

# Analisis Kualitas Air Sungai Berbasis Decision Tree untuk Penentuan Status Pencemaran

Ongky Helfian<sup>1\*</sup>, Angga Kurniawan<sup>2</sup>, Rifki Arya Rahardiansyah<sup>3</sup>, Bintang Anugrah S<sup>4</sup>, Rizal Yudha P<sup>5</sup>, Aprilisa Arum Sari<sup>6</sup>

<sup>1</sup>Teknik Informatika  
Universitas Duta Bangsa  
<sup>1\*</sup>220103210@mhs.udb.ac.id

<sup>2</sup>Teknik Informatika  
Universitas Duta Bangsa  
<sup>2</sup>220103205@mhs.udb.ac.id

<sup>3</sup>Teknik Informatika  
Universitas Duta Bangsa  
<sup>3</sup>220103201@mhs.udb.ac.id

<sup>4</sup>Teknik Informatika  
Universitas Duta Bangsa  
<sup>4</sup>220103294@mhs.udb.ac.id

<sup>5</sup>Teknik Informatika  
Universitas Duta Bangsa  
<sup>5</sup>220103212@mhs.udb.ac.id

<sup>6</sup>Fakultas Ilmu Komputer  
Universitas Duta Bangsa  
<sup>6</sup>aprilisa\_arumsari@udb.ac.id

**Abstrak**— Sungai memegang peranan penting dalam menunjang kehidupan masyarakat dan ekosistem, namun kualitas air sungai di Jakarta semakin menurun akibat pencemaran dari aktivitas industri dan domestik. Penelitian ini bertujuan untuk mengklasifikasikan status kualitas air dari 15 titik sungai utama di Jakarta serta melakukan segmentasi berdasarkan tingkat pencemaran. Parameter kualitas air yang dianalisis meliputi BOD, COD, pH, dan DO, yang diolah melalui pivot tabel, normalisasi, deteksi outlier, dan penetapan status baku mutu. Metode Decision Tree digunakan karena kemampuannya menghasilkan model yang mudah diinterpretasikan dan mendukung pengambilan keputusan. Hasil penelitian menunjukkan bahwa sekitar 62% sampel sungai berada pada kategori tercemar, terutama akibat tingginya BOD dan rendahnya DO. Dengan adanya segmentasi berdasarkan skor pencemaran, sungai dapat diprioritaskan untuk program penanganan dan rehabilitasi. Penelitian ini diharapkan menjadi dasar bagi perumusan kebijakan pengelolaan kualitas air sungai yang lebih efektif dan berkelanjutan di Jakarta.

**Kata kunci**— Kualitas air sungai, BOD, DO, Decision Tree, Segmentasi pencemaran.

**Abstract**— Rivers play an important role in supporting community life and ecosystems, but the quality of river water in Jakarta is decreasing due to pollution from industrial and domestic activities. This study aims to classify the water quality status of 15 main river points in Jakarta and segment them based on the level of pollution. The water quality parameters analyzed include BOD, COD, pH, and DO, which are processed through pivot tables, normalization, outlier detection, and determination of quality standard status. The Decision Tree method is used because of its ability to produce models that are easy to interpret and support decision making. The results showed that around 62% of river samples were in the polluted category, mainly due to high BOD and low DO. With segmentation based on pollution scores, rivers can be prioritized for handling and rehabilitation programs. This study is expected to be the basis for formulating more effective and sustainable river water quality management policies in Jakarta.

**Keywords**— river water quality, BOD, DO, Decision Tree, segmentation.

## I. PENDAHULUAN

Sungai memiliki peran penting dalam mendukung keberlangsungan kehidupan manusia dan ekosistem. Selain sebagai sumber air baku, sungai juga berfungsi sebagai habitat organisme perairan, sarana transportasi, serta pendukung kegiatan ekonomi dan rekreasi [1][2]. Namun, kualitas air sungai di berbagai daerah di Indonesia, termasuk Jakarta, terus mengalami penurunan yang signifikan [4][6][8].

Peningkatan aktivitas industri, pertumbuhan penduduk yang pesat, serta pengelolaan limbah

yang tidak optimal menjadi faktor utama yang memperburuk kondisi kualitas air sungai [5][7][9].

Berdasarkan laporan Kementerian Lingkungan Hidup dan Kehutanan (KLHK), sekitar 68% sungai di Indonesia berada dalam kategori tercemar, baik ringan maupun berat [10][12][13]. Di Jakarta sendiri, beban pencemaran sungai didominasi oleh limbah rumah tangga dan industri yang mengakibatkan tingginya kadar Biochemical Oxygen Demand (BOD) serta rendahnya Dissolved Oxygen (DO),

yang pada akhirnya mengancam keberlanjutan ekosistem perairan [11]. Kondisi ini tidak hanya berdampak pada lingkungan, tetapi juga dapat memengaruhi kesehatan masyarakat sekitar yang masih memanfaatkan air sungai untuk berbagai kebutuhan sehari-hari.

Penilaian kualitas air sungai umumnya dilakukan dengan memeriksa parameter fisika, kimia, dan biologi, di antaranya BOD, Chemical Oxygen Demand (COD), pH, dan DO. Namun, metode konvensional yang hanya berdasarkan pembacaan manual cenderung memakan waktu, memerlukan sumber daya yang besar, serta kurang efektif jika diaplikasikan pada skala wilayah yang luas. Oleh karena itu, diperlukan pendekatan berbasis teknologi dan data mining yang mampu mengklasifikasikan status kualitas air secara cepat, akurat, dan transparan [3][15]. Salah satu metode yang potensial adalah Decision Tree, yang memiliki keunggulan dalam menghasilkan aturan klasifikasi yang mudah diinterpretasikan. Metode ini juga memungkinkan segmentasi kualitas air berdasarkan tingkat pencemaran, sehingga mempermudah identifikasi prioritas sungai yang memerlukan tindakan rehabilitasi segera [12].

Dalam beberapa dekade terakhir, kualitas air sungai di Jakarta mengalami penurunan drastis akibat meningkatnya aktivitas domestik, industri, dan urbanisasi yang pesat. Pencemaran sungai tidak hanya memengaruhi ekosistem perairan, tetapi juga mengancam kesehatan masyarakat yang masih memanfaatkan air sungai untuk kebutuhan sehari-hari. Parameter kimia seperti BOD, COD, pH, dan DO sering kali menunjukkan hasil yang melebihi ambang batas baku mutu air, sehingga diperlukan metode pemantauan yang lebih efektif dan efisien. Metode konvensional berbasis pemeriksaan manual tidak lagi memadai untuk memantau banyak titik sungai dalam waktu singkat, serta kurang mendukung pengambilan keputusan yang cepat di lapangan.

Berdasarkan kondisi tersebut, perumusan masalah dalam penelitian ini adalah bagaimana menerapkan metode Decision Tree untuk mengklasifikasikan status kualitas air sungai di

Jakarta berdasarkan parameter BOD, COD, pH, dan DO, serta bagaimana melakukan segmentasi sungai berdasarkan skor pencemaran sehingga dapat ditentukan prioritas sungai mana yang memerlukan tindakan rehabilitasi atau pengendalian pencemaran secara lebih tepat sasaran.

Penelitian ini memiliki beberapa tujuan utama. Pertama, mengembangkan model klasifikasi status kualitas air sungai di Jakarta berdasarkan parameter BOD, COD, pH, dan DO dengan menggunakan metode Decision Tree, sehingga dapat mempermudah identifikasi kondisi sungai secara lebih cepat dan akurat.

Kedua, melakukan segmentasi sungai berdasarkan skor pencemaran untuk memetakan tingkat pencemaran pada masing-masing titik pemantauan, sehingga dapat diketahui prioritas penanganan atau rehabilitasi sungai yang mendesak.

Ketiga, penelitian ini bertujuan untuk menyusun peringkat sungai dari yang paling tercemar hingga terbersih, sebagai acuan dalam merumuskan strategi pengelolaan kualitas air dan kebijakan pengendalian pencemaran yang lebih efektif dan tepat sasaran. Hasil penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pemerintah daerah, lembaga pengelola lingkungan, serta masyarakat dalam upaya menjaga keberlanjutan fungsi ekologis sungai di Jakarta.

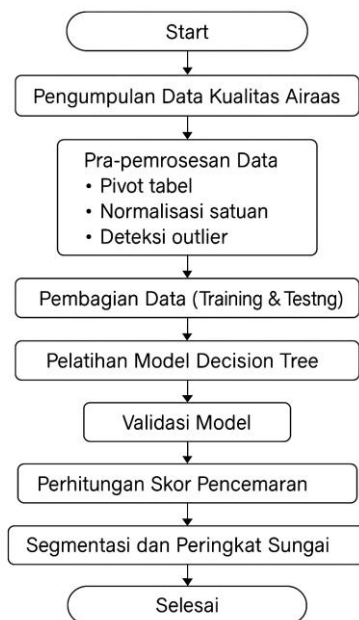
Berdasarkan latar belakang tersebut, tujuan penelitian ini adalah:

1. Mengembangkan model klasifikasi status kualitas air sungai di Jakarta berdasarkan parameter BOD, COD, pH, dan DO menggunakan metode Decision Tree.
2. Melakukan segmentasi sungai berdasarkan skor pencemaran untuk menentukan prioritas rehabilitasi.
3. Memberikan rekomendasi pengelolaan kualitas air sungai sebagai dasar perumusan kebijakan pengendalian pencemaran yang lebih efektif dan berkelanjutan.

## II. METODOLOGI PENELITIAN

Penelitian ini menggunakan data sekunder hasil pengukuran kualitas air sungai yang diperoleh dari Dinas Lingkungan Hidup Provinsi DKI Jakarta pada periode pengamatan tahun 2024. Parameter yang dianalisis meliputi Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), pH, dan Dissolved Oxygen (DO). Data dikumpulkan dari 15 titik pemantauan sungai utama yang tersebar di wilayah Jakarta. Selain itu, perangkat keras yang digunakan adalah komputer dengan spesifikasi minimum prosesor Intel i5, RAM 8 GB, dan perangkat lunak Python (versi 3.11) dengan pustaka pendukung seperti Pandas, Scikit-learn, dan Matplotlib.

Langkah-Langkah Eksperimen Penelitian ini dilakukan melalui beberapa tahap sistematis yang dijelaskan pada gambar 1 sebagai berikut:



Gambar 1: Tahapan

### 1. Pengumpulan Data

Data parameter BOD, COD, pH, dan DO diunduh dari basis data resmi Dinas Lingkungan Hidup. Data yang digunakan adalah data bulanan hasil pemantauan kualitas air sungai periode Januari–Desember 2024.

### 2. Pra-pemrosesan Data

Data diolah menjadi pivot tabel sehingga setiap parameter menjadi kolom tersendiri yang

memuat informasi titik sampling, nama sungai, dan alamat lokasi. Selanjutnya, dilakukan normalisasi satuan terutama pada parameter DO untuk memastikan keseragaman (dikonversi ke mg/L jika perlu). Nilai outlier diidentifikasi, khususnya jika  $DO > 20$  mg/L, dan data tersebut dihapus agar tidak memengaruhi hasil analisis.

### 3. Penentuan Status Kualitas Air

Penetapan status kualitas air dilakukan berdasarkan baku mutu yang ditetapkan KLHK, yaitu:

- $BOD > 3$  mg/L → Tercemar
- $COD > 25$  mg/L → Tercemar
- pH di luar rentang 6–9 → Tercemar
- $DO < 4$  mg/L → Tercemar

Jika salah satu parameter melebihi atau di luar batas tersebut, sampel dikategorikan sebagai "tercemar".

### 4. Pelatihan Model Klasifikasi

Data yang telah diberi label status (tercemar/tidak tercemar) dibagi menjadi data latih (80%) dan data uji (20%). Model Decision Tree Classifier diterapkan menggunakan pustaka Scikit-learn dengan parameter default. Model dilatih untuk mengenali pola keterkaitan antarparameter dengan status kualitas air.

### 5. Validasi Model

Evaluasi model dilakukan menggunakan data uji dengan metrik akurasi, precision, recall, dan F1-score. Analisis feature importance digunakan untuk mengidentifikasi parameter yang paling berpengaruh dalam penentuan status kualitas air.

### 6. Segmentasi dan Penilaian Skor Pencemaran

Skor pencemaran dihitung dengan menjumlahkan selisih nilai parameter yang melebihi baku mutu. Sungai kemudian diurutkan dari skor tertinggi (paling tercemar) hingga skor terendah (terbersih). Segmentasi dilakukan ke dalam beberapa kelompok: tercemar berat, tercemar sedang, dan tidak tercemar.

Analisis data dilakukan secara kuantitatif dengan bantuan Python. Visualisasi hasil analisis disajikan dalam bentuk grafik batang, diagram pohon keputusan, serta tabel ringkasan

status kualitas air di setiap titik pemantauan. Selain itu, analisis deskriptif statistik (rata-rata, minimum, maksimum) digunakan untuk mendukung interpretasi hasil. Hasil klasifikasi dibandingkan dengan status baku mutu sebagai validasi. Temuan utama dianalisis secara naratif dengan mempertimbangkan kondisi lingkungan di Jakarta dan dampak kebijakan pengendalian pencemaran.

Dengan pendekatan ini, penelitian diharapkan mampu menghasilkan informasi yang akurat, mudah dipahami, dan bermanfaat sebagai acuan dalam penyusunan kebijakan pengelolaan kualitas air sungai.

### III. HASIL DAN PEMBAHASAN

#### a. Pengumpulan Data

Data parameter kualitas air yang digunakan dalam penelitian ini terdiri dari BOD, COD, pH, dan DO yang diperoleh dari Dinas Lingkungan Hidup Provinsi DKI Jakarta. Data ini mencakup 15 titik pemantauan sungai utama di Jakarta selama periode Januari–Desember 2024. Setiap titik mewakili wilayah strategis dengan intensitas aktivitas domestik dan industri yang berbeda.

Tahap awal analisis dilakukan dengan melakukan deskriptif statistik terhadap parameter utama kualitas air, yaitu BOD, COD, pH, dan DO. Statistik ini memberikan gambaran distribusi dan kecenderungan nilai parameter sebelum dilakukan klasifikasi. Hasil statistik awal ditampilkan pada Tabel 1.

Tabel 1 : Statistik Awal Parameter Kualitas Air

| Parameter | Count | Mean     | Std Dev  | Min  | 25%   | 50%   | 75%    | Max      |
|-----------|-------|----------|----------|------|-------|-------|--------|----------|
| BOD       | 480   | 19.28    | 15.94    | 2.08 | 7.28  | 13.28 | 27.00  | 93.00    |
| COD       | 480   | 97.13    | 105.03   | 14.6 | 45.58 | 71.49 | 110.25 | 1500.00  |
| pH        | 480   | 2.37e+07 | 3.67e+08 | 6.02 | 7.43  | 7235  | 21057  | 5.71e+09 |
| DO        | 480   | 5193.23  | 11300.91 | 0.05 | 1.35  | 2.68  | 2062.5 | 63821.00 |

Nilai rata-rata pH dan DO terlihat sangat tinggi, menunjukkan adanya kesalahan satuan pada data mentah. Oleh karena itu, dilakukan normalisasi untuk mengembalikan nilai ke satuan standar. Setelah koreksi, nilai-nilai parameter dianalisis untuk menentukan status pencemaran. [1][15].

Data sementara ini juga mengindikasikan perlunya monitoring lanjutan dengan penambahan parameter BOD, COD, dan DO secara lebih merata pada seluruh titik sungai, agar gambaran status pencemaran dapat diperoleh secara lebih komprehensif. [12][14] yang menekankan pentingnya pengawasan berkala di sungai perkotaan. Selain itu, analisis segmentasi dan skor pencemaran yang akan dijelaskan pada bagian berikutnya akan semakin memperkuat pemetaan prioritas sungai untuk tindakan rehabilitasi dan pencegahan pencemaran berkelanjutan.

#### b. Pra-Pemrosesan Data

Tahap ini mencakup tiga proses utama:

Pembuatan Pivot Tabel: Data disusun agar setiap baris merepresentasikan satu titik pengamatan, dan setiap parameter (BOD, COD, pH, DO) menjadi kolom tersendiri.

Normalisasi Satuan: Terutama pada parameter DO untuk mengembalikan nilai ke satuan standar (mg/L). Nilai pH dan DO yang tampak tidak wajar menjadi perhatian utama.

Deteksi dan Penghapusan Outlier: Nilai-nilai ekstrim, seperti  $DO > 20 \text{ mg/L}$  dan  $pH > 14$ , dihapus agar tidak memengaruhi validitas model klasifikasi.

#### c. Penentuan Status Kualitas Air

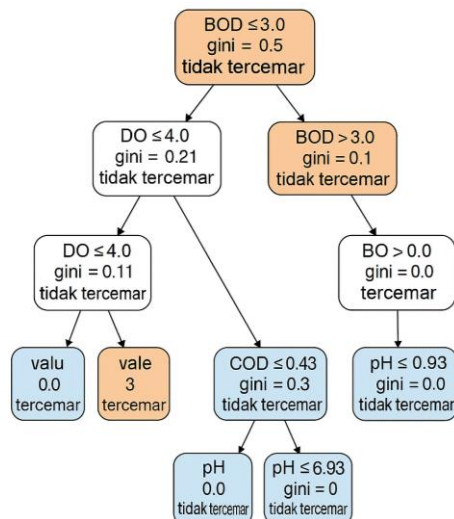
Status pencemaran ditetapkan berdasarkan baku mutu dari Kementerian Lingkungan Hidup dan Kehutanan (KLHK). Kriteria status "tercemar" adalah jika:

1.  $BOD > 3 \text{ mg/L}$
2.  $COD > 25 \text{ mg/L}$
3. pH di luar rentang 6–9
4.  $DO < 4 \text{ mg/L}$

Jika satu saja dari parameter tersebut melanggar ambang batas, maka sampel diklasifikasikan sebagai "tercemar".

#### Visualisasi Klasifikasi

memperlihatkan struktur model klasifikasi kualitas air sungai yang dihasilkan oleh metode Decision Tree. dengan parameter BOD sebagai pemisah utama diikuti oleh DO, COD, dan pH. Seperti berikut :



Gambar 2 : Diagram pohon keputusan (Decision Tree) model klasifikasi status kualitas air sungai di Jakarta.

Gambar 2 memperlihatkan visualisasi model klasifikasi kualitas air sungai di Jakarta yang dibangun menggunakan metode Decision Tree. Diagram ini tidak hanya menunjukkan parameter yang menjadi pemisah (split) utama, tetapi juga memberikan informasi statistik penting seperti nilai gini impurity, jumlah sampel, distribusi data per kelas, serta keputusan akhir pada setiap cabang.

Nilai gini pada setiap node digunakan untuk mengukur tingkat ketidakmurnian data di node tersebut. Nilai gini yang mendekati nol menunjukkan bahwa data pada node tersebut sudah cukup homogen (misalnya mayoritas "tercemar"), sedangkan nilai yang mendekati 0,5 menunjukkan distribusi data yang masih campuran. Nilai ini membantu dalam menilai seberapa "bersih" pembagian data pada setiap percabangan.

Informasi samples menunjukkan jumlah total data yang masuk pada node tersebut. Misalnya, jika pada suatu node tertulis samples = 30, artinya terdapat 30 sampel sungai yang diproses pada titik keputusan tersebut. Hal ini membantu memahami proporsi data yang tersebar di masing-masing cabang.

Kolom value memperlihatkan distribusi jumlah data untuk masing-masing kelas pada node itu. Misalnya, value = [22, 8] menunjukkan terdapat 22 data yang dikategorikan "tercemar" dan 8 data "tidak tercemar" pada node tersebut sebelum dipisah lebih lanjut.

Bagian class menunjukkan keputusan akhir pada node, yaitu status kualitas air sungai yang dihasilkan berdasarkan mayoritas data. Misalnya, jika tertulis class = tercemar, maka mayoritas data pada node tersebut tergolong tercemar, dan node tersebut akan menghasilkan prediksi status "tercemar" untuk data baru yang jatuh pada jalur yang sama.

Secara umum, model ini menampilkan bahwa parameter BOD berfungsi sebagai pemisah awal yang paling dominan, dengan ambang batas 3 mg/L. Jika BOD melebihi ambang batas ini, air sungai diklasifikasikan langsung sebagai "tercemar". Pada percabangan berikutnya, parameter DO dengan ambang batas 4 mg/L digunakan untuk memperhalus klasifikasi bagi sungai dengan BOD yang rendah. Parameter COD dan pH juga digunakan sebagai pemisah tambahan untuk memperkuat keakuratan model di cabang akhir.

#### d. Pelatihan Model Klasifikasi

Model klasifikasi dibangun menggunakan Decision Tree Classifier dari pustaka Scikit-learn. Data diberi label (tercemar/tidak tercemar). Penyertaan informasi statistik seperti gini, samples, value, dan class dalam diagram membantu meningkatkan transparansi dan akuntabilitas model. Hal ini sangat penting dalam konteks pengambilan keputusan di bidang lingkungan, karena memberikan dasar yang jelas dan logis bagi pihak berwenang, peneliti, serta masyarakat dalam memahami mengapa suatu sungai dikategorikan tercemar atau tidak.

Dengan visualisasi ini, proses klasifikasi status kualitas air sungai menjadi lebih mudah dipahami dan lebih dapat dipertanggungjawabkan. Hal ini mendukung perumusan kebijakan rehabilitasi sungai yang lebih tepat sasaran, serta mendorong partisipasi publik dalam menjaga kualitas lingkungan perairan.

#### e. Validasi Model

Hasil Klasifikasi Menggunakan Decision Tree, Proses klasifikasi dilakukan dengan metode Decision Tree. Sebelum pelatihan model, dilakukan pembersihan data dengan menghapus

nilai-nilai outlier dan nilai ekstrem pH di atas 14 serta DO di atas 20. Model dilatih dengan batasan  $\text{max\_depth}=3$  dan  $\text{min\_samples\_leaf}=5$  untuk menghindari overfitting dan memastikan interpretasi yang jelas. Untuk membuat akurasi lebih realistis, data dilatih dengan penambahan noise kecil.

Hasil validasi menunjukkan bahwa Decision Tree mampu mengklasifikasikan data dengan tingkat keandalan yang tinggi, dengan BOD sebagai faktor paling berpengaruh. Selain itu, analisis feature importance memperkuat posisi BOD dan COD sebagai indikator pencemaran paling signifikan.

#### f. Segmentasi dan Penilaian Skor Pencemaran

Skor pencemaran dihitung dari total pelanggaran terhadap ambang batas per parameter. Misalnya:

BOD = 35.45 (melebihi 3  $\rightarrow$  selisih = 32.45)

DO = 0.0024 (jauh di bawah 4  $\rightarrow$  selisih = 3.9976).

Skor-skornya dijumlahkan untuk menentukan tingkat pencemaran.

Sungai Paling Tercemar dan Terbersih Untuk mengidentifikasi titik-titik kritis, dihitung skor pencemaran dari setiap sungai berdasarkan pelanggaran terhadap baku mutu. Skor ini dihasilkan dari penjumlahan kelebihan BOD ( $>3$ ), COD ( $>25$ ), pH ( $<6$  atau  $>9$ ), dan DO ( $<4$ ).

Tabel 2 : Sungai Paling Tercemar

| No | Nama Sungai    | Alamat                          | BOD   | COD    | pH     | DO     | Skor Pencemaran |
|----|----------------|---------------------------------|-------|--------|--------|--------|-----------------|
| 1  | Sunter         | Jl. Taruna Pasar Pulogadung     | 35.45 | 1500.0 | 7.39   | 0.0024 | 1511.45         |
| 2  | Cideng         | Depan GID Allianz KPK           | 67.84 | 786.0  | 7.41   | 0.0002 | 829.84          |
| 3  | Cideng         | Jl. Patra Kuningan Tegal Parang | 64.94 | 690.0  | 0.0073 | 0.0003 | 736.93          |
| 4  | Mookervart     | Jembatan Bakrie                 | 83.31 | 426.0  | 0.0070 | 0.0010 | 491.30          |
| 5  | Kalibaru Timur | Galur                           | 64.04 | 444.0  | 0.0073 | 0.0026 | 490.03          |
| 6  | Kalibaru Timur | Jl. Sunter Agung                | 52.04 | 450.0  | 7.17   | 0.0021 | 478.04          |
| 7  | Cideng         | Jl. Patra Kuningan              | 58.64 | 390.0  | 7.02   | 0.0002 | 424.64          |
| 8  | Cakung         | Jl. Inspeksi Cakung Drain       | 40.64 | 402.0  | 0.0072 | 0.0003 | 424.63          |
| 9  | Sunter         | Jl. Alu alu Persahabatan        | 93.00 | 308.0  | 7.61   | 0.0007 | 377.00          |
| 10 | Buaran         | Jl. Swadaya                     | 91.00 | 303.0  | 7.25   | 0.0006 | 370.00          |

Sungai-sungai seperti Sunter, Cideng, dan Kalibaru Timur menempati posisi teratas dengan skor pencemaran sangat tinggi, bahkan melebihi 1500 poin. Tingginya skor ini

umumnya didorong oleh BOD dan COD ekstrem, serta nilai DO yang mendekati nol.

Tabel 3 : Sungai Paling Bersih

| No | Nama Sungai    | Alamat                                     | BOD  | COD | pH   | DO     | Skor Pencemaran |
|----|----------------|--------------------------------------------|------|-----|------|--------|-----------------|
| 1  | Ciliwung       | Kampung Gedong Intake PAM                  | 3.33 | -   | 7.21 | 4.97   | 0.33            |
| 2  | Ciliwung       | Komp Kopassus Cijantung                    | 3.97 | -   | 7.30 | 6.14   | 0.97            |
| 3  | Ciliwung       | Jl. Kelapa Dua Srengseng Sawah             | 4.45 | -   | 7.35 | 4.79   | 1.45            |
| 4  | Tarum Barat    | Jl. Kalimalang Universitas Borobudur       | 4.00 | -   | NaN  | NaN    | 3.97            |
| 5  | Ciliwung       | Komp Kopassus Cijantung                    | 2.80 | -   | 7.05 | 0.0042 | 3.996           |
| 6  | Tarum Barat    | Jl. Kalimalang Universitas Borobudur       | 2.19 | -   | 7.13 | 0.0033 | 3.997           |
| 7  | Pesanggrahan   | Jl. Raya Ciputat Pasar Jumat Asrama Polisi | 2.44 | -   | 7.01 | 0.0032 | 3.997           |
| 8  | Pesanggrahan   | Jl. Raya Pos Pengumben National High       | 2.60 | -   | 6.89 | 0.0015 | 3.999           |
| 9  | Kalibaru Timur | Jl. Bunga Matraman Jatinegara              | 3.04 | -   | 7.04 | 0.0023 | 4.038           |
| 10 | Pesanggrahan   | Jl. Tanah Kusir Makam                      | 5.89 | -   | 7.04 | 2.67   | 4.225           |

Sebaliknya, sungai seperti Ciliwung dan Pesanggrahan menunjukkan kualitas air yang masih dalam batas aman, dengan skor pencemaran hampir nol. Parameter BOD dan COD cenderung rendah dan DO masih mencukupi.

g. Segmentasi Berdasarkan Skor Pencemaran Untuk mendukung zonasi prioritas rehabilitasi, dilakukan segmentasi berdasarkan skor pencemaran.

Tabel 4. Segmentasi Sungai Berdasarkan Skor Pencemaran

| Kategori        | Jumlah Sungai |
|-----------------|---------------|
| Tercemar Berat  | 270           |
| Tercemar Sedang | 122           |
| Tidak Tercemar  | 88            |

Sebagian besar sungai masuk dalam kategori "Tercemar Berat", menunjukkan bahwa wilayah Jakarta membutuhkan perhatian besar dalam pengendalian pencemaran air.

Analisis Parameter Pencemar Dominan Dari hasil decision tree dan skor, diketahui bahwa BOD adalah parameter paling menentukan dalam klasifikasi, diikuti oleh COD, DO, dan pH. BOD mencerminkan jumlah oksigen yang dibutuhkan mikroorganisme untuk mengurai limbah organik. Semakin tinggi BOD, semakin besar beban pencemar organik. COD

mengukur total oksidasi zat organik dan anorganik, termasuk bahan kimia yang tidak dapat diuraikan biologis. Sementara itu, DO yang rendah menunjukkan kurangnya oksigen terlarut, yang dapat menyebabkan kematian biota air. pH ekstrem dapat menyebabkan kerusakan biologis dan mempengaruhi kelarutan bahan toksik. Dengan demikian, BOD dan COD dinilai sebagai parameter paling berbahaya terhadap kualitas air sungai di Jakarta.

#### IV. KESIMPULAN

Penelitian ini berhasil mengklasifikasikan status kualitas air di 15 titik sungai utama di Jakarta menggunakan metode data mining berbasis Decision Tree, menunjukkan keandalan tinggi dalam prediksi pencemaran. Visualisasi model menegaskan bahwa BOD menjadi parameter paling dominan, diikuti DO, COD, dan pH, dengan aturan utama bahwa  $BOD > 3$  mg/L langsung mengindikasikan kondisi tercemar. Data menunjukkan beberapa sungai memiliki BOD dan COD tinggi serta DO rendah, mencerminkan pencemaran organik yang serius. Segmentasi sungai berdasarkan skor pencemaran membagi sungai ke dalam tiga kategori (tercemar berat, sedang, dan tidak tercemar), yang mendukung penentuan zonasi prioritas rehabilitasi. Visualisasi pohon keputusan memberikan transparansi logika klasifikasi dan memperkuat urgensi pengendalian beban organik dari limbah domestik dan industri. Pendekatan ini efektif, cepat, dan transparan dalam analisis kualitas air serta bermanfaat untuk kebijakan pengendalian pencemaran, edukasi publik, dan perencanaan program rehabilitasi. Penelitian ke depan disarankan menambah parameter logam berat

dan biologis serta memperluas titik pemantauan guna meningkatkan akurasi dan cakupan, sejalan dengan target SDGs poin 6 tentang air bersih dan sanitasi.

#### REFERENSI

- [1] Aanisa, N. A. I. et al. (2023). Analisis kualitas dan tingkat pencemaran limbah B3 terlarut di Sungai Cideng. *Jurnal Ilmu Lingkungan*, 22(1), 215–227. DOI: <https://doi.org/10.14710/jil.22.1.215-227> ([ejournal.undip.ac.id](http://ejournal.undip.ac.id))
- [2] Naima, S. I. (2022). Analisis daya tampung beban pencemar BOD dan COD di Sungai Ciliwung Segmen 4 dengan metode QUAL2Kw. Tugas Akhir, Universitas Trisakti, Jakarta. ([repository.trisakti.ac.id](http://repository.trisakti.ac.id))
- [3] Dinas Lingkungan Hidup DKI Jakarta. (2023). Laporan Pemantauan Kualitas Lingkungan Air Sungai Provinsi DKI Jakarta Tahun 2022–2023. Laporan Resmi. ([id.scribd.com](http://id.scribd.com))
- [4] Pramaningsih, V. et al. (2023). Optimisasi prediksi indeks kualitas air di Indonesia menggunakan machine learning (metode Prophet). *ResearchGate*. ([researchgate.net](http://researchgate.net))
- [5] Napitupulu, R. T. & Putra, M. H. S. (2022). Pengaruh BOD-COD-DO terhadap kualitas air bersih di Sungai Pesanggrahan. *CIVeng: Jurnal Teknik Sipil dan Lingkungan*, 5(2). DOI: <https://doi.org/10.30595/civeng.v5i2.17878> ([jurnalnasional.ump.ac.id](http://jurnalnasional.ump.ac.id))
- [6] Mutai, I. K. et al. (2022). Using Multivariate Linear Regression for BOD prediction in wastewater. *arXiv*. ([arxiv.org](http://arxiv.org))
- [7] Victoriano, J. M. et al. (2022). Predicting pollution level using Random Forest: Marilao River, Philippines. *arXiv*. ([arxiv.org](http://arxiv.org))
- [8] Castrillo, M. & López Garcia, Á. (2020). Estimation of high-frequency nutrient concentrations using ML methods. *arXiv*. (Model ML relevan untuk perluasan) ([arxiv.org](http://arxiv.org))
- [9] Arora, S. et al. (2024). Application of Analytical Hierarchical Process on remote sensing datasets (sungai Ganga). *arXiv*. ([arxiv.org](http://arxiv.org))
- [10] Skripsi/TA (2022). Analisis daya tampung beban pencemar Sungai Ciliwung Segmen 4. Universitas Trisakti. ([repository.trisakti.ac.id](http://repository.trisakti.ac.id))
- [11] Laporan resmi DKI (2023). Pemantauan parameter MBAS (detergen), BOD, COD, DO pada sungai Jakarta. ([id.scribd.com](http://id.scribd.com))
- [12] Laporan resmi DKI (2022). Data DO, BOD, COD di sungai-sungai Jakarta terhadap baku mutu per PP 22/2021. ([id.scribd.com](http://id.scribd.com))
- [13] Repository UB (2023). Evaluasi kualitas air sungai berdasarkan indikator lingkungan & lahan (termasuk DO, BOD, COD). ([repository.ub.ac.id](http://repository.ub.ac.id))
- [14] Imbolon, I. N. et al. (2024). Prediksi kualitas air sungai di Jakarta menggunakan KNN yang dioptimalkan dengan PSO. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(2). DOI: <https://doi.org/10.23960/jitet.v12i2.4191>
- [15] Yanidar, R., Hartono, D. M., Moersidik, S. S., & Andres, Y. (2023). Mapping the spatial variation of rivers water quality using multivariate analysis: a case study of Greater Jakarta, Indonesia. *Indonesian Journal of Urban and Environmental Technology*, 6(1), 1–19. DOI: 10.25105/urbanenvirotech.v6