

Klasifikasi Kualitas Air Sungai Menggunakan Metode KNN Berbasis Web (Studi Kasus Dinas Lingkaran Hidup dan Kebersihan Aceh Utara)

Muhammad Fathan¹, Mulyadi^{2*}, Radhiyatammardiyah³

^{1,2,3} *Jurusan Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe
Jln. B. Aceh Medan Km.280 Buketrata 24301 INDONESIA*

¹fathandark@gmail.com

^{2*}mulyadi@pnl.ac.id

³radhiatammardiyah@pnl.ac.id

Di Indonesia, masalah pelayanan air bersih menjadi perhatian utama karena dampaknya yang serius terhadap kesehatan masyarakat. Berdasarkan data dari WHO (World Health Organization) tahun 2021, setiap tahunnya sekitar 1,7 juta anak meninggal akibat diare yang disebabkan oleh lingkungan yang tidak sehat, terutama akibat air yang tercemar. Aceh Utara menghadapi tantangan serius dengan menurunnya kualitas air sungai akibat pertumbuhan industri dan limbah kelapa sawit, terutama di sungai-sungai utama seperti Peuto, Keuruto, dan Pase. Evaluasi efektivitas algoritma K-Nearest Neighbor (KNN) dalam mengklasifikasikan kualitas air sungai menjadi penting. Algoritma ini diterapkan dengan menggunakan delapan parameter kualitas air, termasuk TSS, BOD, COD, PO4, DO, E. coli, pH, dan NO3-N. Pengujian menggunakan RapidMiner menunjukkan tingkat akurasi KNN sebesar 51,85%, Ketika menggunakan Confusion matrix, akurasi meningkat menjadi 54,28%.

Kata kunci: KNN, Supervised Learning, Machine learning, Klasifikasi air sungai

In Indonesia, the issue of clean water service is a major concern due to its serious impact on public health. According to data from the WHO (World Health Organization) year 2021, around 1.7 million children die each year from diarrhea caused by unhealthy environments, particularly due to contaminated water. Aceh Utara faces significant challenges with declining river water quality due to industrial growth and palm oil waste, especially in major rivers like Peuto, Keuruto, and Pase. Evaluating the effectiveness of the K-Nearest Neighbor (KNN) algorithm in classifying river water quality has become crucial. This algorithm is applied using eight water quality parameters, including TSS, BOD, COD, PO4, DO, E. coli, pH, and NO3-N. Testing using RapidMiner showed a KNN accuracy rate of 51.85%, and when using the Confusion matrix, the accuracy increased to 54.28%.

Keywords: KNN, Supervised Learning, Machine learning, River Water Classification

I. PENDAHULUAN

Di Indonesia, masyarakat masih menghadapi sejumlah permasalahan yang rumit dan belum sepenuhnya terselesaikan hingga saat ini. Salah satu isu yang signifikan adalah tingkat pelayanan air bersih yang masih rendah untuk penduduk. Menurut informasi dari Organisasi Kesehatan Dunia (WHO), setiap tahunnya sekitar 1,7 juta anak meninggal akibat penyakit diare yang disebabkan oleh lingkungan yang tidak sehat, terutama karena air yang terkontaminasi [1]-[2]. Aceh Utara menghadapi tantangan serius terkait kualitas air sungainya yang semakin buruk akibat pertumbuhan industri, pembuangan limbah cair kelapa sawit, dan peningkatan jumlah penduduk. Pencemaran air dapat mengakibatkan penurunan kualitas perairan, yang pada gilirannya akan mengurangi kemampuan perairan untuk mendukung kelangsungan hidup yang ada di dalamnya. Perubahan yang terjadi pada perairan sungai dapat mempengaruhi terhadap organisme perairan yang berada dalam sungai tersebut.[3]-[4]. Sungai-sungai utama di Aceh Utara seperti Sungai Peuto, Keuruto, dan Pase, terkena dampak limbah cair dan bahan kimia dari aktivitas industri dan limbah cair kelapa sawit . Oleh karena itu, penerapan sistem klasifikasi kualitas air yang mampu menyediakan informasi menjadi penting. Klasifikasi kualitas air sungai berdasarkan parameter-parameter tertentu. Parameter-parameter yang sering digunakan untuk

menentukan kualitas air antara lain Total Suspended Solids (TSS), Biological Oxygen Demand (BOD), Chemical Oxygen Demand (COD), fosfat (PO4), 2 Dissolved Oxygen (DO), bakteri E-coli, pH, dan Nitrat (NO3-N). Berdasarkan nilai-nilai dari parameter-parameter ini, kualitas air sungai dapat dikategorikan ke dalam tingkat cemaran ringan, cemaran sedang, dan cemaran berat. Pengembangan sistem ini dengan metode K-Nearest Neighbors(KNN) dan Metode Storet berbasis web di Aceh Utara dianggap sebagai langkah strategis untuk menjaga dan meningkatkan klasifikasi air sungai, serta memastikan kesehatan ekosistem dan masyarakat setempat. Metode K-Nearest Neighbor (KNN) merupakan algoritma Machine learning yang digunakan untuk mengklasifikasikan atau melakukan regresi berdasarkan kedekatan suatu data dengan data lain dalam ruang fitur [5]-[6]. Klasifikasi adalah bentuk pembelajaran yang dipandu (Supervised Learning) yang dimanfaatkan untuk memprediksi kelas suatu objek yang kelasnya belum diketahui [7]-[8]. Water Quality Index (WQI) adalah suatu instrumen yang disederhanakan untuk memungkinkan masyarakat umum mendapatkan informasi tentang kualitas air yang melibatkan data yang kompleks. WQI memberikan nilai tunggal yang mencerminkan kualitas air secara holistik dan digunakan untuk mengevaluasi tingkat baik atau buruknya kualitas air di suatu lokasi. WQI digunakan untuk menghitung nilai indeks kualitas air, memberikan visi terhadap kondisi air di setiap

kelompok lokasi, dan memudahkan penilaian kualitas air secara menyeluruh [9]-[10] Integrasi KNN dan Storet memberikan wawasan yang komprehensif, membantu dalam pengelompokan lokasi berdasarkan kemiripan, dan memberikan evaluasi kualitas air yang akurat. Semua informasi tersebut diakses melalui sistem 3 berbasis web dan memungkinkan pihak-pihak terkait dan masyarakat umum mendapatkan informasi kualitas air secara efisien [11]-[12]. Klasifikasi air bukan hanya tentang mengidentifikasi tingkat pencemaran, melainkan juga menciptakan dasar informasi untuk mendukung kebijakan dan keputusan berkelanjutan. Sungai Keureto, Pase, dan Peuto, sebagai sumber bahan baku PDAM yang digunakan oleh masyarakat Kabupaten Aceh Utara, memiliki beberapa instalasi pengolahan air. Ini menunjukkan kebutuhan penting untuk menjaga kualitas air di sumber guna menjamin pasokan air bersih yang aman dan berkualitas bagi masyarakat setempat. Oleh karena itu, untuk menjawab rasa penasaran mengenai kecocokan algoritma K-Nearest Neighbor (KNN) dengan parameter kualitas air sungai, sistem ini bertujuan memberikan solusi pendekatan dan terintegrasi dalam klasifikasi air di wilayah tersebut.

II. METODOLOGI PENELITIAN

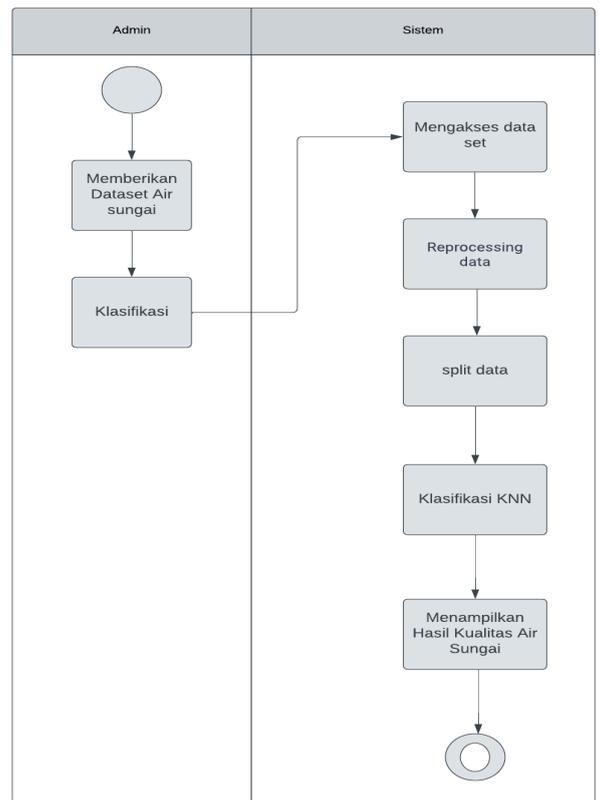
Bagian ini menjelaskan secara rinci tentang penelitian yang dilakukan.

A. Teknik Pengumpulan Data

Data yang telah terkumpul berjumlah 140 data dan awalnya dalam bentuk file PDF. Data ini kemudian diubah menjadi format tabel Excel agar dapat digunakan untuk proses klasifikasi dengan metode K-Nearest Neighbors (KNN). Proses persiapan data dimulai dengan mengkonversi data dari PDF ke Excel dan menyusun data dalam tabel yang rapi. Setelah data dalam bentuk tabel Excel siap, langkah berikutnya adalah melakukan preprocessing data untuk memastikan bahwa data tersebut bersih dan siap untuk analisis lebih lanjut. Proses preprocessing meliputi pembersihan data, penanganan nilai yang hilang, dan normalisasi data. Setelah 18 preprocessing selesai, data dibagi menjadi dua set, yaitu set pelatihan dan set pengujian. pengumpulan data adalah untuk memperoleh dataset yang akan digunakan pada proses data latih dan data uji. Dataset pelatihan digunakan untuk melatih model KNN, sementara dataset pengujian digunakan untuk mengevaluasi kinerja model tersebut. Setelah proses split data selesai, data tersebut kemudian dimasukkan ke dalam algoritma klasifikasi KNN. Hasil klasifikasi ini dapat digunakan untuk pemantauan kualitas air dan pengelolaan sumber daya air di wilayah tersebut.

B. Arsitektur Umum

Perancangan arsitektur umum pada klasifikasi kualitas air sungai menggunakan metode KNN berbasis web (studi kasus dinas lingkungan hidup dan kebersihan aceh utara) pada gambar 2.



Gambar 2. Arsitektur Umum

Dataset yang sudah tersedia akan dilakukan proses *preprocessing data* untuk dilakukan pembersihan data dan dilanjutkan dengan proses *split data* untuk membagi dataset menjadi dua bagian: data pelatihan dan data pengujian. Di Lanjutkan klasifikasi KNN dan mengujinya dengan *confussion matrix*. Setelah itu data akan di masukkan dan di gunakan ke dalam website.

C. Dataset

Dataset yang digunakan dalam penelitian ini berjumlah 140 data. Dataset didapatkan langsung dari kantor Dinas Lingkungan Hidup dan Kebersihan.

D. Preprocessing Data

Preprocessing data dilakukan untuk memastikan bahwa data tersebut bersih dan siap untuk analisis lebih lanjut. Proses *preprocessing* meliputi pembersihan data, penanganan nilai yang hilang, dan normalisasi data.

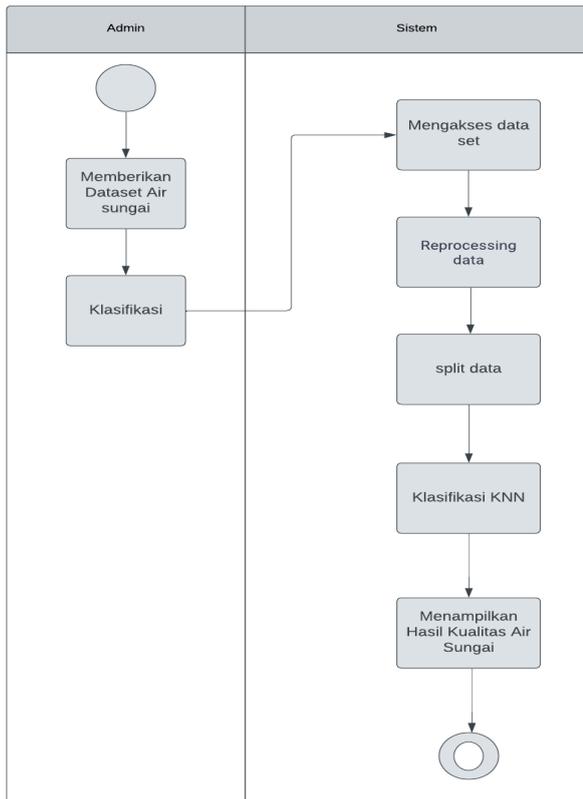
E. Split Data

Proses split data melibatkan pembagian dataset menjadi dua bagian: training set dan test set.

Bagian training berfungsi untuk melatih model, sedangkan test set digunakan untuk mengukur performa model. Biasanya, data dibagi dengan 70-80% untuk pelatihan dan 20-30% untuk pengujian. Langkah ini bertujuan untuk memastikan model bekerja baik pada data baru dan mencegah model terlalu menyesuaikan diri dengan data pelatihan.

F. Klasifikasi Metode KNN

Metode K-Nearest Neighbors (KNN) adalah algoritma yang digunakan untuk klasifikasi dan regresi berdasarkan kedekatan antar data. KNN menentukan kelas data dengan melihat mayoritas dari tetangga terdekat. Jarak antara titik data biasanya dihitung menggunakan Euclidean atau Manhattan distance. Pemilihan nilai K sangat krusial; K kecil membuat model lebih sensitif, sedangkan K besar menghasilkan hasil yang lebih konsisten. Klasifikasi metode KNN adalah proses di mana admin mengelola data kualitas air sungai, yang kemudian diproses menggunakan metode KNN untuk mengklasifikasikan parameter air dan menentukan label kualitasnya.



Gambar 3. Proses klasifikasi KNN

Pada gambar 3. memperlihatkan langkah-langkah klasifikasi data. Proses ini dimulai dengan input dataset oleh admin, diikuti dengan menekan tombol klasifikasi. Sistem kemudian mengakses dataset, melakukan proses pemrosesan data, dan akhirnya menampilkan hasil klasifikasi untuk evaluasi kualitas air.

G. Confusion Matrix

Confusion matrix merupakan salah satu metode yang digunakan untuk mengukur kinerja metode klasifikasi. Rumus pengujian akan ditampilkan pada persamaan 3.1, persamaan 3.2, dan persamaan 3.3 berikut.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \dots\dots\dots 3.1$$

$$\text{Presisi} = \frac{TP}{FP+TP} \times 100\% \dots\dots\dots 3.2$$

$$\text{Recall} = \frac{TP}{FN+TP} \times 100\% \dots\dots\dots 3.3$$

TN : Data negatif yang terdeteksi dengan benar

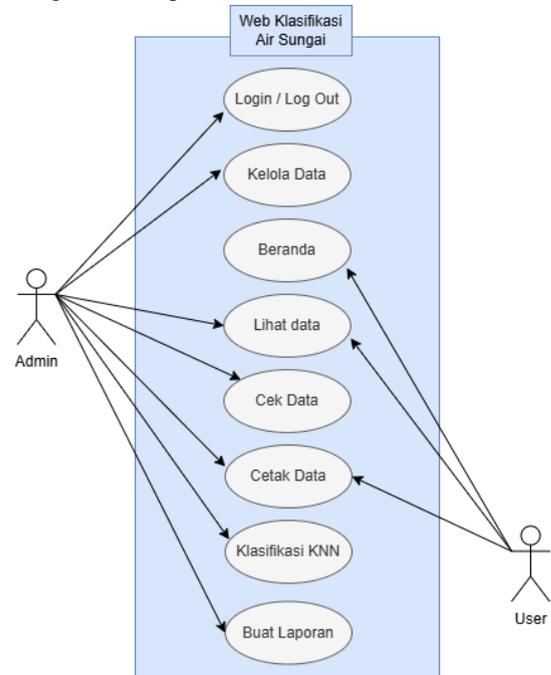
FP : Data negatif namun terdeteksi sebagai data positif

TP : Data positif terdeteksi benar

FN : Data positif terdeteksi sebagai data negative

H. Use Case Diagram

Use case diagram menggambarkan interaksi antara aktor dengan sistem pada Gambar 4.



Gambar 4. Use case diagram

Gambar 4. diatas menggambarkan diagram use case diagram untuk sistem klasifikasi kualitas air sungai. Penjelasan peran Admin dan user dalam konteks ini dapat diuraikan sebagai berikut :

a. Definisi Aktor

Penjelasan mengenai Aktor dalam sistem diterangkan melalui representasi tabel berikut :

1. Admin Aktor yang memiliki hak akses untuk melihat dan mengelola informasi mengenai kualitas air sungai, Parameter air sungai dan dapat mencetaknya jika di perlukan

2. User Aktor yang memiliki akses untuk melihat data kualitas air sungai dan mencetaknya jika diperlukan.

b. Definisi Use case

Defisini untuk use case nya adalah sebagai berikut :

1. Kelola Data Proses di mana admin dapat menambah, memodifikasi, dan menghapus informasi mengenai kualitas air di berbagai lokasi pengambilan air.

2. Beranda Tampilan utama bagi pengguna dan admin yang menyajikan informasi tentang kantor dan kondisi air sungai.

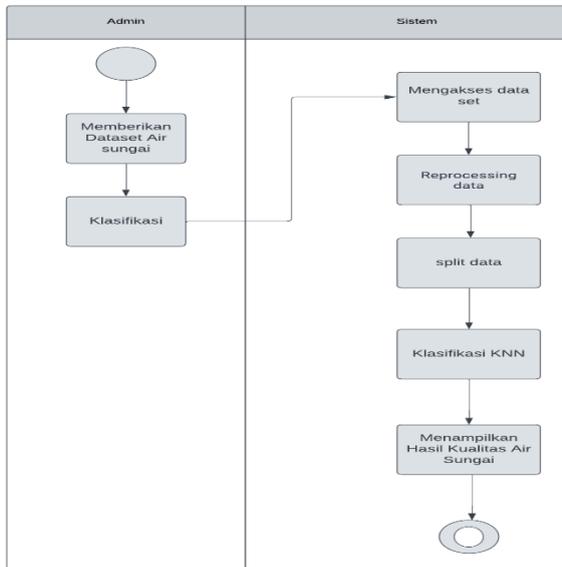
3. Lihat Data Proses di mana User dapat melihat informasi kualitas dan klasifikasi air sungai

4. Klasifikasi KNN Proses di mana admin mengelola data kualitas air sungai, yang kemudian diproses

menggunakan metode KNN untuk mengklasifikasikan parameter air dan menentukan label kualitasnya.

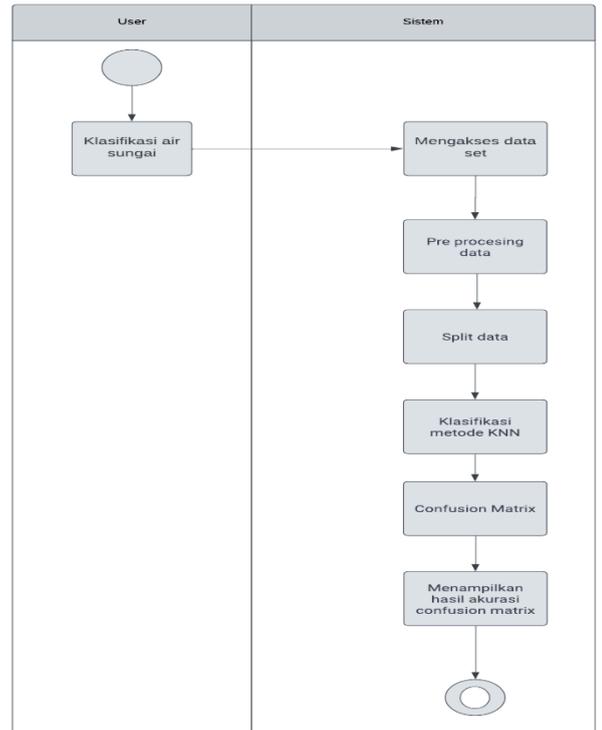
5. Buat Laporan Proses yang memungkinkan admin untuk mencetak informasi kualitas air jika diperlukan.

I. *Activity Diagram* Klasifikasi Metode KNN



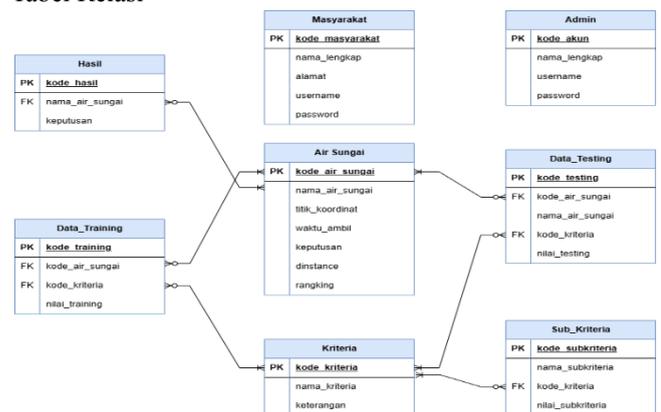
Gambar 5. *Activity Diagram* Klasifikasi KNN
 Pada gambar 5. memperlihatkan langkah-langkah klasifikasi data. Proses ini dimulai dengan input dataset oleh admin, diikuti dengan menekan tombol klasifikasi. Sistem kemudian mengakses dataset, melakukan proses pemrosesan data, dan akhirnya menampilkan hasil klasifikasi untuk evaluasi kualitas air.

J. *Activity Diagram* Confusion Matrix



Gambar 6. *Activity Diagram* Confusion Matrix
 Pada gambar 6. diagram tersebut menggambarkan proses klasifikasi air sungai menggunakan metode KNN (K-Nearest Neighbors), dimulai dari admin yang memulai klasifikasi, dilanjutkan oleh sistem yang mengakses dataset, melakukan pre-prosesing untuk membersihkan dan menormalkan data, membagi dataset menjadi data pelatihan dan data pengujian, menerapkan metode KNN untuk klasifikasi, menghasilkan confusion matrix untuk mengevaluasi kinerja model, dan akhirnya menampilkan hasil akurasi dari confusion matrix untuk menunjukkan seberapa baik model mengklasifikasikan data dengan benar.

K. Tabel Relasi



Gambar 7. *Activity Diagram* Labelling Data
 berikut adalah penjelasan dari struktur tabels relasi yang ada pada gambar 7:

1. Tabel Admin:

Menyimpan informasi admin dengan pengenal unik 'id_admin'. Penggunaan username dan password bertujuan untuk autentikasi admin.

2. Tabel Masyarakat:

Menyimpan data masyarakat dengan identifikasi unik kode_masyarakat. Kolom nama_lengkap, alamat, username, dan password menyimpan informasi detail mengenai masyarakat.

3. Tabel Hasil:

Menyimpan hasil analisis kualitas air dengan identifikasi unik kode_hasil dan kode_air_sungai sebagai foreign key yang merujuk ke tabel Air Sungai. Kolom keputusan mencatat hasil analisis kualitas air.

4. Tabel Air Sungai:

Menyimpan data mengenai sungai dengan identifikasi unik kode_air_sungai. Kolom nama_air_sungai, titik_koordinat, waktu_ambil, keputusan, distance, dan rangking mencakup informasi lengkap mengenai sungai dan pengambilan sampel air.

5. Tabel Data_Testing:

Menyimpan data pengujian kualitas air dengan identifikasi unik kode_testing dan kode_air_sungai sebagai kunci asing yang merujuk ke tabel Air Sungai. Kolom kode_kriteria juga sebagai kunci asing yang merujuk ke tabel Kriteria. Kolom nilai_testing merepresentasikan hasil pengujian kualitas air.

6. Tabel Data_Training:

Menyimpan data pelatihan kualitas air dengan identifikasi unik kode_training dan kode_air_sungai sebagai kunci asing yang merujuk ke tabel Air Sungai. Kolom kode_kriteria juga sebagai kunci asing yang merujuk ke tabel Kriteria. Kolom nilai_training merepresentasikan hasil pelatihan kualitas air.

7. Tabel Kriteria

Menyimpan informasi mengenai kriteria dengan identifikasi unik kode_kriteria. Kolom nama_kriteria dan keterangan mendeskripsikan kriteria yang digunakan dalam analisis kualitas air.

8. Tabel Sub_Kriteria:

Menyimpan data subkriteria dengan identifikasi unik kode_subkriteria dan kode_kriteria sebagai kunci asing yang merujuk ke tabel Kriteria. Kolom nama_subkriteria dan nilai_subkriteria mencatat detail tambahan mengenai kriteria.

Struktur tabel relasi ini menyatukan data antar tabel dengan menggambarkan hubungan dan arus informasi yang efektif dalam sistem manajemen kualitas air. Implementasi semua elemen kunci dimaksudkan untuk memastikan kestabilan dan konsistensi data, membentuk landasan yang kuat untuk pengelolaan informasi yang akurat dan terpercaya. Dengan demikian, tabel relasi ini tidak

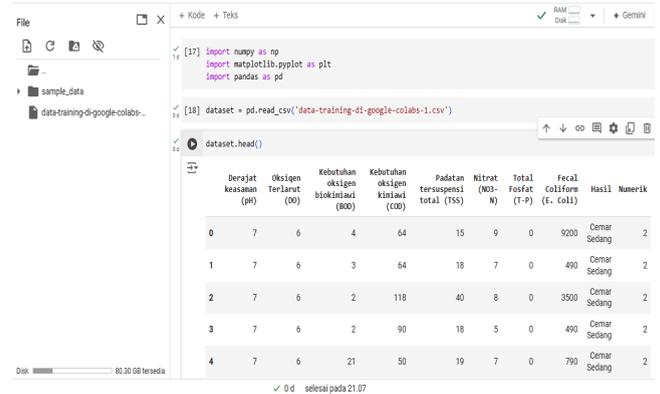
hanya mempermudah pengaturan data, tetapi juga memberikan dasar yang kuat bagi sistem yang efisien dan dapat diandalkan dalam klasifikasi dan manajemen kualitas air.

III. HASIL DAN PEMBAHASAN

A. Implementasi Proses

Implementasi proses menjelaskan langkah-langkah dan prosedur yang terlibat berdasarkan desain model *machine learning* yang telah dirancang:

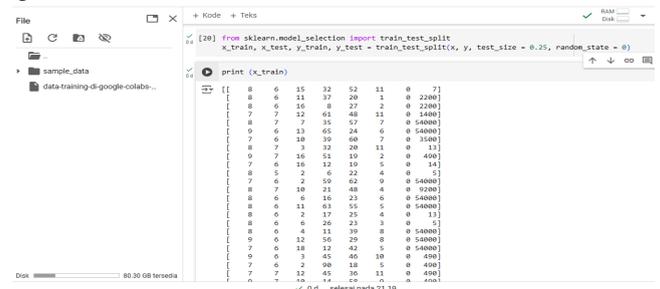
1. Dataset



Gambar 8. Dataset

Gambar 8 memperlihatkan dataset yang telah dikonversi menjadi format CSV. Dataset ini sudah diorganisir sedemikian rupa sehingga siap digunakan dalam proses klasifikasi. Format CSV memudahkan pengolahan data untuk analisis lebih lanjut menggunakan algoritma klasifikasi KNN

2. Split Data



Gambar 9. Split Data

Pada gambar 9 setelah data disiapkan, penggunaan fungsi train_test_split memungkinkan pengembangan untuk membagi dataset menjadi dua bagian: data pelatihan dan data pengujian. dengan cara ini, pengujian kinerja model dapat dilakukan secara terstruktur dan efisien untuk mengevaluasi performa serta akurasi model secara lebih baik.

3. Preprocessing Data

```

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.transform(x_test)

[35] print(x_train)
[[ 0. -0.56972371  1.11544801 -0.25399561  1.87781184  1.43334881
  0. -0.56468993]
 [ 0. -0.56972371  0.37655082  0.68886135 -1.13368018 -1.65204099
  0. -0.60805843]
 [ 0. -0.56972371  1.38816326 -1.51187863 -0.64889429 -1.34432839
  0. -0.60805843]
 [-1.53741223  1.47342338  0.50127267  1.20594637  0.88148339  1.43334881
  0. -0.60805843]
 [ 0. -1.47342338 -0.36234858 -0.89676823  1.4233881  0.10882381
  0. 1.38816326]
 [ 1.53741223 -0.56972371  0.74589532  1.47559354 -0.85722253 -0.10882381
  0. 1.38816326]
 [-1.53741223 -0.56972371  0.18182737  0.11288694  1.63871634  0.10882381
  0. -0.39597548]
 [ 0. -1.47342338 -1.18121118 -0.25399561 -1.13368018  1.43334881
  0. -0.56327214]
 [ 1.53741223  1.47342338  1.38816326  0.74182845 -1.28276959 -1.34432839
  0. -0.54124541]
 [ 1.53741223 -0.56972371  1.38816326 -1.38221446 -1.28276959 -0.41843879
  0. -0.50368981]
 [ 0. -2.61287079 -1.28596383 -1.61678023 -0.90544135 -0.72706859
  0. ]
selesai pada 23:09
    
```

Gambar 10. Preprocessing Data

Pada gambar 10 preproses data merupakan proses awal yang penting sebelum melakukan analisis atau pengolahan lebih lanjut. Langkah-langkah preprossing data mencakup membersihkan data dari nilai-nilai yang hilang atau tidak valid, transformasi data ke format yang lebih sesuai, dan standarisasi atau normalisasi nilai-nilai untuk mempersiapkannya agar dapat digunakan dalam model atau analisis selanjutnya secara efektif dan akurat.

4. Metode Klasifikasi *K-Nearest Neighbors*

```

from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'euclidean', p=2)
classifier.fit(x_train, y_train)
    
```

Gambar 11. Metode Klasifikasi *K-Nearest Neighbors*

Pada Gambar 11 klasifikasi dengan menggunakan metode KNN adalah teknik di mana data yang belum diketahui akan diberi label berdasarkan mayoritas label dari tetangga terdekatnya dalam ruang fitur yang telah ditentukan sebelumnya. Metode ini memanfaatkan data pelatihan yang telah ada untuk memprediksi atau mengklasifikasikan data baru, dengan memperhatikan pola yang terlihat dari data yang telah ada sebelumnya.

5. Pengujian *Confusion matrix*

```

[ ] y_pred = classifier.predict(x_test)

from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)

[[13  4]
 [12  6]]
    
```

Gambar 12. Pengujian *Confusion matrix*

Pada Gambar 12 Confusion matrix adalah sebuah metode evaluasi yang digunakan untuk mengevaluasi kinerja model klasifikasi, termasuk dalam konteks KNN (*K-Nearest Neighbors*). Confusion matrix memberikan gambaran tentang seberapa baik model mampu memprediksi kelas-kelas target dengan membandingkan prediksi yang dihasilkan dengan nilai sebenarnya dari data pengujian. ini membantu untuk mengidentifikasi jumlah prediksi yang benar dan salah untuk setiap kelas,

seperti true positives, true negatives, false positives, dan false negatives. dengan demikian, Confusion matrix sangat penting dalam mengevaluasi akurasi dan keandalan model KNN dalam melakukan klasifikasi data.

6. Hasil Akurasi Metode KNN

Criterion	Value
accuracy	51.85%

	true Cemar Berat	true Cemar Ringan	false Cemar Ringan	class precision
pred. Cemar Berat	6	7	0	46.15%
pred. Cemar Ringan	6	8	0	57.14%
pred. Cemar Ringan	0	0	0	0.00%
class recall	60.00%	63.33%	0.00%	

Gambar 13. Hasil *labelling*

Pada Gambar 13. menampilkan hasil akurasi menggunakan algoritma *K-Nearest Neighbors* (KNN) pada rapidminer untuk klasifikasi air sungai. Klasifikasi ini menggunakan delapan parameter, yaitu TSS, BOD, COD, PO4, DO, E. coli, pH, dan NO3-N, dan menghasilkan akurasi sebesar 51,85%.

7. Hasil Akurasi *Confusion Matrix*

```

[ ] y_pred = classifier.predict(x_test)

from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)

[[13  4]
 [12  6]]
    
```

Gambar 14. Hasil Akurasi *Confusion Matrix*

Pada Gambar 14 ditampilkan hasil akurasi metode *K-Nearest Neighbors* (KNN) menggunakan confusion matrix, yang menghasilkan akurasi sebesar 54,28%. Nilai ini diperoleh dengan menambahkan jumlah true positive (13) dan true negative (6) yang totalnya menjadi 19, kemudian dibagi dengan jumlah dataset yaitu 35, dan dikalikan 100%, sehingga diperoleh hasil akurasi 54,28%.

IV. KESIMPULAN

Berdasarkan hasil dan pembahasan terhadap sistem penulis dapat menyimpulkan beberapa hal sebagai berikut : 1. Aplikasi Klasifikasi Kualitas Air Sungai berhasil dikembangkan untuk memberikan informasi kualitas air sungai di Aceh Utara. 2. Aplikasi ini berhasil menggunakan metode *K-Nearest Neighbors* (KNN) untuk menentukan tingkat pencemaran air sungai berdasarkan parameter TSS, BOD, COD, PO4, DO, E-COLI, pH, dan NO3-N. 3. Aplikasi ini menghasilkan output klasifikasi dengan kategori Cemar Ringan, Cemar Sedang, dan Cemar Berat. Evaluasi model menunjukkan bahwa metode KNN memiliki akurasi sebesar 51% dan Confusion matrix 54%. Hasil K pada KNN diambil dari K5. Pengujian dilakukan dengan menggunakan data sampel air sungai di Aceh Utara. 4. Pengujian sistem yang

dilakukan dengan menggunakan metode black box menunjukkan bahwa fungsi sistem dapat berjalan dengan baik.

REFERENSI

- [1] I. Irfannur and K. Khairan, "Analisis parameter fisika kimia kualitas perairan di Sungai Krueng Mane Aceh Utara," *Arwana J. Ilm. Progr. Stud. Perair.*, vol. 3, no. 1, pp. 16–23, 2021, doi: 10.51179/jipsbp.v3i1.450.
- [2] F. S. Pamungkas, B. D. Prasetya, and I. Kharisudin, "Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python," *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 689–694, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/37875>
- [3] M. Melinda, R. I. Borman, and E. R. Susanto, "Rancang Bangun Sistem Informasi Publik Berbasis Web (Studi Kasus: Desa Durian Kecamatan Padang Cermin Kabupaten Pesawaran)," *J. Tekno Kompak*, vol. 11, no. 1, p. 1, 2018, doi: 10.33365/jtk.v11i1.63.
- [4] C. B. Santoso, "Analisis dan Perancangan Sistem Informasi Produksi," *Teknois J. Ilm. Teknol. Inf. dan Sains*, vol. 6, no. 2, pp. 73–83, 2019, doi: 10.36350/jbs.v6i2.42.
- [5] A. D. Palit, "Mengenali Rambu Lalu Lintas Menggunakan Metode Hog Dan KNN," *J. Ilm. Teknol. Infomasi Terap.*, vol. 3, no. 2, 2017, doi: 10.33197/jitter.vol3.iss2.2017.128.
- [6] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijeit.v5i1.7951.
- [7] S. Monitoring and D. A. N. Peringatan, "Ketinggian Air Berbasis Web Dan Sms Gateway," vol. 5, no. 2, pp. 119–129.
- [8] Yanto et al, "What Makes Gen Z in Indonesia Use P2P Lending Applications: An Extension of Technology Acceptance Model," *Journal Of Information Systems*, vol. 20, issue 1, ISSN: 2088-7043,
- [9] E. Sutanta, "Sistem Informasi Manajemen (eBook)," vol. 1, no. 1, p. xvi+320, 2009, [Online]. Available: <http://grahailmu.co.id/>
- [10] W. Yustanti, "Nihru Nafi' Dzikrulloh1, Indriati2, Budi Darma Setiawan3 2017," *J. Mat. Stat. dan komputasi*, vol. 9, no. 1, pp. 57–68, 2012.
- [11] M. A. Rahman, N. Hidayat, and A. Afif Supianto, "Komparasi Metode Data Mining K-Nearest Neighbor Dengan Naïve Bayes Untuk Klasifikasi Kualitas Air Bersih (Studi Kasus PDAM Tirta Kencana Kabupaten Jombang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.* Vol. 2, No. 12, Desember 2018, hlm. 6346-6353 e-ISSN, vol. 2, no. 12, pp. 925–928, 2018, [Online]. Available: 93 <http://j-ptiik.ub.ac.id>
- [12] A. E. Wijaya and R. B. S. Sukarni, "Sistem Monitoring Kualitas Air Mineral Berbasis Iot (Internet of Things) Menggunakan Platform Node-Red Dan Metode Saw (Simple Additive Weighting)," *J. Teknol. dan Komun. STMIK Subang*, vol. 12, no. 2, pp. 96–106, 2019, doi: 10.47561/a.v12i2.156.
- [13] K. Reza, N. Marlina, and S. Rahmawati, "Penentuan Status Mutu Air Sungai Winongo Pada Parameter Fosfat, Nitrat, dan Amonia Menggunakan Metode Storet, Indeks Pencemaran, CCMEWQI dan BCWQI.," *J. Sains & Teknologi ...*, vol. 15, pp. 1–17, 2023, [Online]. Available: <https://journal.uui.ac.id/JSTL/article/view/22286%0Ah>
<https://journal.uui.ac.id/JSTL/article/download/22286/14829>