah Random Forest yang telah dioptimalkan.
6. Implementasi algoritma menggunakan bahasa pemrograman Python dengan Google Colab sebagai lingkungan pengembangan.

1.4 Tujuan Penelitian

Merujuk pada rumusan masalah yang telah dipaparkan, tujuan penelitian ini adalah:

1. Menerapkan algoritma Random Forest sebagai solusi klasifikasi kualitas air jernih.
2. Mengimplementasikan optimasi Random Forest menggunakan Grid Search untuk penggolongan kualitas air jernih.
3. Menganalisis perbedaan hasil optimasi algoritma Random Forest dengan pendekatan Random Search dan Grid Search pada klasifikasi kualitas air.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini antara lain:

1. Memberikan kontribusi pada pengembangan ilmu pengetahuan, khususnya dalam validasi empiris keberhasilan optimasi algoritma Random Forest dengan Random Search dan Grid Search untuk klasifikasi kualitas air.
2. Memperluas wawasan peneliti dan meningkatkan kompetensi riset, serta menjadi acuan bagi penelitian serupa di masa mendatang.