

Perbandingan Metode K-NN, SVM dan Decision Tree Pada Klasifikasi Kondisi Air Kolam Bioflok Nongsa

**Heru Wijanarko^{1*}, Adlian Jefiza¹, Muhammad Al Fajri¹, Andreas Prasetyo Dowo
Ruron¹, dan Firman Dwi Syah Putra¹**

¹Politeknik Negeri Batam, Teknik Mekatronika, Batam, Indonesia

**Email: wijanarko@polibatam.ac.id*

Received: xxxxx

Revised: xxxxx

Accepted: xxxxx

Abstrak

Kolam bioflok merupakan salah satu teknologi budidaya perikanan yang saat ini sedang berkembang karena dapat memberikan beberapa keuntungan bagi pembudidaya ikan. Dalam budidaya ikan, fungsi air memegang peranan penting dalam keberlangsungan hidup ikan, sehingga kondisi air kolam, termasuk pH, amonia, dan larutan oksigen, harus dipantau. Namun, para pembudidaya ikan saat ini memiliki tantangan untuk memantau kualitas air pada saat itu. Hal tersebut membuat sulitnya menentukan kualitas air kolam bioflok. Jika ketiga hal tersebut tidak dijaga dengan baik, banyak bakteri dan virus yang dapat mengancam kestabilan kondisi air dan kesehatan ikan itu sendiri. Oleh karena itu, pemantauan kondisi air menjadi sangat penting, dan penelitian ini berfokus pada pembuatan alat monitoring untuk mengetahui kadar pH, amonia, dan oksigen pada air kolam. Selain membuat alat monitoring, penelitian ini juga mengklasifikasikan kondisi air kolam. Metode pengklasifikasian menggunakan tiga algoritma: *K-Nearest Neighbors*, *Support Vector Machine*, dan *Decision Tree*. Berdasarkan ketiga metode klasifikasi tersebut, didapatkan hasil performa sebesar 97% untuk KNN, 92% untuk SVM, dan 99,25% untuk *Decision Tree*. Dari perbandingan performa ketiga metode tersebut, didapatkan hasil bahwa metode *Decision Tree* adalah yang terbaik dibandingkan dengan metode KNN dan SVM.

Kata kunci: *Bioflok, Decision Tree, K-Nearest Neighbor, Support Vector Machine*

Abstract

Biofloc pond is a technology of aquaculture that is currently developing because it can provide several benefits for fish farmers. In fish farming, the function of water plays a vital role in the sustainability of fish, so the conditions of the pond, including pH, ammonia, and oxygen solution, must be monitored. However, fish farmers currently have a challenge for monitoring the water quality at that time. That makes it difficult to determine the water quality of biofloc ponds. If these three things are not closely maintained, many bacteria and viruses can threaten the stability of water conditions and the health of the fish themselves. Therefore, monitoring water conditions is very important, and this research focuses on creating a monitoring tool to determine the levels of pH, ammonia, and oxygen solution in the water ponds. In addition to making monitoring tools, this research also classifies the condition of water ponds. The classification method uses three algorithms: K-Nearest Neighbors, Support Vector Machine, and Decision Tree. Based on the three classification methods, the performance is 97% for KNN, 92% for SVM, and 99,25% for Decision Tree. From the comparison of the performance of the three methods, the results show that the Decision Tree method is the best compared to the KNN and SVM methods.

Keywords: *Biofloc, Decision Tree, K-Nearest Neighbor, Support Vector Machine*

Pendahuluan

Teknologi Bioflok merupakan alternatif teknologi untuk kegiatan budidaya ikan air tawar, terutama ikan nila. Ikan nila merupakan salah satu ikan yang mempunyai daya toleransi terhadap perubahan lingkungan yang ada di sekitarnya. Ikan nila juga merupakan pemakan organisme-organisme kecil yang sesuai dengan konsep bioflok itu sendiri. Namun pada penerapan budidaya ikan nila menggunakan kolam bioflok masih ditemukan beberapa permasalahan seperti para pembudidaya ikan tidak mengetahui secara detail kondisi air pada kolam bioflok. Hal ini tentu sangat berpengaruh terhadap kestabilan kondisi air dan ikan nila yang dibudidayakan menggunakan teknologi bioflok.

Untuk mengetahui kondisi air ada beberapa parameter yang perlu diketahui, parameter-parameter kondisi air seperti pH, amonia dan larutan oksigen berkaitan dengan kualitas dan kondisi air [1]. Monitoring kondisi air selalu dilakukan sehingga air memiliki kondisi yang baik untuk melakukan budidaya.

Selain mengukur kondisi air, hal lain yang tentu penting adalah mengklasifikasikan kondisi air di dalam kolam. Melakukan klasifikasi berfungsi untuk mengetahui kondisi air seperti apa yang baik untuk budidaya ikan nila. Penelitian ini menggunakan algoritma K-NN (*K-Nearest Neighbor*), SVM (*Support Vector Machine*), dan *Decision Tree*. Sehingga, pembudidaya dapat menjaga kualitas dari kondisi air dan meningkatkan efisiensi dalam pemantauan kualitas air.

Berdasarkan dari permasalahan di mana pembudidaya kesulitan mengetahui kadar pH, amonia serta larutan oksigen, maka penelitian ini mengukur tiga parameter tersebut dan penelitian ini akan menerapkan tiga algoritma yaitu K-NN, SVM, dan *Decision Tree* untuk mengklasifikasikan kondisi air pada kolam bioflok di daerah Nongsa yang diharapkan dapat memberikan performansi model klasifikasi yang baik.

Metode Penelitian

Kondisi air merupakan salah satu faktor penting dalam kegiatan budidaya yang berkaitan dengan produktivitas ikan menggunakan teknologi bioflok. Parameter kondisi air seperti pH, amonia, dan DO memiliki keterkaitan dengan kondisi air [1].

Kondisi air sangat mempengaruhi kelangsungan hidup ikan nila yang dibudidayakan. Dalam pembudidayaan ikan nila

pH air yang baik memiliki nilai antara 6 - 8,5 [2]. Selain itu, menurut baku standar [3], kadar amonia tidak boleh melebihi 0,02 mg/L. Dan untuk kadar oksigen minimum terlarut yang baik bagi ikan nila adalah 3 mg/L [4].

Dari nilai standar ketiga parameter di atas dilakukan kompilasi untuk meningkatkan efektivitas dari sistem bioflok dalam memonitoring kondisi air dalam budidaya ikan nila untuk dijadikan landasan dalam pemantauan kualitas air. Penelitian ini dilakukan pada air kolam bioflok pembudidaya ikan nila di Nongsa. Parameter kualitas air yang diukur di antaranya pH, amonia, dan larutan oksigen yang dikompilasi menjadi suatu tabel untuk menentukan status kelas dari ketiga parameter

Tabel 1. Kadar Kondisi Air

pH	DO	Amonia	kelas
>=6 - 8	>=3	<0,02	Baik
<6 & >8	<3	>=0,02	Buruk
<6 & >8	>=3	>=0,02	Buruk
>=6 - 8	<3	>=0,02	Buruk
<6 & >8	<3	<=0,02	Buruk
>=6 - 8	>=3	>0,02	Sedang
<6 & >8	>=3	<0,02	Sedang
>=6 - 8	<3	<0,02	Sedang

Pada penelitian ini dibangun alat monitoring dan model klasifikasi kondisi air pada kolam bioflok yang mengukur parameter pH, amonia dan larutan oksigen. Untuk membangun penelitian ini, peneliti telah membuat perancangan *hardware*, metode klasifikasi dan pengujian alat dan metode klasifikasi.

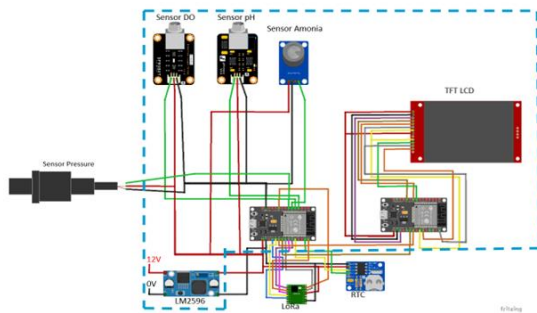
A. Perancangan Hardware

Perancangan ini memuat dua aspek yaitu perancangan mekanikal dan perancangan

elektrikal. Pada perancangan mekanikal membuat wadah sebagai tempat komponen elektrik dan sensor - sensor yang digunakan pada penelitian ini seperti, ESP32, TFT LCD, sensor pH (KIT E-201C), sensor amonia (MQ-137) dan sensor *dissolved oxygen*.



Gambar 1. Desain Mekanikal



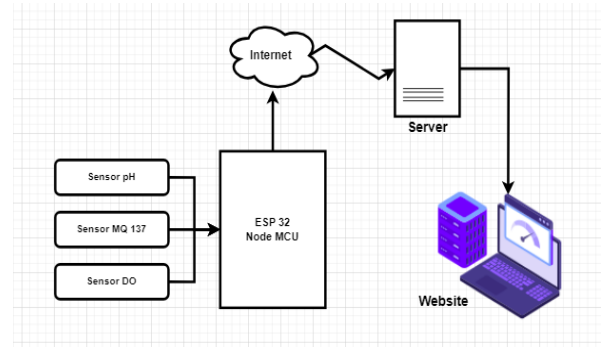
Gambar 2. Rangkaian Elektrikal

NodeMCU ESP32 merupakan mikrokontroler yang berbasis ESP32 yang merupakan *upgrade* dari ESP8266 [5]. ESP32 digunakan agar dapat membaca data dari sensor dan mengirimkan data ke LCD dan *website*. Sensor MQ-137 adalah sensor senyawa kimia yang dapat mendeteksi kadar resistansi tinggi pada amonia, karbon monoksida dan etanol. Resistansi sensor ini berubah ketika terdapat gas amonia yang terdeteksi [6]. Kit E-201C merupakan salah satu jenis sensor pH. *Probe* pH berfungsi mengukur tingkat keasaman dengan skala 0 - 14, di mana skala 0 menjadi tingkat paling asam dan 14 menjadi tingkat paling basa [7]. Sensor Oksigen terlarut (*dissolved oxygen*) atau dikenal juga sebagai sensor DO digunakan sebagai istilah atau ungkapan proses pengukuran jumlah konsentrasi oksigen yang terlarut di dalam satuan unit volume air [8].

B. Website Monitoring data

Pada perancangan *website* berisi proses pengiriman data dari sensor-sensor agar tertampil di dalam *website*. berikut merupakan

topologi *website* dalam pengiriman data-data sensor.



Gambar 3. Topologi Website

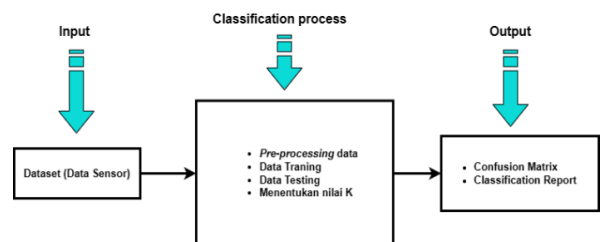
Pada gambar di atas menjelaskan bahwa sensor mengirimkan data hasil pengukurannya ke server melalui jaringan internet. Sensor menggunakan protokol HTTP untuk mengirim datanya. Setelah data diterima pada server, Server mengirimkan data hasil pengukuran dari sensor dalam bentuk konten HTTP. *Website* menerima konten HTTP dari server. *Website* kemudian memproses konten tersebut dan menampilkan data hasil pengukuran dari sensor pada *website*.

C. Metode Klasifikasi

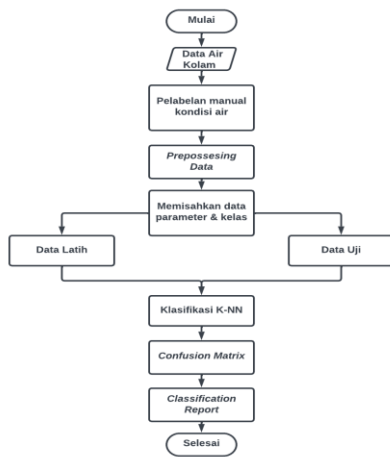
Pada penelitian ini dilakukan juga analisa klasifikasi kondisi air yang menggunakan tiga metode yaitu K-NN, SVM, dan *Decision Tree*. Dan selanjutnya dilakukan perbandingan dari ketiga metode tersebut untuk mendapatkan hasil metode terbaik pada klasifikasi kondisi air kolam bioflok.

1. K-Nearest Neighbor

Pada metode klasifikasi *K-Nearest Neighbors* terdapat arsitektur yang digunakan sebagai landasan dalam proses klasifikasi, terdapat bagian utama, yaitu: bagian *input*, bagian proses klasifikasi dan bagian *output*.



Gambar 4. Arsitektur Metode KNN

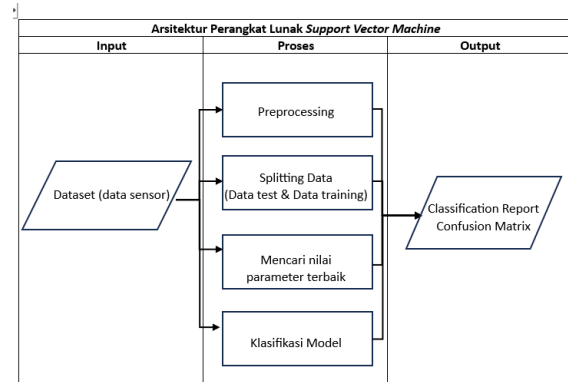


Gambar 5. Diagram Alir Klasifikasi K-NN

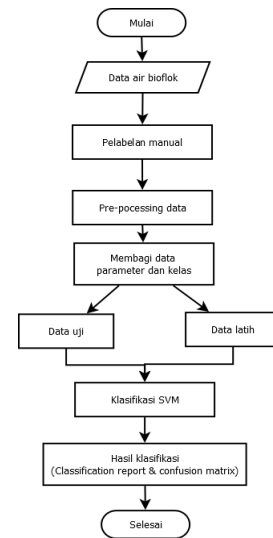
K-NN adalah algoritma *supervised learning* yang mengklasifikasikan *instance* baru berdasarkan mayoritas kategori pada tetangga terdekat. Kelas yang paling sering muncul di antara tetangga yang terdekat akan menjadi kelas hasil klasifikasi [9]. Pada pengukuran ini peneliti membuat hasil parameter menjadi tiga kelas yaitu: kondisi air baik, kondisi air sedang, dan kondisi air buruk dari data parameter pH, amonia dan larutan oksigen. Peneliti juga harus menentukan nilai K dari penyebaran data. Nilai K sangat mempengaruhi hasil klasifikasi. Berdasarkan Gambar 5, proses klasifikasi data menggunakan metode *K-Nearest Neighbors* (K-NN) dimulai dengan pengumpulan *dataset* yang akan digunakan untuk pelatihan dan pengujian. Selanjutnya, peneliti perlu membuat variabel memisahkan data parameter & data kelas. Kemudian dari data variabel yang telah dibuat data latih dan data uji. Setelah itu, menentukan nilai K, dan jarak matriks yang akan digunakan. Nilai K sangat berpengaruh terhadap hasil akurasi yang akan didapatkan. Untuk melihat kebenaran data yang tersebar peneliti membuat matriks konfusi. Kemudian kita menampilkan prediksi data dan hasil akurasi dari data yang telah diuji. Untuk menampilkan Hasil laporan klasifikasi peneliti menampilkan *Classification Report* yang berisi hasil dari nilai *accuracy*, *precision*, *recall*, dan *f1-score*.

2. Support Vector Machine

Pada metode *Support Vector Machine* terdapat arsitektur yang digunakan dalam melakukan klasifikasi.



Gambar 6. Arsitektur Model SVM

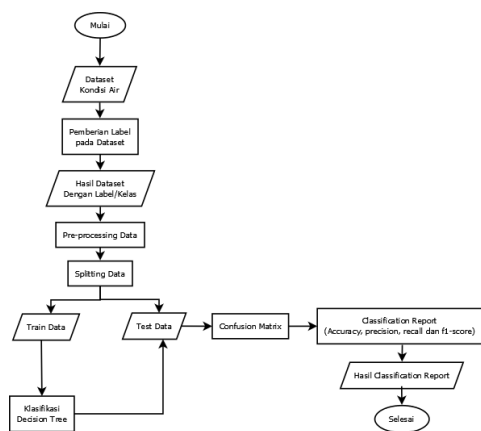


Gambar 7. Diagram Alir Klasifikasi SVM

Support vector machine (SVM) merupakan salah satu metode yang terdapat pada *machine learning* yang biasa digunakan untuk klasifikasi dan regresi. Prinsip kerja dari metode ini yaitu dengan mencari ruang pemisah yang paling optimal dari suatu *dataset* dari setiap kelas yang berbeda [10]. Berdasarkan diagram alir pada Gambar 7 terdapat beberapa proses yang akan dilakukan dalam tahap klasifikasi. Pertama, *import dataset* dari Kualitas air kolam bioflok sebagai *input* data yang akan diklasifikasi, Setelah itu, melakukan pelabelan manual untuk kelas pada *microsoft excel*. Setelah itu, melakukan tahap *preprocessing* untuk meningkatkan kualitas data, selanjutnya memisahkan data dan kelas. Setelah data dan kelas dipisah selanjutnya membagi data uji dan data latih, setelah itu lalu masuk ke klasifikasi menggunakan SVM dimana pada tahap ini menggunakan *multi class* dan memilih *kernel* yang akan digunakan. Setelah itu, data diuji dengan model yang digunakan, setelah pengujian didapat hasil berupa *accuracy*,

precision, recall, dan f1-score serta confusion matrix.

3. Decision Tree



Gambar 8. Diagram Alir Klasifikasi Decision Tree

Decision Tree berbentuk pohon yang terdiri dari beberapa bagian, yaitu *root node*, *internal node*, dan *terminal node*. Data yang melalui *root node* dan *internal node* hingga mencapai *terminal node* merupakan proses klasifikasi dalam metode Decision Tree ini [9]. Berdasarkan diagram alir perancangan klasifikasi metode Decision Tree yang ditunjukkan pada Gambar 8, ada beberapa proses dalam melakukan klasifikasi. Pertama, memasukkan *dataset* kondisi air. Kedua, melakukan pemberian label atau kelas pada *dataset* kondisi air untuk menentukan status kondisi air. Hasil *dataset* yang sudah diberi label digunakan sebagai data untuk melakukan klasifikasi. kemudian pada *dataset* dilakukan *preprocessing* untuk mendapatkan data yang lebih berkualitas. Tahap selanjutnya adalah melakukan *splitting data* atau pemisahan data menggunakan *library train_test_split*. Untuk proses ini, perbandingannya adalah 80% untuk data latihan (*train data*) dan 20% untuk data uji (*test data*). Hasil dari *train data* selanjutnya dilakukan proses klasifikasi menggunakan Decision Tree. Sedangkan hasil dari *test data* dan hasil klasifikasi dilakukan pembuatan *confusion matrix* untuk menentukan performansi dari model klasifikasi dan tahap terakhir adalah pembuatan *classification report* untuk menentukan *accuracy*, *precision*, *recall* dan *f1-score* pada model klasifikasi Decision Tree untuk kondisi air kolam bioflok.

D. Pengujian Alat dan Model Klasifikasi

Sebelum dilakukannya pengambilan data dan monitoring kondisi air pada kolam bioflok, peneliti

harus memastikan bahwa alat yang digunakan memiliki nilai yang hampir identik dengan standar nilai yang ditetapkan untuk alat yang telah teruji.

Pada pengujian sensor pH, dilakukan dengan menggunakan empat jenis air dengan melakukan dua kali pengujian pada masing-masing air, proses dilakukan pada satu hari yang sama dalam jangka waktu yang berbeda-beda. sensor pH yang sudah teruji mengambil data air terlebih dahulu, kemudian dilanjutkan dengan menggunakan sensor pH meter pada *mikrokontroler*, setelah pengambilan data, dibandingkan hasil antara menggunakan pH meter yang sudah teruji dengan pH meter *mikrokontroler*, jika *error* nya lebih tinggi dan mendapatkan hasil yang jauh, maka harus dilakukan kalibrasi ulang sampai mendapatkan *error* yang paling kecil.

Pada pengujian sensor amonia atau MQ-137, dilakukan pada kolam yang akan dimonitoring dengan alat pembanding berupa *amonia test-kit* berupa cairan dan menggunakan tabel *unionized amonia* sebagai penentu nilai amonia sebenar. pengujian dilakukan dengan 5 jenis kolam yang berbeda pada lokasi kolam bioflok.

Pada pengujian sensor DO meter, dilakukan di lab uji BP Batam, Politeknik Negeri Batam. Pengujian dilakukan 5 kali, dikarenakan cairan alat pembanding DO meter yang sudah teruji itu mahal, jenis air yang hanya bisa diujikan yaitu jenis air yang telah disediakan. Sebelum melakukan pembanding, sensor DO *mikrokontroler* harus dikalibrasi terlebih dahulu, terutama suhu pada ruangan tersebut sehingga nilai yang dihasilkan sama dengan Do meter yang telah teruji.

Setelah pengujian alat untuk monitoring kondisi air pada kolam bioflok, peneliti melakukan pengujian metode klasifikasi menggunakan metode KNN, SVM dan Decision Tree. untuk mengevaluasi dan membandingkan kinerja dari ketiga metode tersebut. Metode KNN akan dievaluasi dengan kemampuan dalam mengenali pola-pola kompleks dan memanfaatkan tetangga terdekat. SVM akan diuji sejauh mana kemampuannya memisahkan data-data yang kompleks. Sementara Decision Tree diuji kemampuan dalam menghasilkan aturan keputusan yang *interpretable* dan efektif. Pengujian mencakup seberapa akuratnya hasil dari model klasifikasi untuk mendeteksi status air pada data baru dan dibandingkan dengan status sebenarnya.

Setelah klasifikasi selesai, peneliti melakukan pengujian model klasifikasi menggunakan *predict* yang didapat dari *sklearn* pada klasifikasi KNN, SVM, dan *Decision Tree*. Pengujian Model klasifikasi berfungsi seberapa akurat pengklasifikasian yang telah dilakukan dengan menggunakan *dataset* yang lain selain *dataset* yang digunakan pada saat melakukan pelatihan model klasifikasi namun memiliki fitur yang sama. Berikut merupakan *dataset* yang digunakan dalam pengujian model klasifikasi yang dapat dilihat pada Gambar 9.

pH	o2	amonia	date	Pakan	Kematian
4,76	3	0,01	2023-10-02 01:53:08	0	0 Ekor
4,73	3	0,01	2023-10-02 01:54:09	0	0 Ekor
5,87	4	0,01	2023-10-02 05:40:49	0	0 Ekor
4,75	2	0,01	2023-10-02 07:01:09	0	0 Ekor
4,7	2	0,01	2023-10-02 07:02:09	0	0 Ekor
5,63	2	0,01	2023-10-03 13:11:09	0	0 Ekor
4,63	2	0,01	2023-10-03 13:12:09	0	0 Ekor
5,46	2	0,01	2023-10-03 13:13:09	0	0 Ekor
5,22	2	0,01	2023-10-03 13:14:09	0	0 Ekor
4,6	2	0,01	2023-10-03 13:15:09	0	0 Ekor
6,24	2	0,01	2023-10-03 13:16:09	0	0 Ekor
4,3	2	0,01	2023-10-03 13:17:09	0	0 Ekor
5,24	2	0,01	2023-10-03 13:18:10	0	0 Ekor
5,45	2	0,01	2023-10-03 13:19:09	0	0 Ekor
5,37	2	0,01	2023-10-03 13:20:09	0	0 Ekor
5,89	2	0,03	2023-11-03 10:45:26	0	0 Ekor
6,89	2	0,03	2023-11-03 10:45:56	0	0 Ekor
5,89	2	0,03	2023-11-03 11:02:53	0	0 Ekor
4,76	3	0,01	2023-10-02 01:51:09	0	0 Ekor
6,16	2	0,02	2023-10-03 12:45:10	0	0 Ekor

Gambar 9. Dataset Pengujian Model Klasifikasi.

Hasil dan Diskusi

A. Pengujian Alat Monitoring

Pada data pengujian alat, terdapat beberapa hasil data dari perbandingan sensor dan *kalibrator*-nya. Terdapat data perbandingan yaitu untuk sensor pH, amonia, dan *dissolved oxygen* (larutan oksigen).

Tabel 2. Kalibrasi Kadar pH

pH Meter (PH-009(I)A)	pH Sensor	Error(%)
3,9	4,18	7,18
3,9	4,2	7,69
3,9	4,16	6,67
3,9	4,14	6,15
4,5	4,62	2,67
4,5	4,55	1,11
4,5	4,59	2,00
4,5	4,58	1,78
Rata-Rata Error		4,41

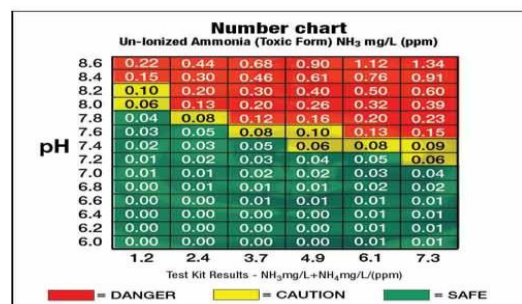
Untuk data pengujian sensor ph dapat dilihat pada Tabel 2. Hasil dari pengujian menentukan bahwa rata-rata *error* dari pH meter dan sensor kit E-201C sekitar 4,41% dan hasil itu membuktikan bahwa perbandingan antara kit E-201C dengan pH meter relatif kecil serta bisa dipastikan bahwa sensor kit E-201C sudah berfungsi dengan baik.

Pada hasil pengujian sensor amonia mendapatkan lima baris tabel dikarenakan pengujian menggunakan lima jenis air yang berbeda. hasilnya dapat dilihat pada Tabel 3.

Tabel 3. Kalibrasi Kadar Amonia

Pengujian ke...	Amonia Test Kit	Sensor MQ-137	Un-ionized Ammonia	Error(%)
1	0	0	0,00	0
2	1	0,01	0,00	0,01
3	3	0,01	0,00	0,01
4	3	0,01	0,00	0,01
5	0	0	0,00	0
Rata-Rata Error				0,006

Pada Tabel di atas, pengujian pada sensor amonia yang dibandingkan dengan *amonia test kit* menghasilkan nilai rata-rata *error* 0,006% dan menggunakan tabel *un-ionized amonia* yang menjadi hasil akhir dari nilai amonia sebenar pada air yang diuji. Hal ini menyatakan bahwa alat sensor amonia MQ-137 memiliki nilai yang akurat dan hanya memiliki *error* yang sangat kecil.



Gambar 10. Tabel Un-Ionized ammonia

Untuk hasil pengujian sensor DO (*dissolved oxygen*) mendapatkan lima hasil perbandingan dengan alat DO meter yang sesungguhnya. Hasil dapat dilihat pada Tabel 4.

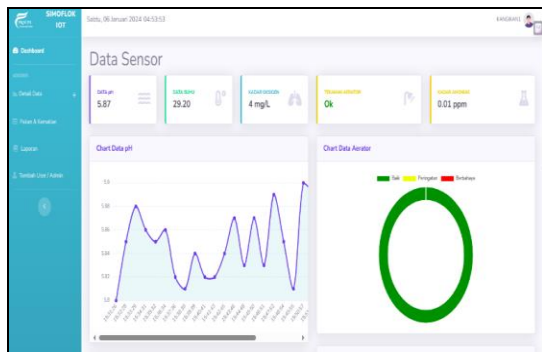
Tabel 4. Kalibrasi kadar DO

Pengujian ke...	DO Meter	DO Sensor	Error(%)
1	5,4	6,02	11,48
2	4,9	5,23	6,73
3	4,6	4,91	6,74
4	4,3	4,41	2,56
5	4,3	4,32	0,47
Rata-Rata Error			5,60

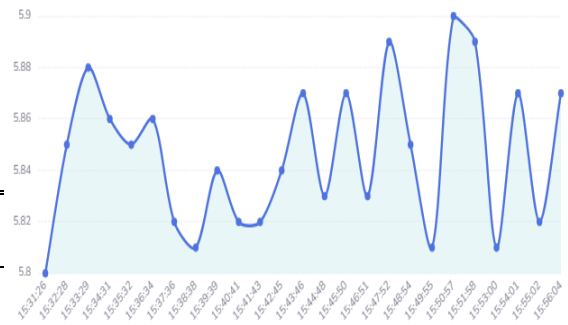
Untuk data pengujian sensor DO dapat dilihat pada Tabel 4 hasil dari kalibrasi sensor DO di atas mendapat rata-rata *error* tidak terlalu tinggi yaitu sekitar 5,60% dan bisa dipastikan sensor DO yang digunakan memiliki nilai yang sesuai.

B. Website Monitoring Kondisi Air

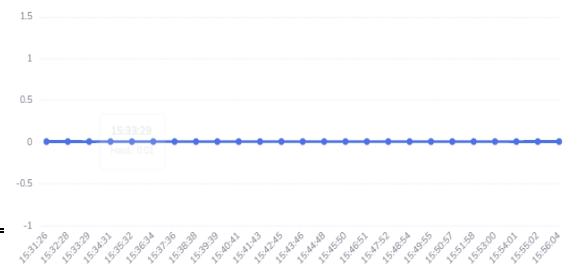
Data kondisi air yang didapat oleh sensor dan diterima ke ESP32, data tersebut dikirim ke *website* yang dilengkapi dengan grafik dari setiap data yang dikirimkan sehingga dapat memantau dan memonitoring kondisi air dari jarak yang berbeda. dilengkapi *history data* yang bisa digunakan sebagai laporan kondisi air pada kolam. data juga dapat digunakan sebagai bahan untuk proses klasifikasi menggunakan metode KNN, SVM, dan *Decision Tree*.



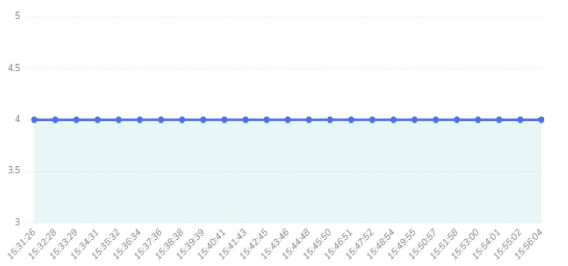
Gambar 11. Dashboard Website



Gambar 12. Grafik Parameter pH



Gambar 13. Grafik Parameter Amonia



Gambar 14. Grafik Parameter Larutan Oksigen

Grafik yang disajikan di atas merupakan grafik nilai pH, amonia, dan larutan oksigen pada *website* untuk memonitoring kondisi air, parameter tersebut yang dapat dilihat oleh pembudidaya untuk memantau perkembangan dari kondisi air pada kolam bioflok di Nongsa.

Gambar 15. Detail Data Keseluruhan Kondisi Air Pada Website

Website juga menyajikan data keseluruhan kondisi air kolam yang di kirim melalui ESP32. Data ini yang digunakan sebagai laporan kondisi

air, data bisa di *download* dalam bentuk excel dan pdf. Data ini yang akan di eksekusi sebagai bahan untuk proses klasifikasi menggunakan metode KNN, SVM, dan *Decision tree*.

C. Dataset Kondisi Monitoring Kondisi Air

Dataset penelitian ini diambil melalui *website*. Data yang diambil dalam waktu 19 hari setelah alat terpasang dengan frekuensi waktu selama satu menit dengan jumlah data 19900 data. kemudian di *preprocessing* untuk mendapatkan data dengan frekuensi selama 30 menit dengan total data sebesar 1716 data.

	pH	o2	amonia	date	Pakan	Kematian	status
0	4.15	1	0.01	2023-10-08 10:00:10	0	0 Ekor	Buruk
1	4.01	1	0.01	2023-10-08 09:30:10	0	0 Ekor	Buruk
2	4.05	1	0.02	2023-10-08 09:00:11	0	0 Ekor	Buruk
3	4.29	2	0.01	2023-10-08 08:30:11	0	0 Ekor	Buruk
4	4.47	1	0.01	2023-10-08 08:00:11	0	0 Ekor	Buruk
...
1711	6.06	4	0.01	2023-09-20 10:00:06	Sabas - 3	2 Ekor	Baik
1712	5.96	4	0.01	2023-09-20 09:30:06	Sabas - 3	2 Ekor	Sedang
1713	5.77	4	0.01	2023-09-20 09:00:05	Sabas - 3	2 Ekor	Sedang
1714	5.79	4	0.01	2023-09-20 08:30:05	Sabas - 3	2 Ekor	Sedang
1715	5.81	3	0.01	2023-09-20 08:00:08	Sabas - 3	2 Ekor	Sedang

1716 rows x 7 columns

Gambar 16. Dataset 19 hari frekuensi 30 menit

D. Hasil pre-processing data

1. Hasil Data Reduction

pH	o2	amonia	date	Pakan	Kematian	Status	Total Data
4.89	2	0.01	2023-10-09 13:37:21	0	0 Ekor	Buruk	19900
4.89	2	0.01	2023-10-09 13:22:05	0	0 Ekor	Buruk	
4.89	2	0.01	2023-10-09 13:19:02	0	0 Ekor	Buruk	
4.89	2	0.01	2023-10-09 13:09:35	0	0 Ekor	Buruk	
4.03	1	0.01	2023-10-08 10:27:11	0	0 Ekor	Buruk	
4.1	1	0.01	2023-10-08 10:26:11	0	0 Ekor	Buruk	
4.02	1	0.01	2023-10-08 10:25:10	0	0 Ekor	Buruk	
3.95	0	0.01	2023-10-08 10:24:11	0	0 Ekor	Buruk	
4	0	0.01	2023-10-08 10:23:11	0	0 Ekor	Buruk	
4.19	1	0.01	2023-10-08 10:22:11	0	0 Ekor	Buruk	
3.81	1	0.01	2023-10-08 10:21:11	0	0 Ekor	Buruk	
4.04	1	0.01	2023-10-08 10:20:10	0	0 Ekor	Buruk	
4.08	1	0.01	2023-10-08 10:19:11	0	0 Ekor	Buruk	
3.9	1	0.01	2023-10-08 10:18:11	0	0 Ekor	Buruk	
4.05	1	0.01	2023-10-08 10:17:10	0	0 Ekor	Buruk	
4.26	1	0.01	2023-10-08 10:16:11	0	0 Ekor	Buruk	
3.93	1	0.01	2023-10-08 10:15:11	0	0 Ekor	Buruk	
3.85	1	0.01	2023-10-08 10:14:11	0	0 Ekor	Buruk	
4.07	1	0.01	2023-10-08 10:13:11	0	0 Ekor	Buruk	

Reduction
Preprocessing

pH	o2	amonia	date	Pakan	Kematian	status	Total Data
4.15	1	0.01	2023-10-08 10:00:10	0	0 Ekor	Buruk	1716
4.01	1	0.01	2023-10-08 09:30:10	0	0 Ekor	Buruk	
4.05	1	0.02	2023-10-08 09:00:11	0	0 Ekor	Buruk	
4.29	2	0.01	2023-10-08 08:30:11	0	0 Ekor	Buruk	
4.47	1	0.01	2023-10-08 08:00:11	0	0 Ekor	Buruk	
4.31	1	0.01	2023-10-08 07:30:11	0	0 Ekor	Buruk	
4.42	1	0.01	2023-10-08 07:00:10	0	0 Ekor	Buruk	
4.41	1	0.01	2023-10-08 06:30:10	0	0 Ekor	Buruk	
4.59	2	0.01	2023-10-08 06:00:11	0	0 Ekor	Buruk	
4.45	2	0.01	2023-10-08 05:30:11	0	0 Ekor	Buruk	
4.52	2	0.01	2023-10-08 05:00:11	0	0 Ekor	Buruk	
4.1	2	0.01	2023-10-08 04:30:10	0	0 Ekor	Buruk	
4	2	0.01	2023-10-08 04:00:10	0	0 Ekor	Buruk	
4.08	2	0.01	2023-10-08 03:30:10	0	0 Ekor	Buruk	
3.85	3	0.01	2023-10-08 03:00:11	0	0 Ekor	Sedang	
3.78	3	0.01	2023-10-08 02:30:11	0	0 Ekor	Sedang	
4.08	3	0.02	2023-10-08 02:00:11	0	0 Ekor	Sedang	
4.05	3	0.01	2023-10-08 01:30:10	0	0 Ekor	Sedang	
3.94	3	0.01	2023-10-08 01:00:10	0	0 Ekor	Sedang	

Gambar 17. Data Pre-processing Reduction

Data reduction mengkompresi data dari awalnya data berjumlah 19900 dikurangi menjadi 1716 data. proses ini menggunakan fitur "sort & filter" pada *microsoft excel* untuk mengambil data dengan range waktu setiap 30 menit.

2. Hasil Transformation

pH	o2	amonia	date	Pakan	Kematian	status
0	4.15	1	0.01	2023-10-08 10:00:10	0	0 Ekor Buruk
1	4.01	1	0.01	2023-10-08 09:30:10	0	0 Ekor Buruk
2	4.05	1	0.02	2023-10-08 09:00:11	0	0 Ekor Buruk
3	4.29	2	0.01	2023-10-08 08:30:11	0	0 Ekor Buruk
4	4.47	1	0.01	2023-10-08 08:00:11	0	0 Ekor Buruk
...
1711	6.06	4	0.01	2023-09-20 10:00:06	Sabas - 3	2 Ekor Baik
1712	5.96	4	0.01	2023-09-20 09:30:06	Sabas - 3	2 Ekor Sedang
1713	5.77	4	0.01	2023-09-20 09:00:05	Sabas - 3	2 Ekor Sedang
1714	5.79	4	0.01	2023-09-20 08:30:05	Sabas - 3	2 Ekor Sedang
1715	5.81	3	0.01	2023-09-20 08:00:08	Sabas - 3	2 Ekor Sedang

Transformation
Preprocessing

pH	o2	amonia	status
0	4.15	1	0.01 Buruk
1	4.01	1	0.01 Buruk
2	4.05	1	0.02 Buruk
3	4.29	2	0.01 Buruk
4	4.47	1	0.01 Buruk
...
1711	6.06	4	0.01 Baik
1712	5.96	4	0.01 Sedang
1713	5.77	4	0.01 Sedang
1714	5.79	4	0.01 Sedang
1715	5.81	3	0.01 Sedang

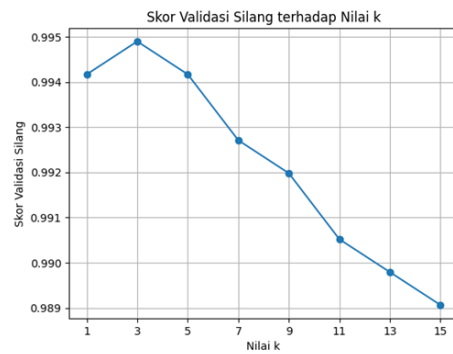
Gambar 18. Dataset Pre-processing Transformation

selanjutnya adalah melakukan data transformation. Proses ini menghilangkan fitur *date*, *Pakan*, dan *Kematian* yang tidak digunakan untuk klasifikasi data kondisi air.

E. Hasil Model Klasifikasi

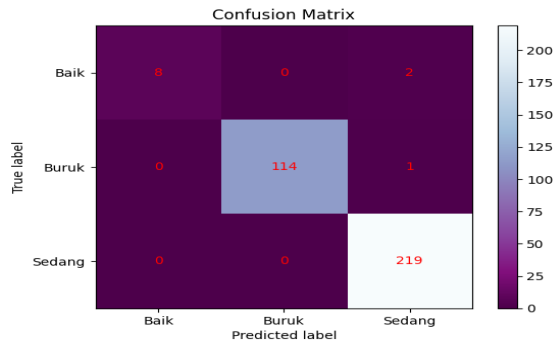
1. Hasil Klasifikasi KNN

Pada tahap klasifikasi menggunakan metode *K-Nearest Neighbors* didapat hasil yang menjadi keluaran klasifikasi, yaitu: grafik menentukan nilai K terbaik, *confusion matrix*, *classification report*, grafik penyebaran data.



Gambar 19. Nilai K terbaik

Tahap pertama ialah menentukan nilai K terbaik menggunakan validasi silang untuk mengurangi *overfitting*. Pada gambar 6. terlihat bahwa hasil pencarian nilai K terbaik adalah K=3. Selanjutnya adalah membangun model klasifikasi K-NN dengan menggunakan nilai K yang sudah didapatkan. Hasil keluaran *confusion matrix* dapat dilihat pada gambar di bawah:



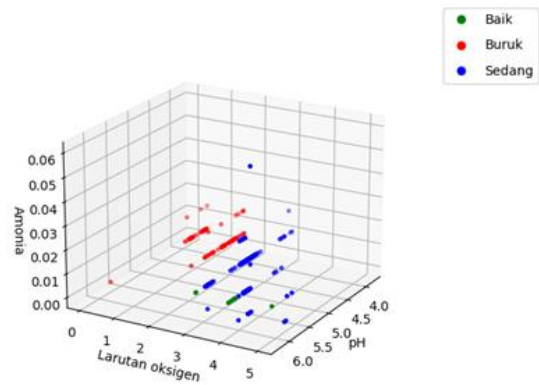
Gambar 20. Confusion Matrix

Jumlah data keseluruhan yang ditampilkan pada gambar 20. sebesar 344 data yang merupakan 20% data yang merupakan data pelatihan dari 1715 data keseluruhan. dari 344 data yang di testing, terdapat 8 data kelas “baik” terprediksi “baik”, 2 data kelas baik terprediksi “sedang”, 114 data kelas “buruk” terprediksi “buruk”, 1 data kelas “buruk” terprediksi “buruk”, dan terdapat 219 data kelas “sedang” terprediksi “sedang”. Hasil *classification report* dapat dilihat pada gambar di bawah:

	precision	recall	f1-score	support
Baik	1.00	0.80	0.89	10
Buruk	1.00	0.99	1.00	115
Sedang	0.99	1.00	0.99	219
accuracy			0.99	344
macro avg	1.00	0.93	0.96	344
weighted avg	0.99	0.99	0.99	344

Gambar 21. Classification Report

Gambar 21 menampilkan laporan hasil klasifikasi yang sudah dilakukan. pada gambar di atas menampilkan hasil nilai presisi, nilai recall, f1_score dan nilai akurasi. pada metode K-NN yang digunakan mendapatkan sebesar 99,6% nilai presisi, 93% nilai recall, 96% nilai f1_score dan nilai akurasi sebesar 99%. hasil ini yang akan dibandingkan dengan dua metode lainnya.

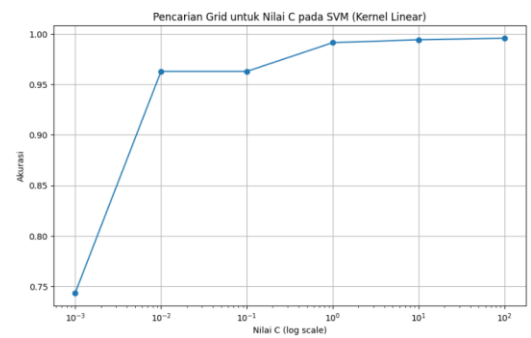


Gambar 22. Keterkaitan Data pada Metode K-NN

Pada penyebaran data di atas merupakan data yang telah dilatih dan diuji, kemudian menggunakan nilai k yang telah ditentukan bernilai 3 sebagai penentuan keputusan untuk keterkaitan antar data. Untuk warna pada grafik menunjukkan bahwa warna yang hijau merupakan daya yang teridentifikasi baik, warna merah teridentifikasi bernilai buruk dan warna biru merupakan data yang teridentifikasi kondisi air sedang. Dari grafik tersebut terdapat data yang seharusnya buruk akan tetapi terbaca sedang, pada nilai amonia = 0,04, pH = 4,5 dan larutan oksigen < 3. Hal ini dikarenakan pengaruh keterkaitan KNN dalam memilih tetangga terdekatnya.

2. Hasil Klasifikasi SVM

Pada Klasifikasi menggunakan 1715 data yang tersusun pada file dengan format .csv. Pada klasifikasi ini menggunakan kernel linear dengan nilai C terbaik dan menggunakan pendekatan multi class “one versus one atau ovo”. Berikut merupakan gambar untuk menentukan nilai C pada kernel linear di SVM:



Gambar 23. Grafik Nilai C Terbaik

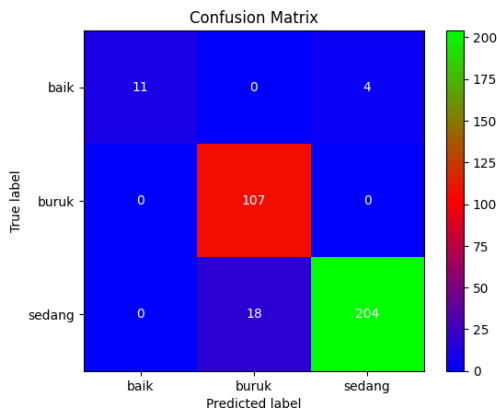
Gambar di atas menunjukkan bahwa dari nilai 0.001, 0.01, 0.1, 1, 10, 100 yang coba dibandingkan, didapatkan nilai C yang terbaik

adalah nilai C = 100. Nilai yang didapatkan dimasukkan dalam tahap klasifikasi untuk mencari *classification report*. Berikut merupakan gambar dari *classification report*.

	precision	recall	f1-score	support
0	1.00	0.73	0.85	15
1	0.86	1.00	0.92	107
2	0.98	0.92	0.95	222
accuracy			0.94	344
macro avg	0.95	0.88	0.91	344
weighted avg	0.94	0.94	0.94	344

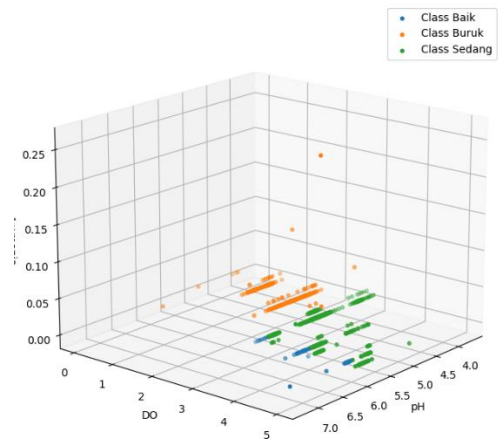
Gambar 24. Classification Report SVM

Dari gambar di atas terdapat nilai akurasi, presisi, *recall*, dan *f1-score* yang akan dibandingkan dengan dua metode lainnya. Pada metode ini didapat nilai akurasi 94%, dan nilai rata - rata presisi 94,4%, rata - rata *recall* 88,3% dan rata - rata *f1-score* 90,6%.



Gambar 25. Confusion Matrix SVM

Pada gambar di atas tertampil *confusion matrix*. Data yang tertampil pada *confusion matrix* berjumlah 344 data yang didapat dari 20% data dari 1715 data yang didapat dari *train test split* pada proses sebelumnya. Pada *confusion matrix* di atas dapat dilihat terdapat 11 data kelas "baik" yang terprediksi "baik", sebanyak 107 data kelas "buruk" yang terprediksi "buruk" dan 204 data kelas "sedang" yang terprediksi "sedang".

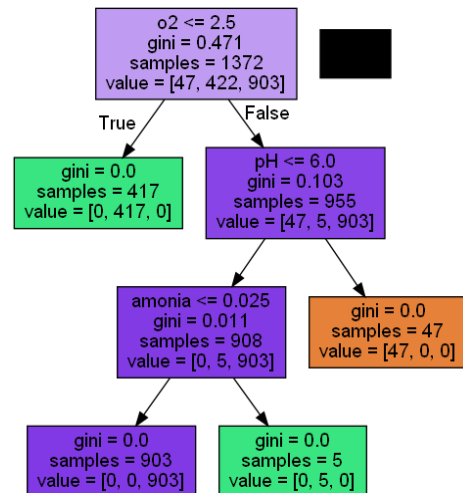


Gambar 26. Penyebaran Data Berdasarkan Kelas

Dapat dilihat di mana plot pada Gambar 26 mengambil data dari tiga parameter yang digunakan dan memberikan warna yang berbeda untuk setiap kelasnya, di mana warna biru untuk kelas "Baik", warna jingga untuk kelas "Buruk" dan warna Hijau untuk kelas "Sedang".

3. Hasil Klasifikasi Decision Tree

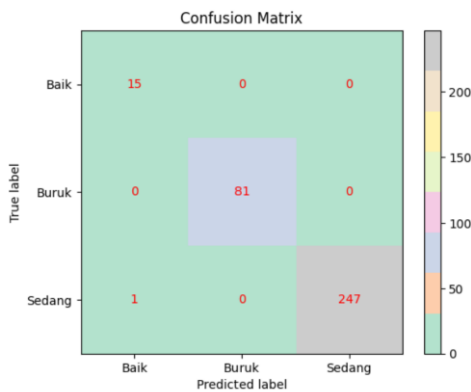
Hasil pengolahan pertama yaitu pembuatan plot *Decision Tree*. Untuk membuat plot menggunakan *library tree* yang disediakan oleh *sklearn*. Sedangkan untuk menampilkan plot menggunakan *library graphviz*.



Gambar 27. Plot Decision Tree

Gambar di atas menjelaskan hasil dari pembuatan plot *Decision Tree*. Dengan menggunakan tingkat kedalaman (*max depth*) tiga tingkatan didapatkan dua *internal node* dan empat *terminal node*. pada *root node* atau titik awal pengambilan keputusan menggunakan fitur

o₂ (oksigen) dengan kriteria ≤ 2.5 . Nilai lain yang terdapat di dalam *root node* yaitu nilai *gini*. Jika *gini* = 0.471 yang berarti fitur tersebut masih belum murni sehingga memerlukan proses *splitting* selanjutnya. Selain *gini*, pada *root node* juga terdapat *samples* dengan nilai 1372 yang menunjukkan jumlah data latih pada proses *splitting data*. *Root node* kemudian dilakukan *splitting* lagi untuk menentukan keputusan selanjutnya dan pada tahap ini menghasilkan satu *internal node* dan satu *terminal node*. Keputusan *true* menghasilkan *terminal node* dengan nilai *gini* = 0.0 yang berarti keputusan tersebut sudah murni dan tidak akan melakukan penentuan keputusan lagi. Informasi yang bisa diperoleh selanjutnya bahwa terdapat 417 data yang termasuk ke dalam kelas Buruk. Keputusan *false* menghasilkan *internal node* dengan *gini* = 0.103 yang menunjukkan apabila data o₂ (oksigen) > 2.5 maka akan dilakukan pengambilan keputusan selanjutnya dengan menggunakan fitur yang lain. Proses pengambilan keputusan akan dilakukan terus-menerus hingga *gini* bernilai 0 dan didapatkan nilai murni untuk kelas Baik dan Sedang.



Gambar 28. Confusion Matrix Decision Tree

Hasil pengolahan kedua yaitu *confusion matrix*. Untuk hasilnya dapat dilihat pada Gambar 28. Dari 344 data, terdapat 15 data yang berada di kelas baik dan terprediksi baik. Kemudian terdapat 81 data yang termasuk ke dalam kelas buruk dan terprediksi buruk. Selanjutnya terdapat 247 data yang termasuk ke dalam kelas sedang dan terprediksi sedang. Terakhir, terdapat satu data yang termasuk ke dalam kelas Sedang namun terprediksi Baik.

	precision	recall	f1-score	support
Baik	0.94	1.00	0.97	15
Buruk	1.00	1.00	1.00	81
Sedang	1.00	1.00	1.00	248
accuracy			1.00	344
macro avg	0.98	1.00	0.99	344

Gambar 29. Classification Report Decision Tree

Hasil pengolahan selanjutnya adalah hasil pengolahan *classification report*. Pada pengolahan ini kita bisa mendapatkan nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang mana nantinya akan dilakukan perbandingan metode menggunakan keempat nilai ini. Untuk hasil *classification report* dapat dilihat pada Gambar 29.

F. Pengujian Model Klasifikasi

Pengujian model klasifikasi dilakukan dengan data baru sebanyak 20 data yang disusun ke dalam *Microsoft excel*. Berikut merupakan hasil pengujian model klasifikasi KNN, SVM, dan *Decision Tree*.

Status sebenar	Status menggunakan metode KNN
Sedang	Sedang
Sedang	Sedang
Sedang	Sedang
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Baik
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Sedang
Sedang	Baik

Gambar 30. Pengujian Data Baru Metode KNN

Pada data status gambar 30 terdapat dua data yang berstatus sedang dibaca baik dan terdapat 1 data berstatus buruk bernilai baik. Data tersebut dapat dilihat pada data yang berwarna kuning.

Pelabelan manual	Hasil Uji SVM
Sedang	Sedang
Sedang	Sedang
Sedang	Sedang
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Baik
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Sedang
Sedang	Buruk

Gambar 31. Pengujian Data Baru Metode SVM

Pada Gambar 31 menunjukkan pengujian model klasifikasi ini menggunakan dataset yang berisikan 20 data, dari 20 data yang diujikan terdapat 18 data yang terprediksi sesuai dengan kelasnya dan dua data yang tidak terprediksi sesuai kelasnya yaitu data "Sedang" terbaca "Baik" maupun terbaca "Buruk".

Kelas Sebenarnya	Klasifikasi menggunakan Decision Tree
Sedang	Sedang
Sedang	Sedang
Sedang	Sedang
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Buruk	Buruk
Sedang	Sedang
Sedang	Buruk

Gambar 32. Pengujian Data Baru Metode Decision Tree

Gambar 32 Merupakan hasil perbandingan antara pengklasifikasian dataset awal dengan pengklasifikasian menggunakan model Decision Tree. Dua data yang tidak sesuai tersebut yaitu data yang memiliki kelas Sedang namun dinyatakan Buruk oleh model.

G. Perbandingan Performansi Waktu

Dalam performansi waktu, setiap metode diuji seberapa cepat masing-masing metode melakukan klasifikasi terhadap data parameter yang di-input, hasil performansi ini yang akan

dibandingkan dari masing-masing metode untuk melihat metode mana yang lebih cepat mengeksekusi dan mempelajari karakteristik data.

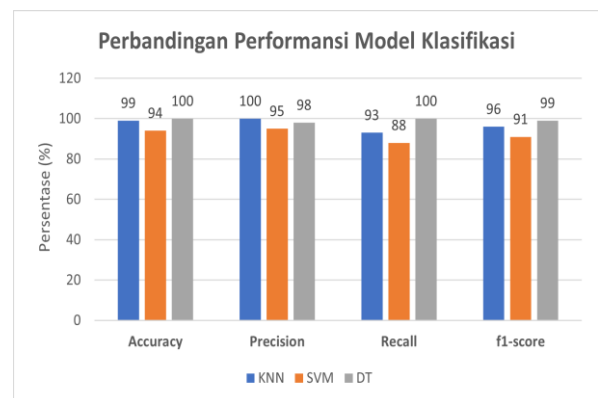
Metode	Jumlah Data					Keterangan
	100	300	500	1000	1716	
K-Nearest Neighbors	1.41s	1.44s	2s	2.87s	3.74s	s (Second)
Support Vector Machine	1.18s	1.39s	1.33s	1.50s	2.31s	
Decision Tree	0,79s	0,85s	0,96s	1,02s	1,20s	

Gambar 33. Hasil Pengujian Performansi Waktu

Pengujian performansi waktu dilakukan menggunakan beberapa skenario data, yaitu : 100 data, 300 data, 500 data, 1000 data dan keseluruhan data yang diambil selama 19 hari setiap 30 menit. Perbandingan performansi waktu setiap data dari masing-masing metode dilakukan untuk mengevaluasi kecepatan pemahaman dan eksekusi data. Hasil yang tercatat dalam tabel di atas menunjukkan bahwa metode K-Nearest Neighbors (KNN) mengalami keterlambatan waktu yang lebih signifikan dalam eksekusi data. Sebaliknya, metode Decision Tree (DT) mampu mengeksekusi data dengan lebih cepat dibandingkan metode Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN).

H. Perbandingan Performansi dari Metode Klasifikasi

Untuk perbandingan pada gambar di atas berdasarkan hasil accuracy, precision, recall, dan f1-score yang telah didapatkan sebelumnya. Namun untuk nilai precision, recall, dan f1-score diambil menggunakan sistem macro average yang juga terdapat pada classification report, hal ini dikarenakan nilai precision, recall, dan f1-score pada classification report ditampilkan nilai per kelas.



Gambar 34. Perbandingan Performansi

Berdasarkan nilai *accuracy*, *precision*, *recall*, dan *f1 - score* diatas didapatkan rata - rata performansi metode K-NN 97%, SVM 92%, dan *Decision Tree* 99,25.

Simpulan

Berdasarkan dari hasil penelitian ini, bahwa penelitian ini telah mampu untuk mengetahui data dari parameter - parameter kondisi air yang dibutuhkan seperti pH, amonia, dan larutan oksigen melalui pengukuran oleh sensor dan dimonitoring melalui *website*. Data yang dimonitoring melalui *website*, dijadikan *dataset* untuk dilakukan proses klasifikasi dengan membangun model dari tiga metode klasifikasi, yaitu: metode KNN, Metode SVM dan Metode *Decision Tree*. Dari hasil perbandingan performansi dari tiga metode tersebut mendapatkan hasil bahwa metode *Decision tree* menjadi yang paling baik dibandingkan dengan metode KNN dan SVM. hal ini dikarenakan metode *Decision Tree* memiliki nilai tertinggi dari performansi dari masing-masing matriks dengan nilai 99,25%, sedangkan Metode KNN memiliki nilai 97% dan SVM memiliki nilai 92%.

Daftar Pustaka

- [1]. D. Azhari, N. Isye Mose, A. Martina Tomaso, P. Studi Teknologi Budidaya Ikan, and J. Perikanan dan Kebaharian, "KAJIAN KUALITAS AIR (SUHU, DO, Ph, Aminia, Nitrat) PADA SISTEM AKUAPONIK UNTUK BUDIDAYA IKAN NILA (*Oreochromis niloticus*)."
- [2]. P. A. Indriati and H. Hafiludin, "Manajemen Kualitas Air Pada Pembenihan Ikan Nila (*Oreochromis niloticus*) Di Balai Benih Ikan Teja Timur Pamekasan," *Juvenil:Jurnal Ilmiah Kelautan dan Perikanan*, vol. 3, no. 2, pp. 27–31, Aug. 2022, doi: 10.21107/juvenil.v3i2.15812.
- [3]. "SNI 7550:2009 Standar Nasional Indonesia," 2009.
- [4]. "DKP - BUDIDAYA IKAN NILA MENGGUNAKAN SISTEM BIOFLOK." Accessed: Mar. 23, 2023. [Online]. Available: <https://dkp.kulonprogokab.go.id/detil/201/budidaya-ikan-nila-menggunakan-sistem-bioflok>.
- [5]. G. P. Putra, Y. Divayana, P. Rahardjo, J. Kampus, and B. Jimbaran, "RANCANG BANGUN SISTEM SMART HOME PADA RUMAH KOS BERBASIS INTERNET OF THINGS," 2022.
- [6]. L. Nul Hakim, A. Taqwa, I. Ziad, J. Teknik Elektro, P. Studi Teknik Telekomunikasi Politeknik Negeri Sriwijaya Jl Sriwijaya Negara, and B. Besar Palembang, "Rancang Bangun Pendeteksi Kebocoran Gas Konsentrasi Amonia (NH3) menggunakan Modul Wifi ESP8266," 2019.
- [7]. "Electrode Specifications".
- [8]. A. Debatara, R. V Manurung, and D. Hiskia, "Mikrotranduser Deteksi Kadar Oksigen Terlarut Aplikasi Monitoring Kualitas Air," 2011.
- [9]. A. Tangkelayuk and E. Mailoa, "Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes Dan Decision Tree," vol. 9, no. 2, pp. 1109–1119, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [10]. T. B. Sasongko, "Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Klasifikasi Jalur Minat SMA)," 201