

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara berkembang dengan jumlah penduduk yang besar. Sehingga hal ini menyebabkan permintaan air bersih di Indonesia meningkat. Air merupakan sebuah kebutuhan kedua yang sangat diperlukan oleh manusia setelah udara. Di dalam air memiliki unsur mineral yang dibutuhkan untuk perkembangan fisik manusia. Sungai merupakan sumber air permukaan yang dimanfaatkan oleh manusia untuk kebutuhan hidup. Kualitas sungai dapat berubah tergantung dari aktivitas manusia. Pencemaran sungai dapat disebabkan oleh lingkungan sekitar, baik sungai tersebut maupun perilaku masyarakat pengguna sungai [1]. Pencemaran sungai sering terjadi disebabkan pola kehidupan dalam memanfaatkan alam seperti pencemaran yang berasal dari limbah *industry*, hal ini dapat mengakibatkan turunnya kualitas air sungai sehingga menyebabkan air tidak berguna sesuai manfaatnya [2]. Kualitas Air secara umum menunjukkan kondisi air yang memenuhi persyaratan sanitasi. Berdasarkan keputusan Menteri Kesehatan Nomor 416 menyatakan bahwa air bersih harus memenuhi persyaratan yang telah ditetapkan yaitu keadaan fisik, kimia dan biologi [3] .

PDAM Barito Kuala merupakan IPA (Instalasi Pengolahan Air) yang menjadi salah satu sarana dan prasarana pengolahan air untuk menyediakan kebutuhan masyarakat akan sumber air bersih yang sesuai dengan baku mutu yang ditentukan [4]. Kabupaten Barito Kuala berbatasan dengan dua sungai besar sehingga mempengaruhi tata air yang ada. Selama musim hujan, sungai Barito dapat membanjiri kawasan yang luas sehingga menyebabkan banjir di permukaan terus menerus dan air pasang juga mempengaruhi saluran air. Oleh karena itu diperlukan upaya untuk menjaga kualitas air melalui pemantauan dan pengendalian air secara berkala sesuai dengan keputusan pemerintah No. 82 Tahun 2001 [5].

Kualitas air dapat ditentukan dengan menguji air menggunakan parameter Fisika, Kimia dan Biologi seperti Ph air, Tds, kekeruhan, suhu dan lain-lain. Pengujian kualitas air secara umum menggunakan cara manual dengan metode *STORET* dan *WQI (Water Quality Index)*. Metode ini dilakukan dengan menghitung satu-persatu dari parameter yang digunakan. Proses ini memakan waktu yang cukup lama untuk menerima dan menghasilkan hasil tes kesimpulan kualitas air [6].

Sehingga dengan permasalahan yang ada peneliti mengusulkan penggunaan metode *data mining* untuk melakukan klasifikasi kualitas air.

Salah satu teknologi yang terkenal yaitu *machine learning* yang merupakan suatu cabang ilmu pengetahuan kecerdasan buatan (*Artificial Intelligent/AI*) dalam meningkatkan kemampuan dan kecanggihan *machine* dalam mengidentifikasi informasi serta memecahkan masalah pada sistem yang baru dikenali dari sebuah data tanpa memprogram ulang. Beberapa metode *machine learning* dapat digunakan untuk mengklasifikasi, salah satunya *Modified K-Nearest Neighbor*. Terdapat penelitian yang menyatakan bahwa algoritma *Support Vector Machine*, *K-Nearest Neighbor*, *Neural Network*, *Logistic Regression* dan *Modified K-Nearest Neighbor* merupakan algoritma klasifikasi yang terkenal digunakan dalam bidang sosial [7]. *K-Nearest Neighbor* adalah salah satu algoritma klasifikasi untuk menemukan kasus baru dengan mencocokkan bobot fitur yang ada untuk menghitung kedekatan kasus lama dan baru, sedangkan *Modified K-Nearest Neighbor* adalah algoritma yang dikembangkan dari algoritma *K-Nearest Neighbor*. Algoritma *Modified K-Nearest Neighbor* menambahkan proses klasifikasi dengan menghitung nilai validitas dan bobot. Pada algoritma *Modified K-Nearest Neighbor* nilai validitas dan bobot mencapai akurasi yang lebih baik ketika nilai validitas dan bobot tinggi [8].

Menurut penelitian sebelumnya yang dilangsungkan tahun 2019. Pada penelitian ini memakai 167 data yang berasal dari PDAM Malang. Penelitian ini melakukan klasifikasi air dengan metode SVM menghasilkan akurasi sebesar 78,70%, sedangkan menggunakan metode lain yaitu *Naïve Bayes* memberikan akurasi sebesar 85, 78% [2]. Penelitian yang dilakukan M Ikhasan Perdana Putra dkk (2019) menggunakan metode *M K-NN* untuk klasifikasi penyakit kanker payudara. Penelitian ini menggunakan data dari “*UCI- Machine Learning Repository Breast Cancer Wisconsin*”. Data tersebut diklasifikasikan menggunakan dua kelas yaitu kanker payudara jinak dan kanker payudara ganas. Hasil dari pengujian menunjukkan bahwa nilai *k* mempengaruhi akurasi. Penelitian ini menghasilkan nilai akurasi tinggi sebesar 97.61% [9]. Pada penelitian lainnya [10], penelitian tersebut berfokus menganalisis pengaruh metode normalisasi *Min-max* dan *StandardScaler* terhadap berbagai algoritma penambangan data seperti algoritma KNN, *Naïve Bayes*, ANN, dan SVM kernel RBF dan kernel liner. Pada

pengujian ini algoritma *Naïve Bayes* dengan normalisasi *Min-max* maupun *StandartScaler* menghasilkan performa yang stabil dengan akurasi 70,67%. Sedangkan KNN, Performanya masih cukup stabil dibandingkan SVM dan ANN dengan nilai akurasi sebesar 77,33% pada normalisasi *Min-max* dan 76% pada normalisasi *StandartScaler*. Namun pada SVM performa terbaik didapatkan jika menggunakan kernel RBF dengan normalisasi *Min-max* menghasilkan akurasi sebesar 82,67%.

Berdasarkan penelitian terdahulu yang telah diuraikan di atas, dapat disimpulkan bahwa algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor* masing-masing memiliki keunggulan. Ada penelitian yang telah membandingkan algoritma tersebut dengan algoritma *Naïve Bayes* dan *Decision Tree* [11], hasilnya menunjukkan bahwa nilai akurasi *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor* lebih unggul. Oleh karena itu peneliti berinisiatif untuk melakukan perbandingan antara algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor* untuk klasifikasi kualitas air. Penggunaan algoritma *K-Nearest Neighbor* ini dikarenakan algoritma tersebut memiliki performa yang baik. Sedangkan untuk penggunaan algoritma *Modified K-Nearest Neighbor* dikarenakan algoritma tersebut merupakan algoritma pengembangan dari *K-Nearest Neighbor* dimana pada klasifikasinya menambahkan dua perhitungan yaitu nilai validitas dan bobot. Belum ada penelitian sebelumnya yang membandingkan algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor* dalam mengklasifikasikan kualitas air. Maka dari itu diharapkan dengan adanya penelitian ini dapat memberikan informasi dan dapat menjadi acuan dan referensi serta pengembangan ilmu pengetahuan untuk penelitian selanjutnya.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Berdasarkan uraian latar belakang permasalahan dari penelitian ini adalah dalam penentuan kualitas air masih menggunakan metode tradisional manual dan memakan waktu yang cukup lama untuk mendapatkan hasil klasifikasi kualitas air.

1.2.2 Metode Usulan

Metode yang disarankan untuk memecahkan masalah di atas adalah dengan menggunakan algoritma klasifikasi *Modified K-Nearest Neighbor* dan algoritma *K-Nearest Neighbor*

1.2.3 Pertanyaan Penelitian

Berdasarkan masalah dan metode yang disajikan, pertanyaan penelitian sebagai berikut :

1. Bagaimana perbandingan hasil tingkat akurasi menggunakan algoritma *Modified K-Nearest Neighbor* dan algoritma *K-Nearest Neighbor* dengan menggunakan metode normalisasi *Min-max* dan *StandartScaler* untuk klasifikasi kualitas air?

1.3 Batasan Masalah

Batasan masalah dari penelitian ini yaitu:

1. Data yang dipakai pada penelitian ini hanya berfokus pada *dataset* yang diperoleh dari PDAM Barito Kuala, Kalimantan Selatan.
2. Data yang dipakai sebanyak 210 data, yang terdiri dari 7 atribut yaitu (bau air, tds air, kekeruhan air, suhu air, ph air, sisa klor air dan CO2 air).
3. Kelas data yaitu memenuhi syarat (ms) dan tidak memenuhi syarat (tms).
4. Metode yang digunakan untuk penelitian ini yaitu *Modified K-Nearest Neighbor* dan *K-Nearest Neighbor*.

1.4 Tujuan Penelitian

Berdasarkan masalah di atas, tujuan dari penelitian ini adalah:

1. Menghitung dan membandingkan hasil tingkat akurasi antar algoritma *Modified K-Nearest Neighbor* dan *K-Nearest Neighbor* dengan normalisasi *Min-max* dan *StandartScaler* untuk klasifikasi kualitas air.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Membantu memudahkan pihak PDAM Barito Kuala dalam proses pengecekan klasifikasi Kualitas Air.

2. Bisa dijadikan referensi dalam mengembangkan suatu sistem klasifikasi menggunakan algoritma *Modified K-Nearest Neighbor* dan *K-Nearest Neighbor*.
3. Mengetahui hasil kinerja dari algoritma *Modified K-Nearest Neighbor* dan *K-Nearest Neighbor* dalam melakukan klasifikasi pada data kualitas air.

1.6 Sistematika Penulisan

Sistem penulisan usulan penelitian Tugas Akhir yakni:

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika skripsi.

BAB II KAJIAN PUSTAKA

Pada bab ini membahas teori yang dibutuhkan untuk penelitian, seperti algoritma *Modified K-Nearest Neighbor* dan *K-Nearest Neighbor* dalam klasifikasi kualitas air, serta beberapa penelitian sebelumnya.

BAB III METODE PENELITIAN

Bab ini menjelaskan perancangan sistem, dimulai dengan perangkat lunak dan perangkat keras yang diperlukan serta rancangan modelnya.

BAB IV HASIL DAN PEMBAHASAN

Bab ini mencakup lingkungan pengujian sistem, proses pembuatan sistem, yang terdiri dari *source code*, *interface*, dan hasil pengujian metode.

BAB V KESIMPULAN

Berisi kesimpulan dari semua penelitian yang telah dilakukan serta saran yang ingin disampaikan di penelitian selanjutnya untuk menghasilkan klasifikasi yang lebih baik.

DAFTAR PUSTAKA

Berisi semua referensi yang digunakan untuk penyusunan laporan skripsi.