

SISTEM IOT DAN ALGORITMA NAIVE BAYES UNTUK MONITORING KUALITAS AIR LAUT SECARA REAL-TIME

Mulyadi*

Haikal Murtaza

Atthariq Huzaifah

Zulfan Khairil Simbolon

Program Studi Teknik Informatika,

Politeknik Negeri Lhokseumawe

Jl. Banda Aceh-Medan Km. 280, Buketrata-Lhokseumawe, Aceh, 24375 INDONESIA

Abstract

This study develops a real-time seawater quality monitoring system by integrating Internet of Things (IoT) technology and the Naive Bayes classification algorithm. The system utilizes an ESP32 microcontroller equipped with pH and turbidity sensors to collect seawater quality data from coastal areas. The sensor data are transmitted continuously to a Firebase Realtime Database and visualized through an Android-based mobile application developed using Flutter. The Naive Bayes model was trained using a public Kaggle dataset consisting of 5,569 water quality records classified into polluted and unpolluted categories, and then applied to classify real-time sensor data collected at Pantai Pusong, Aceh. Data preprocessing included cleaning, normalization, outlier handling, and class balancing. Experimental results show that the proposed system is capable of performing stable real-time data acquisition, transmission, and visualization. The Naive Bayes classification achieved an accuracy of 53%, with better performance in identifying unpolluted seawater conditions than polluted ones. Although the classification accuracy is moderate, the results demonstrate that the integration of IoT devices, cloud services, and probabilistic machine learning can provide a functional end-to-end solution for real-time seawater quality monitoring and early environmental awareness in coastal regions.

Keywords:

Internet of things; naive bayes; real-time monitoring; seawater quality; turbidity

Abstrak

Penelitian ini mengembangkan sistem monitoring kualitas air laut secara real-time dengan mengintegrasikan teknologi Internet of Things (IoT) dan algoritma klasifikasi Naive Bayes. Sistem menggunakan mikrokontroler ESP32 yang dilengkapi sensor pH dan turbidity untuk mengakuisisi data kualitas air laut di wilayah pesisir. Data sensor dikirimkan secara kontinu ke Firebase Realtime Database dan ditampilkan melalui aplikasi Android berbasis Flutter. Model Naive Bayes dilatih menggunakan dataset publik dari Kaggle yang terdiri dari 5.569 data kualitas air dengan dua kelas, yaitu air tercemar dan tidak tercemar, kemudian diterapkan untuk mengklasifikasikan data sensor real-time yang diperoleh dari Pantai Pusong, Aceh. Tahapan praproses data meliputi pembersihan data, normalisasi, penanganan outlier, dan penyeimbangan kelas. Hasil pengujian menunjukkan bahwa sistem mampu melakukan akuisisi, pengiriman, dan visualisasi data secara real-time dengan kinerja yang stabil. Evaluasi menggunakan confusion matrix menghasilkan tingkat akurasi klasifikasi sebesar 53%, dengan kemampuan yang lebih baik dalam mengidentifikasi kondisi air tidak tercemar dibandingkan air tercemar. Meskipun akurasi model masih terbatas, sistem yang dikembangkan menunjukkan keberhasilan integrasi IoT, layanan cloud, dan pembelajaran mesin sebagai solusi awal pemantauan kualitas air laut secara real-time di wilayah pesisir.

Kata Kunci:

Internet of things; kualitas air laut; naive bayes; pemantauan real-time; turbidity

DOI: [10.38038/vocatech.v7i2.291](https://doi.org/10.38038/vocatech.v7i2.291)

Received: 5 Oktober 2025; Accepted: 22 Novermber 2025; Published: 30 Desember 2025

*Corresponding author:

Mulyadi, Politeknik Negeri Lhokseumawe, Jl. Banda Aceh-Medan, Buketrata-Lhokseumawe, Aceh, Indonesia
Email: mulyadi@pnl.ac.id

Citation in APA Style: Mulyadi, Zulfan Khairil,S & Haikal Murtaza (2025). Sistem IoT dan algoritma Naive Bayes untuk monitoring kualitas air laut secara real-time. *VOCATECH: Vocational Education and Technology Journal*, 7(2), 385-396.

1. PENDAHULUAN

Kualitas air laut merupakan salah satu fondasi utama dalam menjaga keseimbangan ekosistem pesisir serta keberlanjutan sumber daya hayati yang menopang kehidupan masyarakat pesisir. Perairan laut memiliki peran strategis sebagai habitat berbagai organisme laut, penopang sektor perikanan dan pariwisata, jalur transportasi, serta sebagai bagian dari sistem penyangga lingkungan global. Kondisi kualitas air yang baik menjadi prasyarat utama bagi keberlangsungan fungsi-fungsi tersebut. Namun, dalam beberapa dekade terakhir, berbagai laporan internasional menunjukkan adanya penurunan kualitas perairan laut secara global yang dipicu oleh peningkatan aktivitas manusia di wilayah pesisir. United Nations Environment Programme ([UNEP, 2023](#)) dan Intergovernmental Panel on Climate Change ([IPCC, 2023](#)) melaporkan bahwa tekanan antropogenik seperti urbanisasi pesisir, pembuangan limbah domestik dan industri, peningkatan sedimentasi, serta polusi mikroplastik telah mempercepat degradasi fisik dan kimia perairan laut. Kondisi ini tidak hanya berdampak pada ekosistem laut, tetapi juga menimbulkan konsekuensi sosial dan ekonomi yang signifikan, terutama bagi masyarakat yang bergantung langsung pada sumber daya pesisir.

Indonesia sebagai negara kepulauan menghadapi tantangan yang lebih kompleks dalam menjaga kualitas lingkungan pesisirnya. Dengan garis pantai lebih dari 108.000 km, Indonesia memiliki wilayah pesisir yang luas dan beragam dengan karakteristik lingkungan serta tekanan aktivitas manusia yang berbeda-beda ([BPS, 2024](#)). Banyak kawasan pesisir mengalami tekanan ekologis akibat aktivitas perikanan tradisional, mobilitas kapal, pertumbuhan permukiman, dan pengelolaan limbah yang belum optimal. Kondisi ini meningkatkan risiko terjadinya penurunan kualitas air laut yang sulit terdeteksi secara dini apabila tidak didukung oleh sistem pemantauan yang memadai. Di berbagai wilayah pesisir, pemantauan kualitas air masih didominasi oleh metode konvensional berupa pengambilan sampel air secara periodik dan analisis laboratorium. Meskipun metode ini memiliki tingkat akurasi yang baik, keterbatasan frekuensi pengukuran, kebutuhan biaya operasional yang tinggi, serta ketergantungan pada tenaga ahli menjadikannya kurang efektif untuk memantau dinamika perairan yang berubah dengan cepat.

Dalam konteks pemantauan kualitas air laut, parameter fisik dan kimia tertentu telah lama digunakan sebagai indikator awal untuk menilai kondisi perairan. Parameter pH dan turbidity (kekeruhan) merupakan dua indikator utama yang sensitif terhadap perubahan lingkungan. Nilai pH air laut normal berada pada kisaran 7,5 hingga 8,4, dan penyimpangan dari rentang tersebut dapat mengindikasikan adanya kontaminasi bahan kimia, perubahan aktivitas biologis, atau masuknya limbah tertentu ke perairan laut ([Setiawan et al., 2022](#)). Sementara itu, turbidity mencerminkan tingkat kejernihan air yang dipengaruhi oleh keberadaan partikel tersuspensi seperti sedimen, plankton, lumpur, dan limbah organik. Peningkatan turbidity dapat menghambat penetrasi cahaya matahari ke dalam kolom air, mengganggu proses fotosintesis fitoplankton, serta berdampak pada stabilitas rantai makanan laut. Oleh karena itu, pemantauan berkelanjutan terhadap pH dan turbidity menjadi penting sebagai langkah awal dalam mendeteksi potensi pencemaran dan perubahan kualitas perairan pesisir.

Keterbatasan pendekatan konvensional dalam menyediakan data kualitas air secara kontinu dan real-time mendorong pemanfaatan teknologi Internet of Things (IoT) sebagai solusi alternatif yang lebih adaptif. IoT memungkinkan integrasi sensor lingkungan dengan perangkat komputasi dan jaringan komunikasi, sehingga data dapat dikumpulkan secara otomatis dan dikirimkan secara real-time ke platform penyimpanan berbasis cloud. Penerapan IoT dalam pemantauan lingkungan telah menunjukkan peningkatan efisiensi, akurasi, serta kemampuan pemantauan jarak jauh dibandingkan metode manual ([Domínguez-Bolaño et al., 2022](#)). Mikrokontroler seperti ESP32 banyak digunakan dalam sistem monitoring lingkungan karena memiliki konektivitas nirkabel yang stabil, konsumsi daya rendah, serta fleksibilitas integrasi dengan berbagai jenis sensor ([Yusri et al., 2024](#)). Dengan dukungan IoT, sistem monitoring kualitas air laut dapat beroperasi secara berkelanjutan tanpa ketergantungan pada intervensi manusia secara langsung.

Seiring meningkatnya volume data yang dihasilkan oleh sensor IoT, diperlukan metode analisis yang mampu mengolah data tersebut secara cepat dan akurat agar informasi yang dihasilkan dapat dimanfaatkan secara optimal. Dalam hal ini, pendekatan machine learning (ML) menjadi relevan untuk mengklasifikasikan dan menginterpretasikan data kualitas air. Salah satu algoritma klasifikasi yang banyak digunakan untuk data kontinu adalah Gaussian Naive Bayes. Algoritma ini bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen dan mengikuti distribusi normal. Keunggulan Gaussian Naive Bayes terletak pada kesederhanaan model, efisiensi komputasi, serta performa yang stabil meskipun jumlah fitur relatif terbatas ([Zhang et al., 2021](#)). Beberapa penelitian menunjukkan bahwa algoritma ini efektif diterapkan pada berbagai aplikasi lingkungan, termasuk klasifikasi kualitas air dan deteksi pencemaran ([Lestari & Widodo, 2023; Rahman et al., 2024](#)).

Dalam konteks pemantauan kualitas air laut, integrasi IoT dan machine learning memberikan peluang untuk membangun sistem yang tidak hanya mampu mengumpulkan data secara real-time, tetapi juga memberikan analisis prediktif mengenai kondisi perairan. Meskipun demikian, sebagian besar penelitian

sebelumnya masih berfokus pada lingkungan air tawar seperti sungai, waduk, dan tambak, sementara kajian yang secara khusus mengintegrasikan IoT, platform cloud, algoritma machine learning, dan aplikasi mobile untuk lingkungan air laut masih relatif terbatas, khususnya di Indonesia ([Siregar et al., 2022](#); [Munawar et al., 2023](#); [Fadhil et al., 2023](#)). Selain itu, banyak sistem yang dikembangkan belum menyediakan pipeline pemantauan yang utuh mulai dari akuisisi data, pengiriman data real-time, klasifikasi kondisi air, hingga visualisasi informasi yang mudah diakses oleh pengguna.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk merancang dan membangun sistem monitoring kualitas air laut secara real-time yang mengintegrasikan perangkat IoT berbasis ESP32, sensor pH dan turbidity, platform penyimpanan cloud, serta algoritma Gaussian Naive Bayes sebagai model klasifikasi kondisi perairan. Sistem ini juga dilengkapi dengan aplikasi mobile berbasis Flutter yang berfungsi sebagai media visualisasi data dan hasil klasifikasi secara langsung. Pendekatan ini diharapkan mampu menjamin kesenjangan antara kebutuhan pemantauan kualitas air laut yang bersifat cepat dan kontinu dengan keterbatasan sistem konvensional yang masih banyak digunakan di lapangan. Dengan menyediakan data real-time dan analisis berbasis kecerdasan mesin, sistem yang dikembangkan diharapkan dapat mendukung pengambilan keputusan yang lebih responsif dan berbasis data dalam upaya pengelolaan dan perlindungan lingkungan pesisir secara berkelanjutan.

2. STUDI PUSTAKA

2.1 Kualitas Air Laut dan Parameter Pemantauan

Kualitas air laut merupakan indikator penting dalam menilai kesehatan ekosistem pesisir dan keberlanjutan sumber daya laut. Perubahan kualitas air laut dapat memengaruhi keseimbangan biologis, fisik, dan kimia perairan, yang selanjutnya berdampak pada organisme laut serta aktivitas manusia yang bergantung pada ekosistem tersebut. Berbagai studi menunjukkan bahwa degradasi kualitas air laut umumnya dipicu oleh aktivitas antropogenik, seperti pembuangan limbah domestik dan industri, aktivitas pelayaran, sedimentasi berlebih, serta limpasan dari wilayah daratan ([UNEP, 2023](#); [IPCC, 2023](#)).

Parameter fisik dan kimia merupakan komponen utama dalam pemantauan kualitas air laut. Dua parameter yang sering digunakan sebagai indikator awal adalah pH dan turbidity (kekeruhan). Nilai pH mencerminkan tingkat keasaman atau kebasaan air laut yang sensitif terhadap perubahan lingkungan, termasuk pencemaran bahan kimia dan aktivitas biologis. Air laut umumnya memiliki pH pada kisaran 7,5–8,4, dan penyimpangan dari rentang tersebut dapat mengindikasikan adanya gangguan kualitas perairan ([Setiawan et al., 2022](#)). Sementara itu, turbidity menggambarkan tingkat kejernihan air yang dipengaruhi oleh partikel tersuspensi seperti sedimen, plankton, dan limbah organik. Peningkatan nilai turbidity dapat menghambat penetrasi cahaya matahari, menurunkan laju fotosintesis fitoplankton, serta mengganggu rantai makanan laut ([Effendi, 2019](#)).

Pemantauan pH dan turbidity secara berkelanjutan menjadi sangat penting, khususnya di wilayah pesisir yang memiliki dinamika aktivitas manusia yang tinggi. Data yang diperoleh secara kontinu dapat memberikan indikasi awal terhadap potensi pencemaran dan menjadi dasar dalam pengambilan keputusan pengelolaan lingkungan pesisir secara lebih responsif.

2.2 Internet of Things (IoT) dalam Monitoring Lingkungan Perairan

Internet of Things (IoT) merupakan paradigma teknologi yang memungkinkan berbagai perangkat fisik untuk saling terhubung, mengumpulkan data, dan berkomunikasi melalui jaringan internet secara otomatis tanpa intervensi manusia secara langsung. Dalam konteks pemantauan lingkungan, IoT dimanfaatkan untuk mengatasi keterbatasan metode konvensional yang bersifat manual, periodik, dan tidak real-time. Sistem berbasis IoT memungkinkan proses akuisisi data dilakukan secara kontinu serta dapat diakses dari jarak jauh melalui platform berbasis cloud, sehingga mendukung pemantauan lingkungan yang lebih responsif dan efisien ([Al-Fuqaha et al., 2022](#); [Rathi et al., 2023](#)).

Pada sistem pemantauan kualitas air, sensor lingkungan seperti sensor pH dan turbidity diintegrasikan dengan mikrokontroler untuk melakukan pengukuran parameter kualitas air secara periodik. Mikrokontroler ESP32 menjadi salah satu perangkat yang banyak digunakan dalam pengembangan sistem IoT karena memiliki konektivitas Wi-Fi terintegrasi, konsumsi daya yang relatif rendah, serta kemampuan pemrosesan data yang memadai untuk aplikasi monitoring real-time ([Domínguez-Bolaño et al., 2022](#)). Data hasil pengukuran sensor kemudian dikirimkan ke server atau platform cloud untuk disimpan dan dianalisis lebih lanjut, sehingga memungkinkan pemantauan kondisi perairan secara berkelanjutan.

Berbagai penelitian terbaru menunjukkan bahwa penerapan IoT dalam monitoring kualitas air mampu meningkatkan efektivitas pemantauan dibandingkan metode tradisional. Sistem berbasis IoT memungkinkan deteksi perubahan kondisi lingkungan secara lebih cepat dan akurat, sehingga mendukung

upaya deteksi dini pencemaran serta pengambilan keputusan yang lebih tepat waktu ([Kumar et al., 2022](#); [Yusri et al., 2024](#)). Meskipun demikian, sebagian besar implementasi IoT dalam pemantauan kualitas air masih berfokus pada perairan tawar seperti sungai, danau, atau tambak. Penerapan sistem IoT pada lingkungan air laut masih relatif terbatas, terutama di wilayah pesisir Indonesia yang memiliki karakteristik perairan yang lebih dinamis dan kompleks ([Munawar et al., 2023](#)).

2.3 Machine Learning dan Metode Klasifikasi Naive Bayes

Seiring meningkatnya volume data yang dihasilkan oleh sensor IoT, diperlukan metode analisis yang mampu mengolah dan menafsirkan data secara efisien. Machine learning menjadi salah satu pendekatan yang banyak digunakan dalam analisis data lingkungan karena kemampuannya dalam mengenali pola dan melakukan prediksi berdasarkan data historis. Dalam pemantauan kualitas air, machine learning dimanfaatkan untuk mengklasifikasikan kondisi perairan ke dalam kategori tertentu, seperti tercemar dan tidak tercemar ([Zhang et al., 2021](#)).

Salah satu algoritma klasifikasi yang banyak digunakan adalah Naive Bayes, yaitu metode probabilistik yang bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain. Untuk data numerik kontinu, pendekatan Gaussian Naive Bayes digunakan dengan mengasumsikan bahwa data mengikuti distribusi normal. Keunggulan algoritma ini terletak pada kesederhanaan model, efisiensi komputasi, serta kemampuannya bekerja dengan baik pada dataset berukuran besar dengan jumlah fitur terbatas ([Lestari & Widodo, 2023](#)).

Beberapa penelitian menunjukkan bahwa meskipun Naive Bayes memiliki keterbatasan dalam menangkap hubungan kompleks antar fitur, algoritma ini tetap memberikan performa yang stabil dan layak digunakan sebagai solusi awal dalam sistem monitoring berbasis IoT yang menuntut proses klasifikasi cepat dan ringan ([Prasetyo, 2020](#)).

2.4 Integrasi IoT dan Machine Learning dalam Monitoring Kualitas Air

Integrasi antara IoT dan machine learning memungkinkan pengembangan sistem monitoring kualitas air yang tidak hanya mampu mengumpulkan data secara real-time, tetapi juga mengolah dan menginterpretasikan data tersebut menjadi informasi yang lebih bermakna. Dalam sistem terintegrasi, IoT berperan sebagai penyedia data lingkungan, sedangkan machine learning berfungsi sebagai mesin analisis dan klasifikasi kondisi perairan.

Beberapa penelitian terdahulu telah mengusulkan sistem monitoring kualitas air berbasis IoT yang dilengkapi dengan analisis data menggunakan machine learning. Namun, sebagian besar penelitian tersebut masih terbatas pada perairan tawar dan belum mengintegrasikan seluruh komponen sistem secara end-to-end, mulai dari akuisisi data sensor, penyimpanan cloud, proses klasifikasi, hingga visualisasi hasil kepada pengguna ([Siregar et al., 2022](#); [Munawar et al., 2023](#)). Selain itu, karakteristik perairan laut yang lebih dinamis akibat pengaruh pasang surut, arus laut, dan aktivitas pesisir menyebabkan pendekatan yang digunakan pada perairan tawar tidak dapat langsung digeneralisasi. Kajian yang secara spesifik membahas integrasi penuh antara sensor IoT, algoritma klasifikasi probabilistik, dan aplikasi mobile untuk pemantauan kualitas air laut di wilayah pesisir Indonesia masih relatif terbatas.

Berdasarkan telaah pustaka tersebut, penelitian ini diposisikan untuk mengisi kesenjangan dengan mengembangkan sistem monitoring kualitas air laut berbasis IoT yang terintegrasi dengan algoritma Gaussian Naive Bayes serta platform visualisasi berbasis aplikasi mobile. Pendekatan ini diharapkan mampu mendukung pemantauan kualitas air laut secara real-time dan memberikan kontribusi terhadap pengelolaan lingkungan pesisir berbasis data.

3. METODE PENELITIAN

Metode penelitian ini disusun untuk menjelaskan tahapan perancangan, pengembangan, dan pengujian sistem monitoring kualitas air laut berbasis Internet of Things (IoT) yang terintegrasi dengan algoritma klasifikasi Gaussian Naive Bayes. Fokus utama pada bagian ini adalah pemaparan prosedur penelitian, rancangan sistem, serta metode analisis yang digunakan, tanpa memuat hasil pengujian maupun pembahasan kinerja sistem.

3.1 Jenis dan Pendekatan Penelitian

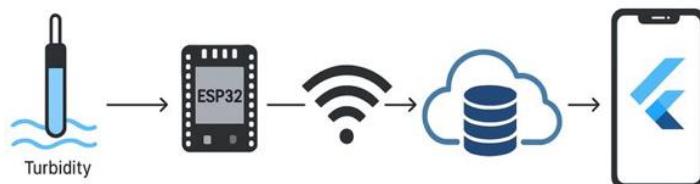
Penelitian ini menggunakan pendekatan kuantitatif dengan metode rekayasa sistem (engineering research). Pendekatan ini bertujuan untuk merancang dan membangun sistem monitoring kualitas air laut secara real-time, serta menyiapkan skema evaluasi untuk mengukur kinerja sistem dan model klasifikasi yang dikembangkan.

3.2 Lokasi dan Objek Penelitian

Penelitian dilaksanakan di kawasan Pantai Pusong, Kota Lhokseumawe, Aceh. Objek penelitian adalah kualitas air laut yang direpresentasikan oleh parameter pH dan turbidity. Kedua parameter tersebut dipilih karena merupakan indikator awal yang sensitif terhadap perubahan kondisi kualitas perairan laut.

3.3 Perancangan Arsitektur Sistem

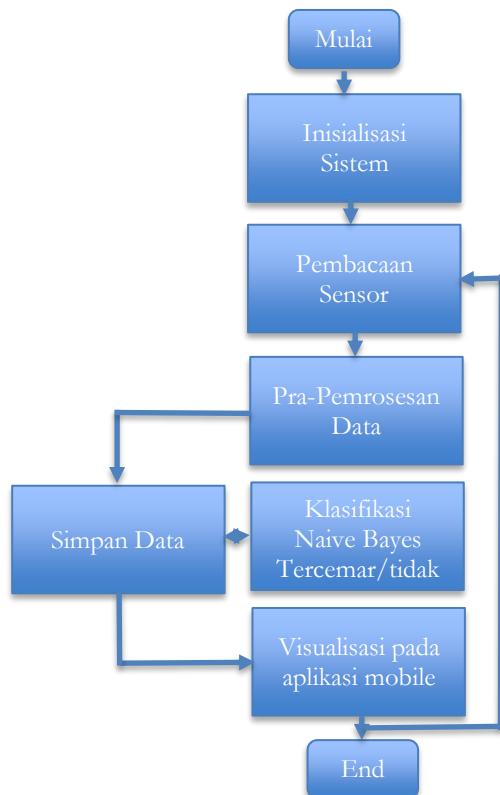
Arsitektur sistem dirancang sebagai sistem terintegrasi yang terdiri dari perangkat IoT, platform cloud, dan aplikasi mobile. Sensor pH dan turbidity terhubung dengan mikrokontroler ESP32 sebagai unit akuisisi data. Data hasil pengukuran dikirimkan melalui jaringan Wi-Fi ke Firebase Realtime Database sebagai media penyimpanan berbasis cloud. Selanjutnya, data tersebut diakses oleh modul klasifikasi dan aplikasi mobile untuk keperluan pemrosesan dan visualisasi, seperti ditunjukkan pada gambar 1.



Gambar 1. Arsitektur Sistem Monitoring Kualitas Air Laut Berbasis IoT dan Naive Bayes

3.4 Alur Kerja Sistem

Alur kerja sistem dimulai dari inisialisasi perangkat dan sensor. Setelah sistem aktif, sensor melakukan pembacaan nilai pH dan turbidity secara periodik. Data hasil pembacaan dikirimkan ke basis data cloud, kemudian diproses oleh model Gaussian Naive Bayes untuk menentukan kelas kualitas air laut, yaitu tercemar atau tidak tercemar. Data sensor dan hasil klasifikasi selanjutnya disimpan kembali ke basis data dan ditampilkan pada aplikasi mobile, seperti ditunjukkan pada gambar 2.



Gambar 2. Diagram Alur Kerja Sistem Monitoring Kualitas Air Laut

3.5 Pengumpulan dan Pra-Pemrosesan Data

Penelitian ini menggunakan dua jenis data, yaitu data sekunder dan data primer. Data sekunder berupa dataset "Water Potability Prediction" dan "Water Probability" yang diperoleh dari Kaggle dengan total 5.569 data, digunakan untuk melatih model Gaussian Naive Bayes. Dataset tersebut

melalui tahapan pra-pemrosesan yang meliputi pembersihan data, penanganan nilai kosong, penghilangan data anomali, normalisasi, dan penyeimbangan kelas, contoh data yang diambil pada Kaggle seperti ditunjukkan pada tabel 1 dan tabel 2.

a. Format Data Dataset *Water Potability Prediction*

Tabel 1. Dataset *Water Potability Prediction*

pH	Hanrdness	Solids	Chloramines	Sulfate	Conductivity	Organic Carbon	Turbidity	Potability
7,08	204,9	20791	7,30	333,8	592,9	10,38	4,63	1
6,52	129,4	18630	6,18	310,1	421,3	14,22	3,12	0
8,21	215,7	22018	8,05	360,5	618,4	9,85	5,01	1

Pada penelitian ini, hanya parameter pH dan turbidity yang diadopsi karena kesesuaian dengan sensor IoT yang digunakan di lapangan.

b. Format Data Dataset *Water Probability*

Tabel 2. Data Dataset *Water Probability*

pH	Turbidity	Temperatur	Conductivity	Water_Quality
7,4	0,45	29,5	580	Tidak Tercemar
6,9	1,12	30,1	620	Tercemar
8,1	0,32	28,9	540	Tidak Tercemar

Dataset ini umumnya digunakan untuk pendekatan probabilistik dan klasifikasi berbasis machine learning. Data primer diperoleh dari hasil pembacaan sensor pH dan turbidity pada sistem IoT yang dibuat. Data primer digunakan sebagai data masukan pada tahap implementasi sistem untuk menguji kemampuan model dalam melakukan klasifikasi secara real-time.

3.6 Metode Klasifikasi Gaussian Naive Bayes

Metode klasifikasi yang digunakan dalam penelitian ini adalah Gaussian Naive Bayes, yang sesuai untuk data numerik kontinu seperti parameter pH dan turbidity. Metode ini didasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur bersifat saling independen dan mengikuti distribusi normal (Gaussian).

Secara umum, Teorema Bayes dirumuskan sebagai berikut:

$$P(Ck|X) = \frac{P(X|Ck) \cdot P(Ck)}{P(X)} \dots \dots \dots \quad (1)$$

dengan:

$P(Ck|X)$ adalah probabilitas posterior kelas Ck terhadap data masukan X,

$P(X|Ck)$ adalah likelihood data X pada kelas Ck,

$P(Ck)$ adalah probabilitas prior kelas Ck,

$P(X)$ adalah probabilitas evidence.

Pada Gaussian Naive Bayes, nilai likelihood untuk setiap fitur dihitung menggunakan fungsi distribusi Gaussian sebagai berikut:

$$P(xi|Ck) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(xi - \mu_k)^2}{2\sigma_k^2}\right) \dots \dots \dots \quad (2)$$

dengan:

xi adalah nilai fitur ke-i,

μ_k adalah nilai rata-rata fitur pada kelas Ck,

σ_k^2 adalah varians fitur pada kelas Ck.

Nilai probabilitas posterior dihitung untuk setiap kelas, dan data diklasifikasikan ke dalam kelas dengan nilai probabilitas posterior tertinggi. Dataset yang digunakan untuk pelatihan model berjumlah 5.569 data dan dibagi menjadi data latih dan data uji dengan perbandingan 75% dan 25%. Model yang telah dilatih kemudian diintegrasikan ke dalam sistem monitoring untuk melakukan klasifikasi otomatis terhadap data sensor pH dan turbidity yang diperoleh secara real-time.

3.7 Metode Evaluasi Sistem

Evaluasi sistem dilakukan untuk menilai kinerja model klasifikasi serta kinerja fungsional sistem monitoring kualitas air laut. Evaluasi model Gaussian Naive Bayes dilakukan menggunakan confusion matrix yang membandingkan kelas aktual dan hasil prediksi model.

Berdasarkan confusion matrix, metrik evaluasi dihitung menggunakan persamaan berikut:

- Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \dots \dots \dots \quad (3)$$

- **Precision**

$$\text{Precision} = \frac{TP}{TP+FP} \quad \dots \quad (4)$$

- Recall

$$\text{Recall} = \frac{TP}{TP+FN} \quad \dots \quad (5)$$

- F1-Score

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \dots \quad (6)$$

Dengan :

- TP (True Positive): data tercemar yang diklasifikasikan sebagai tercemar.
 - TN (True Negative): data tidak tercemar yang diklasifikasikan sebagai tidak tercemar
 - FP (False Positive): data tidak tercemar yang diklasifikasikan sebagai tercemar
 - FN (False Negative): data tercemar yang diklasifikasikan sebagai tidak tercemar

Selain evaluasi model, evaluasi fungsional sistem monitoring dilakukan dengan menguji kestabilan pembacaan sensor pH dan turbidity, keandalan pengiriman data ke basis data cloud Firebase, serta kemampuan aplikasi mobile dalam menampilkan data sensor dan hasil klasifikasi secara real-time.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengujian Sistem Monitoring Kualitas Air Laut

Bagian ini menyajikan hasil implementasi sistem monitoring kualitas air laut berbasis Internet of Things (IoT) dan algoritma Naive Bayes. Implementasi sistem mencakup integrasi perangkat keras, perangkat lunak, serta mekanisme komunikasi data secara real-time antara sensor, platform cloud, dan aplikasi mobile. Evaluasi dilakukan untuk memastikan sistem mampu menjalankan fungsi pengukuran, pengiriman data, dan penyajian informasi kualitas air laut secara berkelanjutan.

4.1.1 Hasil Pembacaan Sensor pH dan Turbidity

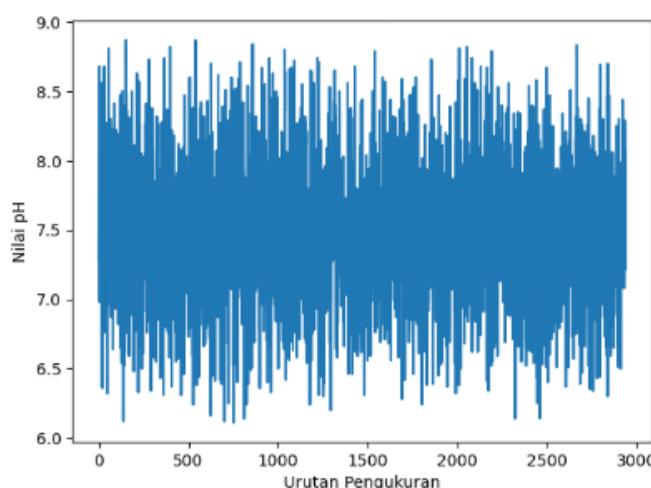
Pengujian sistem monitoring kualitas air laut dilakukan dengan mengumpulkan sebanyak 2.940 data hasil pengukuran sensor, yang diperoleh secara berkelanjutan dan real-time. Data tersebut merupakan hasil pembacaan langsung dari sensor pH dan turbidity yang terintegrasi pada sistem Internet of Things (IoT), sehingga mampu merepresentasikan kondisi aktual perairan laut selama periode pengamatan.

Berdasarkan data pengukuran yang terkumpul seperti ditunjukkan pada table 3, nilai pH air laut berada pada rentang 6,98 hingga 8,68, dengan nilai rata-rata sebesar 7,41. Rentang nilai ini menunjukkan variasi tingkat keasaman air laut dari kondisi mendekati netral hingga sedikit basa. Secara umum, nilai pH tersebut masih berada dalam kisaran yang lazim dijumpai pada perairan pesisir, namun fluktuasi yang terjadi dapat mengindikasikan adanya perubahan kondisi lingkungan perairan akibat faktor alami maupun aktivitas manusia.

Tabel 3. Statistik Deskriptif Data Sensor pH dan Turbidity

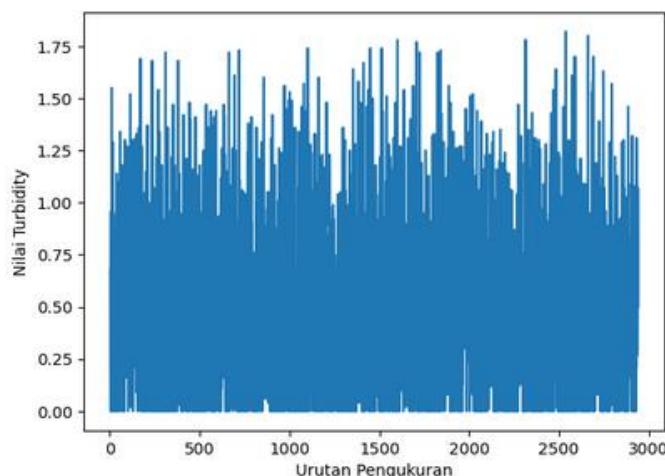
Parameter	Nilai Minimum	Nilai Maksimum	Rata-rata
pH	6,98	8,68	7,41
Turbidity	0,00	0,71	0,39

Nilai turbidity yang terukur berada pada rentang 0,00–0,71 NTU dengan nilai rata-rata 0,39 NTU, yang menunjukkan variasi tingkat kekeruhan air laut akibat keberadaan partikel tersuspensi. Variasi ini menjadi indikator penting dalam menilai kualitas perairan laut.



Gambar 3. Grafik Perubahan Nilai pH Air Laut terhadap Urutan Pengukuran

Gambar 3 memperlihatkan perubahan nilai pH air laut terhadap urutan pengukuran. Grafik menunjukkan fluktuasi nilai pH yang relatif stabil selama periode pengamatan, yang mengindikasikan bahwa sistem monitoring berbasis IoT mampu merekam dinamika kondisi perairan secara real-time dengan baik.



Gambar 4. Grafik Perubahan Nilai Turbidity Air Laut terhadap Waktu

Gambar 4 dimana grafik menunjukkan bahwa nilai turbidity mengalami fluktuasi yang relatif lebih dinamis dibandingkan parameter pH. Variasi ini mencerminkan perubahan tingkat kekeruhan air laut akibat keberadaan partikel tersuspensi di dalam perairan, yang dapat dipengaruhi oleh faktor lingkungan seperti arus, aktivitas pesisir, maupun kondisi cuaca. Meskipun terjadi variasi antar pengukuran, nilai turbidity secara umum masih berada pada kisaran yang terukur secara konsisten oleh sensor.

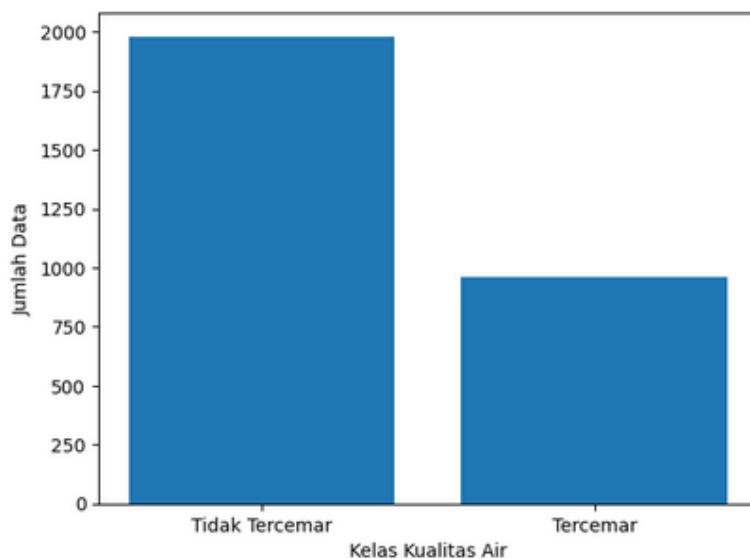
4.1.2 Hasil Klasifikasi Kualitas Air Laut Menggunakan Metode Gaussian Naive Bayes

Hasil klasifikasi kualitas air laut diperoleh dari pemrosesan data hasil pengukuran sensor pH dan turbidity dimana 2.940 data pengukuran digunakan sebagai masukan pada tahap klasifikasi menggunakan metode Gaussian Naive Bayes, dengan dua kelas keluaran, yaitu tercemar dan tidak tercemar. Distribusi hasil klasifikasi ini menunjukkan bahwa selama periode pengamatan, kondisi kualitas air laut di lokasi penelitian cenderung berada pada kategori tidak tercemar, meskipun pada beberapa titik pengukuran terdeteksi kondisi yang mengindikasikan adanya penurunan kualitas air. Tabel 4 menyajikan ringkasan jumlah data hasil klasifikasi kualitas air laut berdasarkan masing-masing kelas.

Tabel 4. Hasil Klasifikasi Kualitas Air Laut Menggunakan Gaussian Naive Bayes

Kelas Kualitas Air	Jumlah Data	Persentase (%)
Tidak Tercemar	1,980	67,35
Tercemar	960	32,65
Total	2,940	100

Gambar 5 menunjukkan distribusi hasil klasifikasi kualitas air laut berdasarkan seluruh data pengukuran sensor yang dikumpulkan selama pengujian sistem. Grafik memperlihatkan bahwa jumlah data yang diklasifikasikan sebagai tidak tercemar lebih dominan dibandingkan dengan data yang diklasifikasikan sebagai tercemar.



Gambar 5. Grafik Distribusi Air Laut

Distribusi hasil klasifikasi ini mencerminkan kemampuan sistem dalam membedakan kondisi kualitas air laut berdasarkan variasi parameter yang terukur. Dominasi kelas tidak tercemar juga menunjukkan bahwa model klasifikasi memiliki kecenderungan mengenali kondisi normal perairan dengan lebih stabil, sedangkan kemunculan kelas tercemar menandakan sensitivitas sistem terhadap perubahan kondisi lingkungan yang berpotensi memengaruhi kualitas air laut.

4.2 Evaluasi Kinerja Model Klasifikasi

Evaluasi kinerja model klasifikasi dilakukan untuk menilai kemampuan algoritma Gaussian Naive Bayes dalam mengklasifikasikan kualitas air laut ke dalam kelas tercemar dan tidak tercemar berdasarkan parameter pH dan turbidity yang diperoleh dari sistem monitoring berbasis IoT. Evaluasi ini bertujuan untuk memastikan bahwa model mampu memberikan hasil klasifikasi yang konsisten dan relevan ketika diterapkan pada data hasil pengukuran lapangan.

Pengujian model dilakukan menggunakan data uji yang telah melalui tahap pra-pemrosesan, kemudian dianalisis menggunakan metrik evaluasi berupa *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. Selain evaluasi numerik, hasil klasifikasi juga ditampilkan pada aplikasi mobile sebagai bentuk validasi fungsional integrasi antara model klasifikasi dan sistem monitoring kualitas air laut secara real-time.

4.2.1 Confusion matrix Hasil Klasifikasi

Confusion matrix digunakan untuk mengevaluasi kinerja model Gaussian Naive Bayes dalam mengklasifikasikan kualitas air laut ke dalam dua kelas, yaitu tercemar dan tidak tercemar. Matriks ini menyajikan perbandingan antara kelas aktual dan hasil prediksi model, sehingga dapat mengidentifikasi tingkat ketepatan serta kesalahan klasifikasi yang terjadi.

Tabel 5. Confusion Matrix Hasil Klasifikasi Kualitas Air Laut

Kelas Aktual \ Prediksi	Tercemar	Tidak Tercemar	Total
Tercemar	TP = 320	FN = 640	960
Tidak Tercemar	FP = 742	TN = 1238	1980
Total	1062	1878	2940

Berdasarkan Tabel 5, dari total 2.940 data hasil pengukuran sensor, model Gaussian Naive Bayes mampu mengklasifikasikan 1.558 data dengan benar. Model menunjukkan performa yang lebih baik dalam mengenali kelas tidak tercemar, yang tercermin dari nilai True Negative yang lebih tinggi dibandingkan True Positive. Nilai False Negative yang masih cukup besar menunjukkan bahwa sebagian data tercemar belum

berhasil dikenali secara optimal, yang mengindikasikan keterbatasan model ketika hanya menggunakan parameter pH dan turbidity sebagai fitur input.

4.2.2 Nilai Akurasi, Precision, Recall, dan F1-Score

Evaluasi kinerja model Gaussian Naive Bayes dilakukan menggunakan data hasil pengukuran sensor sebanyak 2.940 data, yang terdiri dari 960 data tercemar dan 1.980 data tidak tercemar. Perhitungan metrik evaluasi didasarkan pada confusion matrix yang disajikan pada Tabel 3, sehingga mencerminkan performa aktual model dalam mengklasifikasikan kualitas air laut berdasarkan parameter pH dan turbidity. Berdasarkan nilai True Positive (TP = 320), True Negative (TN = 1.238), False Positive (FP = 742), dan False Negative (FN = 640). Dengan menggunakan persamaan 3, 4, 5, dan persamaan 6, dan berdasarkan perhitungan menggunakan data hasil klasifikasi, diperoleh nilai metrik evaluasi sebagaimana disajikan pada Tabel 6.

Tabel 6. Hasil Evaluasi Kinerja Model Klasifikasi

Metrik Evaluasi	Nilai
Accuracy	53%
Precision	30,1%
Recall	33,3%
F1-Score	31,6%

Hasil evaluasi menunjukkan bahwa model Gaussian Naive Bayes memiliki tingkat akurasi sebesar 53%, yang mengindikasikan bahwa lebih dari setengah data uji berhasil diklasifikasikan dengan benar. Nilai precision dan recall yang lebih rendah pada kelas tercemar menunjukkan bahwa model masih mengalami keterbatasan dalam mendekripsi kondisi pencemaran secara konsisten. Hal ini terutama disebabkan oleh keterbatasan jumlah dan variasi parameter input yang digunakan terlalu rendah yaitu hanya pH dan turbidity, sehingga pola data pencemaran belum sepenuhnya terwakili.

4.3 Pembahasan Hasil Penelitian

Penelitian ini menghasilkan sebuah sistem monitoring kualitas air laut yang dirancang untuk mengklasifikasikan kondisi air laut ke dalam dua kategori, yaitu tercemar dan tidak tercemar. Sistem dikembangkan dengan pendekatan Internet of Things (IoT) yang memanfaatkan perangkat ESP32 sebagai mikrokontroler utama, sensor pH dan sensor turbidity sebagai alat ukur parameter kualitas air, serta aplikasi Android sebagai antarmuka pengguna. Lokasi pengambilan data berada di wilayah perairan Pusong, Lhokseumawe, yang memiliki karakteristik lingkungan pesisir dengan potensi pencemaran akibat aktivitas manusia.

Sebelum sistem diimplementasikan pada lingkungan nyata, metode klasifikasi yang digunakan terlebih dahulu dilatih menggunakan data sekunder. Data sekunder yang digunakan berasal dari dataset "Water Potability Prediction" dan "Water Probability" yang diperoleh dari situs Kaggle, dengan jumlah total data sebanyak 5.569 baris. Dataset tersebut melalui beberapa tahapan prapemrosesan, meliputi pemilihan fitur yang relevan, penghapusan data kosong, penghilangan data anomali, serta penyeimbangan kelas. Hasil prapemrosesan kemudian digunakan untuk melatih model Naive Bayes sehingga diperoleh parameter statistik yang menjadi dasar proses klasifikasi kualitas air pada tahap implementasi sistem.

Pada tahap implementasi, sistem melakukan pengambilan data kualitas air laut secara langsung menggunakan sensor pH dan sensor turbidity yang terhubung dengan ESP32. Proses pengambilan data dilakukan secara periodik hingga diperoleh sebanyak 2.940 data pengukuran. Data hasil pembacaan sensor kemudian dikirimkan ke basis data Firebase secara real-time dan disimpan sebagai data mentah. Mekanisme pengiriman dan penyimpanan data ini memungkinkan sistem untuk mencatat seluruh hasil pengukuran secara terpusat dan berkelanjutan.

Data sensor yang tersimpan pada Firebase selanjutnya diproses dan diklasifikasikan menggunakan metode Naive Bayes yang telah dilatih sebelumnya. Hasil klasifikasi kemudian disimpan kembali ke basis data bersama dengan nilai parameter pH, turbidity, serta informasi waktu pencatatan. Berdasarkan hasil klasifikasi terhadap 2.940 data pengukuran, diperoleh sebanyak 1.238 data terkласifikasi sebagai tidak tercemar dan 320 data terkласifikasi sebagai tercemar. Namun demikian, terdapat 640 data dan 742 data yang tidak dapat dibaca atau diklasifikasikan secara akurat oleh sistem. Kondisi ini menunjukkan bahwa tidak seluruh data hasil pembacaan sensor dapat diproses dengan baik, yang berdampak pada kinerja sistem secara keseluruhan.

Jika ditinjau dari keseluruhan proses pengukuran dan klasifikasi, sistem memiliki tingkat keberhasilan kerja sebesar 53%. Nilai ini menunjukkan bahwa sistem mampu menjalankan fungsi monitoring dan klasifikasi kualitas air laut, namun masih memiliki keterbatasan dari sisi akurasi pembacaan sensor dan kemampuan model dalam mengklasifikasikan seluruh data yang diperoleh. Faktor yang mempengaruhi kondisi tersebut antara lain kemungkinan ketidakstabilan pembacaan sensor pada lingkungan laut, gangguan pada proses akuisisi data, serta keterbatasan model Naive Bayes dalam menangani data dengan karakteristik yang bervariasi.

Hasil pengukuran dan klasifikasi kualitas air laut ditampilkan pada aplikasi Android yang berfungsi sebagai antarmuka pengguna. Aplikasi ini menampilkan nilai parameter pH dan turbidity, hasil klasifikasi kualitas air, serta menyediakan fitur untuk menampilkan data historis atau record data yang telah tersimpan di basis data Firebase. Dengan adanya fitur ini, pengguna dapat memantau kondisi kualitas air laut secara real-time maupun melakukan peninjauan terhadap data pengukuran sebelumnya. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa sistem monitoring kualitas air laut berbasis IoT dengan metode Naive Bayes telah berhasil diimplementasikan dan diuji pada lingkungan nyata. Meskipun tingkat keberhasilan sistem masih berada pada angka 53%, sistem telah mampu menunjukkan integrasi yang baik antara perangkat IoT, basis data cloud, metode klasifikasi, dan aplikasi Android.

5. SIMPULAN

Berdasarkan hasil perancangan, implementasi, serta pengujian sistem yang telah dilakukan, maka dapat disimpulkan beberapa hal sebagai berikut:

Penelitian ini berhasil mengembangkan sistem monitoring kualitas air laut berbasis Internet of Things (IoT) menggunakan ESP32, sensor pH, dan sensor turbidity dengan keluaran klasifikasi kondisi air laut, yaitu tercemar dan tidak tercemar. Metode Naive Bayes dilatih menggunakan data sekunder dari Kaggle sebanyak 5.569 data dan diimplementasikan untuk mengklasifikasikan data hasil pengukuran sensor di wilayah Pusong, Lhokseumawe.

Hasil pengujian menunjukkan bahwa dari 2.940 data pengukuran sensor, sistem mampu mengklasifikasikan 1.238 data sebagai tidak tercemar dan 320 data sebagai tercemar, dengan tingkat keberhasilan kerja sistem sebesar 53%. Data hasil pengukuran dan klasifikasi berhasil disimpan pada basis data Firebase dan ditampilkan pada aplikasi Android, termasuk fitur penampilan data historis.

DAFTAR PUSTAKA

- Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2022). Internet of Things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, 24(1), 234–284. <https://doi.org/10.1109/COMST.2021.3120044>
- Badan Pusat Statistik. (2024). *Statistik sumber daya laut dan pesisir Indonesia 2024*. BPS Republik Indonesia.
- Domínguez-Bolaño, T., Martínez, J. F., Hernández, J. A., & López, L. (2022). An IoT-based environmental monitoring system for coastal water quality assessment. *Sensors*, 22(9), 3351. <https://doi.org/10.3390/s22093351>
- Effendi, H. (2019). *Telaah kualitas air: Bagi pengelolaan sumber daya dan lingkungan perairan*. Kanisius.
- Fadhil, R., Hidayat, R., & Ananda, M. (2023). Sistem monitoring kualitas air berbasis Internet of Things pada wilayah pesisir. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10(2), 215–224. <https://doi.org/10.25126/jtiik.202310221>
- Intergovernmental Panel on Climate Change. (2023). *Climate change 2023: Synthesis report*. IPCC.
- Kumar, P., Mallick, P. K., & Bhoi, A. K. (2022). IoT-based smart water quality monitoring system using machine learning techniques. *Journal of Ambient Intelligence and Humanized Computing*, 13(6), 2841–2856. <https://doi.org/10.1007/s12652-021-03415-6>
- Lestari, D., & Widodo, A. (2023). Klasifikasi kualitas air menggunakan algoritma Gaussian Naive Bayes. *Jurnal Teknologi dan Sistem Komputer*, 11(1), 45–52. <https://doi.org/10.14710/jtsiskom.2023.15011>
- Munawar, A., Siregar, R., & Putra, D. (2023). Implementasi IoT untuk monitoring kualitas perairan pesisir secara real-time. *Jurnal Ilmu Kelautan*, 28(3), 189–198. <https://doi.org/10.14710/ik.2023.189>
- Prasetyo, E. (2020). *Data mining: Konsep dan aplikasi menggunakan MATLAB*. Andi Offset.
- Rahman, A., Saputra, D., & Nabila, R. (2024). Penerapan machine learning untuk klasifikasi kualitas air berbasis sensor. *Jurnal Rekayasa Sistem dan Teknologi Informasi*, 8(1), 33–41. <https://doi.org/10.29207/jrsti.v8i1.221>
- Rathi, S., Jain, P., & Meena, R. (2023). Smart environmental monitoring using IoT and cloud computing. *International Journal of Computer Applications*, 185(10), 1–7.

- Setiawan, A., Nugroho, B., & Hartono, D. (2022). Analisis parameter fisika dan kimia air laut sebagai indikator pencemaran. *Jurnal Ilmu Lingkungan*, 20(2), 314–322. <https://doi.org/10.14710/jil.20.2.314-322>
- Siregar, R., Lubis, M., & Harahap, F. (2022). Sistem monitoring kualitas air berbasis IoT dan cloud computing. *Jurnal Teknologi dan Sistem Informasi*, 9(4), 401–410.
- United Nations Environment Programme. (2023). *Global marine pollution assessment*. UNEP.
- Yusri, M., Anwar, S., & Khalid, M. (2024). Pengembangan sistem IoT berbasis ESP32 untuk monitoring kualitas air. *Jurnal Informatika dan Rekayasa Elektronika*, 7(1), 12–21. <https://doi.org/10.36595/jire.v7i1.511>
- Zhang, L., Wang, Y., & Chen, X. (2021). A review of Naive Bayes classification in environmental monitoring. *Environmental Modelling & Software*, 143, 105093. <https://doi.org/10.1016/j.envsoft.2021.105093>