

## SISTEM PENDUKUNG KEPUTUSAN MENENTUKAN HASIL BUDIDAYA UDANG VANAME DENGAN METODE ALGORITMA C4.5 (PT ANUGERAH SUMBER LAUT JAYA)

Muchamad Iqbal<sup>1</sup>, Wendi Usino<sup>2</sup>, Triono<sup>3</sup>

<sup>1</sup>STMIK Bina Sarana Global, <sup>2</sup>Universitas Budi Luhur, <sup>3</sup>STMIK Bina Sarana Global

<sup>1</sup>miqbal@stmikglobal.ac.id

<sup>2</sup>wendi.usino@budiluhur.ac.id

<sup>3</sup>triono@stmikglobal.ac.id

**Abstrak** - Kondisi geografis Indonesia dengan dua per tiganya perairan dapat memberikan keuntungan tersendiri bagi masyarakat Indonesia yang berprofesi sebagai nelayan. Salah satu komoditas budidaya air laut yang memiliki nilai ekonomi dan permintaan ekspor yang tinggi salah satunya adalah udang vaname (*Liopenaeus Vannamei*), yang sering disebut sebagai udang putih. Salah satu metode klasifikasi adalah algoritma C4.5, metode algoritma C4.5 dalam pemilihan atribut dapat dilakukan dengan cara menggunakan *Gain* dengan tujuan mencari nilai *Entropy*. Algoritma C4.5 dapat membagi-bagi data berdasarkan kriteria yang dipilih agar terbentuk sebuah pohon keputusan (*Decision Tree*) yang menggunakan pendekatan secara *top down*. Dari data-data yang diperoleh maka bisa dihitung dari hasil perhitungan pada luas tambak, bibit, pakan, dan pupuk akan didapatkan hasil optimasinya. Hasil dari program yang dibuat berdasarkan data-data kualitas air yang dianalisa dengan teknik *data mining* dapat dimanfaatkan untuk mendapatkan informasi yang berguna untuk pihak perusahaan dan dari pemodelan pohon keputusan dengan metode Algoritma C4.5 didapatkan hasil *accuracy* sebesar 76,47%, *precision* sebesar 72,72% dan *recall* sebesar 88,88%. Dari hasil pengujian tersebut diketahui bahwa klasifikasi termasuk dalam kategori *Fair Classification*.

**Kata Kunci:** *Optimasi Hasil Panen, Algoritma C4.5, Udang Vaname, Gain, Nilai Entropy, Decision Tree, Fair Classification*

**Abstract** - Indonesia's geographical condition, with two-by-one of its waters, can provide its own benefits for Indonesians who are fishermen. One of the sea water cultivation commodities which has economic value and high export demand one of them is shrimp Vaname (*Liopenaeus Vannamei*), which is often referred to as white shrimp. One method of classification is algorithm C 4.5, algorithm method C 4.5 in attribute selection can be done by using *Gain* with the purpose of finding the value of *Entropy*. The C 4.5 algorithm can divide data based on criteria chosen to form a decision Tree using a top down approach. From the data obtained, it can be calculated from the results of calculations on the area of pond, seedlings, feed, and fertilizer will be obtained the results of its optimization. The results of the programs created based on water quality data analyzed by data mining techniques can be utilized to obtain useful information for the company and from the modeling of decision tree with algorithm method C 4.5 Obtained the accuracy of 76.47%, precision at 72.72% and recall by 88.88%. From the test results it is known that the classification belongs to the *Fair Classification* category.

**Keywords:** *harvest Optimization, C4.5Algorithm, Vannamei Shrimp, Gain, Entropy value, Decision Tree, Fair Classification*

### 1. PENDAHULUAN

Produksi perikanan budidaya komoditas udang vaname selama lima tahun (2009 – 2013) memiliki kecenderungan terus naik. Produksi udang vaname pada tahun 2009 sebesar 170.969 ton, tahun 2010 sebesar 206.578 ton, tahun 2011 sebesar 246.420 ton, tahun 2012 sebesar 251.763 ton, dan pada tahun 2013 sebesar 386.314 ton (Sumber: DJPB Kementerian Kelautan dan Perikanan). Sejalan dengan tren produksi udang yang terus menaik, volume ekspor udang vaname juga menunjukkan peningkatan yang signifikan. Artinya peningkatan produksi udang sejalan dengan peningkatan ekspor.

Di Indonesia udang vaname merupakan salah satu jenis udang yang telah mengalami perkembangan yang pesat karena beberapa keunggulan yang

dimiliki antara lain dapat tumbuh dengan cepat, nilai konsumsi pakan yang rendah dan mampu beradaptasi terhadap kisaran salinitasi yang tinggi serta dapat dipelihara pada tebar yang tinggi. Kualitas air mempunyai peran yang sangat penting pada keberhasilan budidaya udang. Air sebagai media hidup udang, berpengaruh langsung terhadap kesehatan dan pertumbuhan udang. Kualitas air menentukan keberadaan berbagai jenis ekosistem yang ada di dalam tambak, baik terhadap udang maupun biota lainnya. Kualitas air yang jauh dari nilai maksimal dapat menyebabkan kegagalan pertumbuhan udang, sebaliknya kualitas air yang maksimal dapat mendukung pertumbuhan udang.

PT Anugerah Sumber Laut Jaya adalah perusahaan budidaya udang vaname yang berlokasi di Kalianda, Lampung Selatan. Saat ini hasil budidaya udang vaname tidak bagus dalam hal *size* udang vaname,

ini disebabkan karena kualitas air yang kurang optimal. Dengan kondisi tersebut menyebabkan tidak optimalnya keuntungan yang didapat oleh perusahaan. Maka dari itu dibutuhkan suatu sistem yang dapat membantu perusahaan dalam pengambilan keputusan, dalam hal mengelola kualitas air.

Penelitian yang dilakukan oleh (Aditya Wirawan dan Azhari Azhari, 2014) menghasilkan kesimpulan secara umum, kondisi garam dan ketinggian wilayah sangat berpengaruh kepada rekomendasi jenis ikan, kondisi pH dan kandungan O<sub>2</sub> tidak begitu berpengaruh kepada rekomendasi jenis ikan, Fuzzy Mamdani bisa diterapkan untuk penyelesaian permasalahan perikanan sebagaimana telah ada penelitian sebelumnya di bidang pertanian.

Sementara itu Penelitian yang dilakukan oleh (LUMENTUT & Prof. Sri Hartati, M.Sc, 2015) yang menghasilkan kesimpulan secara umum, tentang Sistem Pendukung Keputusan (SPK) yang dibuat dengan kesesuaian kondisi air dan perhitungan analisis finansial sudah sesuai seperti yang diharapkan. Yaitu dapat memilih jenis budidaya ikan yang sesuai harapan serta menguntungkan untuk dibudidayakan. Metode perhitungan TOPSIS dapat membantu dalam penyederhanaan perhitungan dengan melakukan perbandingan pada setiap alternatif yang ada dengan parameter kesesuaian air dan analisis finansial yang didapat.

Sedangkan Penelitian yang dilakukan oleh (Setyorini dan Ratnawati, 2017) menghasilkan kesimpulan secara umum, Dengan dibuatnya sistem untuk menghitung optimasi hasil panen udang vaname menggunakan fuzzy mamdani ini dapat disimpulkan antara lain:

1. Berdasarkan hasil penelitian dapat disimpulkan bahwa fuzzy mamdani bisa digunakan untuk menentukan nilai optimal, normal atau tidak optimal budidaya udang vaname.
2. Dari contoh perhitungan yang ada, didapatkan kesimpulan yaitu menunjukkan hasil 29680 ekor sedangkan dari data realnya diperoleh hasil 29470 ekor. Dengan demikian perbandingan antara data real (data panen) dan setelah proses optimasi memiliki perbedaan 210 ekor.
3. Dari Pengujian menyeluruh dapat dilihat bahwa antara data real dan hasil sistem optimasi memiliki perbedaan terkecil sebesar 27 ekor dan tertinggi sebesar 722 Ekor.

Dan penelitian yang dilakukan oleh (Nono Sudarsono, 2016), menghasilkan kesimpulan secara umum, yaitu dengan membuat aplikasi sistem pendukung keputusan budidaya ikan air tawar. Aplikasi yang di buat menghasilkan informasi berupa rekomendasi pemilihan bibit ikan air tawar

dan mempermudah pengelola dalam pengambilan keputusan budidaya ikan air tawar. Dengan adanya aplikasi yang sudah dibuat, maka tidak adanya hambatan dalam proses pembesaran bibit ikan, dan panen hasil budidaya sesuai dengan waktu yang telah di targetkan. Saran aplikasi sistem penunjang keputusan budidaya ikan air tawar, untuk kedepannya dapat dilakukan pengembangan, seperti penambahan kriteria-kriteria penilaian tiap bibit ikan perlu perawatan, evaluasi dan analisa secara rutin, sehingga perlu diadakan perbaikan atau penyempurnaan terhadap aplikasi yang dibuat.

Berdasarkan latar belakang di atas, yang dapat penulis lakukan adalah dengan cara menganalisa data-data yang berhubungan dengan budidaya udang vaname, dalam sebuah laporan penelitian dengan judul “Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5 (Studi Kasus di PT Anugerah Sumber Laut Jaya)”.

## 2. KAJIAN PUSTAKA DAN RUMUSAN HIPOTESIS

### A. Udang vaname

Udang vaname berasal dari Amerika, dimana populasi ini dapat ditemui di perairan pantai serta laut Amerika Latin seperti Meksiko, dan Nikaragua, serta Puerto Riko. Udang vaname digolongkan dalam famili Penindae. Penggolongan udang vaname secara lengkap sebagai berikut:

Filum : Arthropoda  
Kelas : Crustacea  
Ordo : Decapoda  
Famili : Penaide  
Genus : Litopenaeus  
Spesies : Litopenaeus vaname



Gambar 1. Contoh Udang Vaname

Pada Gambar 1 dapat dilihat bentuk udang vaname yang bersifat *nocturnal* atau jenis hewan yang aktif dalam mencari makan pada malam hari. Sifat lain pada udang vaname adalah kanibalisme, sifat ini sering muncul pada udang yang sehat atau yang tidak sedang ganti kulit dengan sasaran udang yang sedang ganti kulit.

### B. Kualitas Air

Untuk mengetahui cara mengukur kualitas air dapat dilakukan dengan cara visual, yaitu dengan cara melihat kecerahan warna air dan tinggi permukaan

Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5  
(Pt Anugerah Sumber Laut Jaya)

air, atau cara lain dengan menggunakan alat ukur kualitas air. Tabel 1 merupakan standar kualitas air yang di keluarkan oleh tim WWF-Indonesia.

Tabel 1. Tabel Kualitas air  
(Sumber: Tim Perikanan WWF-Indonesia)

Paramater	Optimal	Toleransi
DO (Disscolved Oxygen)	>4 ppm	>3 ppm
Temperatur	28 – 32 <sup>o</sup> C	26 – 35 <sup>o</sup> C
Salinitasi	15 – 25 ppt	0 – 35<35 ppt
pH	7,5 – 8	7 – 8,5
NH <sub>3</sub>	0 ppm	0,1 – 0,5 ppm
No <sub>2</sub>	0 ppm	0,1 – 1 ppm
H <sub>2</sub> S	0 ppm	0,001 ppm
Alkalinitas	100- 120 ppm	>100 ppm
Kecerahan	25 – 40 cm	
Pestisida/insektiida	0 ppb	
Warna air	Hijau kecoklatan	

Dalam budidaya udang pengamatan atau pengecekan kualitas air dilakukan setiap hari dan dicatat dalam formulir pemantauan kualitas air. Pengamatan harian terhadap paramater air yaitu: (Badrudin, 2014)

C. Pengertian Sistem Pendukung Keputusan

Menurut Bonczek, dkk., dikutip dari Turban, *Decision Support Systems And Intelligent Systems*, dikutip dari Nofriansyah, *Sistem Pendukung Keputusan vs Data Mining*, (Dicky Nofriansyah, 2015) mendefinisikan sistem pendukung keputusan sebagai berikut: “Sistem berbasis komputer terdiri dari tiga komponen yang saling berinteraksi, yaitu : sistem bahasa (suatu mekanisme untuk memberikan komunikasi antara pengguna dan komponen sistem pendukung keputusan lain), sistem pengetahuan (repositori pengetahuan merupakan domain masalah yang ada pada sistem pendukung keputusan bisa juga disebut sebagai data atau sebagai prosedur), dan sistem pemrosesan masalah (merupakan hubungan antara dua komponen lainnya, yang terdiri dari satu atau lebih kapabilitas manipulasi masalah umum yang diperlukan untuk pengambilan keputusan).”

D. Data mining

Menurut (Dicky Nofriansyah, 2015) pengertian *data mining* adalah, suatu istilah yang digunakan untuk mengurai penemuan pengetahuan di dalam *database*. *Data mining* merupakan suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

Menurut (Hermawati, 2017) definisi *data mining* adalah proses yang mengerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*)

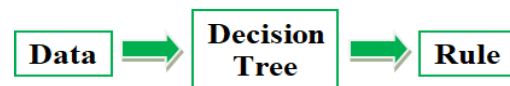
dengan tujuan menganalisis dan mengekstraksi pengetahuan (*knowledge*).

E. Teknik Klasifikasi Dalam *Data mining*

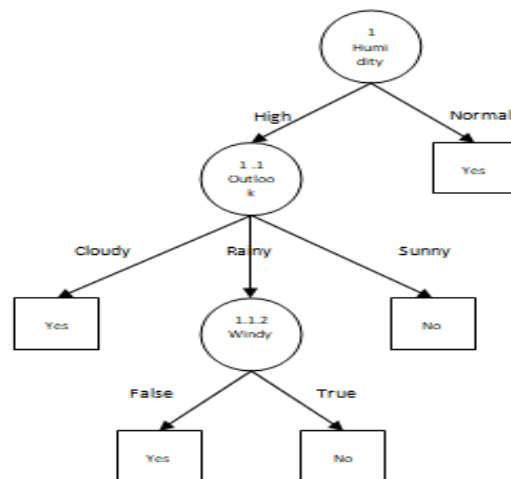
Menurut (Prabowo Pudjo Widodo, Rahmadya Trias Handayanto, 2013) untuk menentukan nilai suatu model baik atau buruk, diperlukan elemen-elemen kunci, diantaranya, Akurasi Prediksi, Kecepatan, *Robustness*, Skalabilitas, *Interpretability*, Kesederhanaan,

F. Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan salah satu metode klasifikasi yang sangat populer karena mudah untuk diinterpretasi oleh manusia. Tingkat akurasi sangat baik, selama data yang akan dijadikan patokan merupakan data yang akurat. Konsep dari pohon keputusan ini adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Terlihat pada Gambar 2



Gambar 2. Contoh Konsep Pohon Keputusan  
(Sumber: Hermawati)



Gambar 3. Contoh Pohon Keputusan  
(Sumber: Hermawati)

Pengertian tentang pohon keputusan (*decision tree*) menurut (Hermawati, 2017) adalah sebagai berikut : Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan sebagai suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon keputusan. Atribut memiliki nilai-nilai yang dinamakan dengan *instance*. Pada pohon keputusan terdapat 3 jenis node, yaitu: *Root Node*, *Internal Node*, *Leaf Node* atau *Terminal Node*.

G. Algoritma C4.5

Algoritma C4.5 adalah salah satu algoritma yang dapat digunakan untuk membuat klasifikasi atau segmentasi. Secara umum penggunaan algoritma C4.5 untuk pengambilan keputusan adalah sebagai berikut: pertama pilih atribut sebagai akar, lalu membuat cabang untuk tiap-tiap nilainya, selanjutnya membagi kasus kedalam cabang, Ulangi proses itu untuk setiap cabang, sampai keseluruhan kasus pada cabang dapat memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti dalam Persamaan 1.

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

- S : himpunan kasus
- A : atribut
- N : jumlah partisi atribut A
- |S<sub>i</sub>| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam S

Sedangkan, perhitungan nilai entropy dapat dilihat pada Persamaan 2.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

Keterangan:

- S : himpunan Kasus
- A : fitur
- n : jumlah partisi S
- P<sub>i</sub> : proporsi dari S<sub>i</sub> terhadap S

H. Gini Index

Salah satu teknik untuk menangani atribut yang bersifat *numeric* atau kontinyu adalah dengan Gini. Beberapa nilai *splitting* dimana jumlah nilai *splitting* = jumlah nilai yang berbeda. Tiap nilai *splitting* mempunyai sebuah *matrix* perhitungan yang berhubungan dengan nilai tersebut.

Rumus mencari nilai GINI =  $1 - \sum_j [p(j | t)]^2$

I. Confusion Matrix

*Confusion matrix* adalah suatu cara pengujian untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan. *Confusion matrix* menggambarkan hasil yang benar dan salah dari suatu model klasifikasi, nilai *confusion matrix* biasanya ditunjukkan dalam satuan persen (%) seperti pada Tabel 2.

Tabel 2. Tabel Confusion Matrix (Sumber: Musthafa, dkk.)

		Predicted Class	
		Bagus	Tidak Bagus
Actual Class	Bagus	True Positive (TP)	False Negative (FN)
	Tidak Bagus	False Positive (FP)	True Negative (TN)

Precision =  $TP / (TP + FP) \times 100\%$   
 Recall =  $TP / (TP + FN) \times 100\%$   
 Accuracy =  $(TP + TN) / \text{Total Sampel} \times 100\%$

Tabel 3. Tabel hasil Confusion Matrix

Akurasi	Hasil
90% - 100%	Excellent Clasification
70% - 80%	Fair Clasification
60% - 70%	Poor Clasification
50% - 60%	Failure

J. Rumusan Hipotesis

Mengacu pada penjelasan tentang faktor-faktor apa saja yang dapat mempengaruhi pengambilan keputusan. Diduga dalam hal metode penelitian yang digunakan untuk sistem pendukung keputusan dengan algoritma C4.5 dapat mengetahui atribut atau kriteria yang tepat dalam budidaya udang vaname, dan dapat menganalisa data-data yang berhubungan dengan budidaya udang, sehingga sistem pendukung keputusan yang dibuat dapat menentukan hasil budidaya udang yang mampu memberikan keputusan terbaik dengan menggunakan metode algoritma C4.5

3. METODE PENELITIAN

Pada metode sistem pendukung keputusan terdapat beberapa informasi yang perlu dipelajari lebih lanjut lagi sehingga informasi-informasi yang diharapkan dan dibutuhkan dalam pemecahan sebuah masalah dapat terselesaikan dengan baik. Pada metode sistem pendukung keputusan, terdapat sebuah metode penambangan data dengan tujuan untuk menyelesaikan masalah dan memberikan informasi-informasi yang dibutuhkan, metode penambangan data tersebut dikenal dengan istilah teknik *data mining*. Penambangan data merupakan suatu proses analisis data dengan tujuan menemukan suatu pola dari kumpulan beberapa data, untuk mendapatkan informasi berupa pola yang mempunyai arti bagi pendukung keputusan.

Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5  
(Pt Anugerah Sumber Laut Jaya)

Salah satu metode yang dapat digunakan pada metode penambangan data diantaranya adalah metode klasifikasi, dimana metode klasifikasi merupakan suatu proses untuk menentukan model data yang dapat menjelaskan konsep, membedakan konsep serta kelas data, dengan tujuan untuk mendapatkan perkiraan kelas dari suatu objek, yang kelasnya tidak diketahui. Salah satu metode klasifikasi yang dapat digunakan adalah algoritma C4.5, pada algoritma C4.5 pemilihan atribut dapat dilakukan dengan cara menggunakan *Gain*, dengan cara mencari nilai *Entropy* terlebih dahulu. Algoritma C4.5 dapat membagi data berdasarkan kriteria yang dipilih untuk membuat sebuah pohon keputusan (*Decision Tree*) yang menggunakan pendekatan secara *top down*.

#### 4. ANALISIS DAN PERANCANGAN

##### A. Pengelompokan dan analisa data

Di dalam penulisan penelitian ini akan melakukan :

1. Mencari *node* dan *leaf*, serta *rule* yang dihasilkan dari *decision tree*.
2. Tidak semua atribut pada kasus ini akan digunakan, penulis hanya menggunakan data sebanyak 6 atribut dan 1 atribut target saja.
3. Uji data inputan
4. Akurasi data
5. Dalam analisis data mining menggunakan 76 data latih dan 17 data uji

Data yang digunakan oleh penulis dalam penelitian ini merupakan data primer yang diperoleh dari PT Anugerah Sumber Laut Jaya, Kalianda Lampung Selatan. Data yang didapat mempunyai 93 data, terdiri dari 76 data sebagai data latih, 17 data sebagai data uji. Data tersebut terdiri dari 11 atribut, serta memiliki 1 atribut target, atribut-atribut ini semua dapat terlihat pada Tabel 4.

Langkah awal yang harus dilakukan sebelum proses klasifikasi adalah dengan cara melakukan *preprocessing data*, untuk pra proses dan analisa data digunakan sebanyak 76 data latih dan 17 data tes.

##### 1. Data selection

Yaitu dengan cara pemilihan data dari sekumpulan data operasional, tahapan ini dilakukan sebelum tahap penggalan informasi dalam KDD dimulai. Data hasil seleksi akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, dan terpisah dari basis data operasional.

Tabel 4. Tabel Data Atribut

Atribut	Keterangan
Pond_local	ID tambak atau kolam merupakan atribut dari data kualitas air, atribut ini tidak dipakai karena hanya sebagai ID nama tambak.
Tgl_data	Merupakan atribut dari data kualitas air, atribut ini tidak digunakan karena hanya sebagai mengetahui waktu pencatatan dalam parameter kualitas air.
Doc	Merupakan atribut dari data kualitas air, doc adalah umur udang dalam budidaya udang dan atribut ini tidak digunakan.
Tinggi air	Merupakan atribut dari data kualitas air, tinggi air mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
Salinitas	Merupakan atribut dari data kualitas air, salinitas mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
DO (Dissolved Oxygen)	Merupakan atribut dari data kualitas air, DO mempengaruhi biota tambak seperti plankton sehingga sangat mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
Ph	Merupakan atribut dari data kualitas air, pH mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
Suhu	Merupakan atribut dari data kualitas air, Suhu mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
Alkalinitas	Merupakan atribut dari data kualitas air, Alkalinitas mempengaruhi dalam budidaya udang sehingga atribut ini dipakai. Data bersifat numeric atau kontinyu dan akan dikategorikan menggunakan GINI.
Kecerahan	Kecerahan atau transparan dari air tambak, atribut ini termasuk dalam data kualitas air tetapi atribut ini tidak digunakan karena terdapat banyak nilai yang kosong atau <i>missing value</i> .
Warna	Warna air termasuk dalam data kualitas air tetapi atribut ini tidak digunakan karena terdapat banyak nilai yang kosong atau <i>missing value</i> .
Output	Atribut target budidaya udang dengan kategori BAGUS yaitu size udang vaname dengan size 50 ekor/kg atau $\geq 20$ gr/ekor dan TIDAK BAGUS yaitu size udang vaname dengan size di atas 50 ekor./kg atau $< 20$ gr/ekor

##### 2. Cleaning

pada umumnya data didapat dari objek yang diteliti, diantaranya memiliki isian-isian yang tidak sempurna, seperti data yang hilang, serta

terdapat data yang tidak terisi. Selain itu, ada juga atribut-atribut data yang tidak relevan, sehingga lebih baik data tersebut dibuang, karena keberadaannya dapat mengurangi mutu atau akurasi dari hasil *data mining* yang diharapkan nantinya.

Dari data yang ada, atribut atau variabel yang dapat digunakan adalah:

1. Tinggi air
2. Salinitas
3. DO
4. pH
5. Suhu
6. Alkalinitas
7. Output

Penggunaan atribut atau variabel tersebut dengan pertimbangan bahwa atribut dari data kualitas air mempengaruhi hasil budidaya udang sehingga diharapkan hasil *output* masuk dalam satu klasifikasi dan dapat berjalan dengan baik.

Tabel 5. Atribut Kualitas Air yang digunakan

Tinggi Air	Salinitas	DO	pH	Suhu	Alkalinitas	Output
130,12	26,23	4,64	7,48	28,73	110	TIDAK BAGUS
130,96	25,85	4,63	7,5	28,69	110,96	TIDAK BAGUS
130,19	25,77	4,77	7,49	28,72	109,81	BAGUS
131,96	26,5	4,67	7,5	28,7	109,42	TIDAK BAGUS
130	25,31	4,83	7,52	28,83	108,08	BAGUS
128,54	25,54	4,74	7,51	28,82	108,85	BAGUS
127,46	26,08	4,85	7,48	28,87	106,92	BAGUS
125,92	26,23	4,87	7,5	28,88	108,46	BAGUS
127,88	26,19	4,92	7,52	28,81	107,5	BAGUS
128,19	26	4,82	7,49	28,78	106,73	TIDAK BAGUS
128,77	25,85	4,9	7,49	28,74	106,35	TIDAK BAGUS
129,23	25,88	4,64	7,53	28,73	110,19	TIDAK BAGUS
131,54	25,73	4,53	7,48	28,73	109,42	TIDAK BAGUS
130,27	26,31	4,56	7,46	28,7	109,23	TIDAK BAGUS
130,96	25,92	4,54	7,48	28,76	108,85	BAGUS
128,81	25,81	4,57	7,46	28,77	109,23	TIDAK BAGUS
129,81	26,51	4,62	7,44	28,75	108,85	TIDAK BAGUS
129,04	25,85	4,6	7,44	28,7	106,92	TIDAK BAGUS
128,88	26,15	4,52	7,49	28,68	109,62	TIDAK BAGUS
126,96	26,75	4,67	7,51	28,8	113,46	TIDAK BAGUS
130,19	26,54	4,78	7,51	28,83	110,58	BAGUS
128,46	26,46	4,81	7,47	28,87	108,65	TIDAK BAGUS

131,15	26,73	4,71	7,5	28,9	111,15	BAGUS
128,96	26,15	4,78	7,52	29,04	114,62	BAGUS
128,81	26,5	4,79	7,51	29,03	109,04	BAGUS
127,15	26,62	5,09	7,48	29,04	109,23	TIDAK BAGUS
126,35	26,88	4,87	7,53	29,18	109,23	BAGUS
126,54	26,58	4,88	7,52	29,07	110,19	BAGUS
127,65	26,31	4,94	7,5	29,02	109,42	BAGUS
129,04	26,5	5,01	7,49	28,93	110,38	TIDAK BAGUS
128,85	26,35	4,76	7,53	28,95	111,15	BAGUS
129,62	26,73	4,74	7,49	28,92	110	TIDAK BAGUS
129,5	26,73	4,68	7,46	28,9	111,15	TIDAK BAGUS
128,27	26,58	4,75	7,48	28,98	110,77	TIDAK BAGUS
127	26,35	4,78	7,49	29,01	110	TIDAK BAGUS
128,08	26,62	4,83	7,45	29,02	111,35	TIDAK BAGUS
128,08	26,54	4,62	7,45	28,88	110	TIDAK BAGUS
127,12	26,35	4,61	7,52	29	113,46	TIDAK BAGUS
119,46	26,73	4,72	7,6	29,21	125	BAGUS
122,85	26,42	4,71	7,61	29,15	124,23	TIDAK BAGUS
123,46	26,35	4,74	7,6	29,1	120	BAGUS
124,62	27,04	4,66	7,58	29,38	121,54	TIDAK BAGUS
119,08	26,15	4,73	7,56	29,22	127,5	BAGUS
121,35	26,69	4,72	7,6	29,17	124,04	BAGUS
123,69	26,38	4,69	7,58	29,03	119,64	BAGUS
124,04	26,08	4,69	7,58	28,91	117,5	TIDAK BAGUS
121,31	26,31	4,89	7,61	28,78	122,69	BAGUS
119,54	26,42	4,74	7,6	28,82	125,58	BAGUS
124,54	26,96	4,72	7,61	28,82	123,65	BAGUS
125,96	26,52	4,69	7,6	28,81	131,54	TIDAK BAGUS
120,38	26	4,9	7,61	28,83	121,15	BAGUS
122,12	26,81	4,7	7,59	28,83	125,58	BAGUS
124,54	26,31	4,94	7,6	28,92	123,46	BAGUS
124,42	27	4,64	7,61	28,92	117,5	TIDAK BAGUS
124,54	26,77	4,72	7,6	29,05	123,46	BAGUS
125,19	26,77	4,72	7,61	29,08	122,31	BAGUS
122,58	26,73	4,72	7,6	28,98	118,65	BAGUS
129,04	25,85	4,6	7,44	28,7	106,92	TIDAK BAGUS
129,81	26,15	4,62	7,44	28,75	108,85	TIDAK BAGUS
130,96	25,92	4,54	7,48	28,76	108,85	BAGUS
130,27	26,31	4,56	7,46	28,7	109,23	TIDAK BAGUS
128,19	26	4,82	7,49	28,78	106,73	BAGUS

Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5  
(Pt Anugerah Sumber Laut Jaya)

127,46	26,08	4,85	7,48	28,87	106,92	BAGUS
128,54	25,54	4,74	7,51	28,82	108,85	BAGUS
130,12	26,23	4,64	7,48	28,73	110	TIDAK BAGUS
127,15	26,62	5,09	7,48	29,03	109,23	TIDAK BAGUS
126,54	26,58	4,88	7,52	29,06	110,19	BAGUS
128,85	26,35	4,76	7,53	28,95	111,15	BAGUS
128,27	26,58	4,75	7,48	28,98	110,77	TIDAK BAGUS
128,08	26,62	4,83	7,45	29,02	111,35	TIDAK BAGUS
122,58	26,73	4,72	7,6	28,98	118,65	BAGUS
124,23	26	4,9	7,61	28,84	121,15	BAGUS
119,54	26,42	4,74	7,6	28,82	125,58	BAGUS
124,04	27,04	4,65	7,57	29,38	121,54	TIDAK BAGUS
123,08	30,12	4,73	7,57	29,08	119,62	BAGUS
129,23	25,88	4,64	7,53	28,73	110,19	BAGUS

Pada data ini, atribut yang bersifat *numeric* atau kontinyu akan ditransformasi datanya, karena data yang akan digunakan harus dirubah terlebih dahulu ke dalam kategori untuk selanjutnya dapat diproses ke dalam data mining. Data-data yang harus ditransformasi adalah tinggi air, salinitas, DO, pH, suhu, alkalinitas dirubah kedalam bentuk kategori. Cara mengategorikan atribut yang bersifat kontinyu yaitu dengan mencari nilai GINI. Untuk rumus mencari nilai GINI

$$GINI = 1 - \sum_j [p(j | t)]^2 \quad (1)$$

Langkah-langkah menghitung Gini untuk atribut kontinyu menggunakan Persamaan (1) sebagai berikut:

- Urutkan atribut berdasarkan nilainya
- Secara linier, pindai nilai-nilai ini setiap waktu perbarui *countmatrix* dengan *mean* antara kedua atribut yang mengapitnya dan hitung gini index
- Pilih posisi split position yang mempunyai gini index paling kecil.

- Hasil perhitungan gini index untuk split 122,965 dan 126,75

$$\begin{aligned} Gini (< 122,965) &= 1 - (10/11)^2 - (1/11)^2 = 1 - 0,826 - 0,008 = 0,165 \\ Gini (>= 122,965) &= 1 - (30/65)^2 - (35/65)^2 = 1 - 0,213 - 0,289 = 0,497 \\ \text{Total Gini index} &\text{ adalah: } Gini = (11/76)*0,165 + (65/76)*0,497 = 0,449 \\ Gini (< 126,75) &= 1 - (22/28)^2 - (6/28)^2 = 1 - 0,617 - 0,045 = 0,336 \end{aligned}$$

$$\begin{aligned} Gini (>= 126,75) &= 1 - (18/48)^2 - (30/48)^2 = 1 - 0,14 - 0,39 = 0,47 \\ \text{Total Gini index} &\text{ adalah: } Gini = (28/76)*0,336 + (48/76)*0,47 = 0,42 \end{aligned}$$

Setelah melakukan perhitungan keseluruhan data tinggi air maka split position  $< 126,75$  dan  $\geq 126,75$  menggantikan data baru karena diantara data tersebut 126,75 memiliki gini index terkecil yaitu 0,42

- Hasil perhitungan gini index untuk split 25,635 dan 26,92

$$\begin{aligned} Gini (< 25,635) &= 1 - (3/3)^2 - (0/3)^2 = 1 - 1 - 0 = 0 \\ Gini (>= 25,635) &= 1 - (37/73)^2 - (36/73)^2 = 1 - 0,257 - 0,243 = 0,5 \\ \text{Total Gini index} &\text{ adalah: } Gini = (3/76)*0 + (73/76)*0,5 = 0,48 \\ Gini (< 26,92) &= 1 - (38/71)^2 - (33/71)^2 = 1 - 0,286 - 0,216 = 0,498 \\ Gini (>= 26,92) &= 1 - (2/5)^2 - (3/6)^2 = 1 - 0,16 - 0,36 = 0,48 \\ \text{Total Gini index} &\text{ adalah: } Gini = (71/76)*0,498 + (5/76)*0,48 = 0,496 \end{aligned}$$

Setelah melakukan perhitungan keseluruhan data Salinitas maka split position  $< 25,635$  dan  $\geq 25,635$  menggantikan data baru karena diantara data tersebut 25,635 memiliki gini index terkecil yaitu 0,48

- Hasil perhitungan gini index untuk split 4,625 dan 4,695

$$\begin{aligned} Gini (< 4,625) &= 1 - (2/13)^2 - (11/13)^2 = 1 - 0,024 - 0,716 = 0,26 \\ Gini (>= 4,625) &= 1 - (38/63)^2 - (25/63)^2 = 1 - 0,364 - 0,157 = 0,479 \\ \text{Total Gini index} &\text{ adalah: } Gini = (13/76)*0,26 + (63/76)*0,479 = 0,441 \\ Gini (< 4,695) &= 1 - (4/27)^2 - (23/27)^2 = 1 - 0,022 - 0,725 = 0,253 \\ Gini (>= 4,695) &= 1 - (36/49)^2 - (13/49)^2 = 1 - 0,539 - 0,07 = 0,391 \end{aligned}$$

$$\text{Total Gini index} \text{ adalah: } Gini = (28/76)*0,253 + (48/76)*0,391 = 0,341$$

Setelah melakukan perhitungan keseluruhan data DO maka split position  $< 4,695$  dan  $\geq 4,695$  menggantikan data baru karena diantara data tersebut 4,695 memiliki gini index terkecil yaitu 0,341

- Hasil perhitungan gini index untuk split 7,495 dan 7,575

$$\begin{aligned} Gini (< 7,495) &= 1 - (6/31)^2 - (25/31)^2 = 1 - 0,037 - 0,65 = 0,313 \\ Gini (>= 7,495) &= 1 - (34/45)^2 - (11/45)^2 = 1 - 0,57 - 0,059 = 0,371 \end{aligned}$$

Total Gini index adalah:  $Gini = (31/76)*0,313 + (45/76)*0,371 = 0,346$   
 $Gini (< 7,575) = 1 - (24/55)^2 - (31/55)^2 = 1 - 0,19 - 0,317 = 0,493$   
 $Gini (>= 7,575) = 1 - (16/21)^2 - (5/21)^2 = 1 - 0,58 - 0,056 = 0,364$   
 Total Gini index adalah:  $Gini = (55/76)*0,493 + (21/76)*0,364 = 0,456$

Setelah melakukan perhitungan keseluruhan data pH maka split position  $< 7,495$  dan  $>=7,495$  menggantikan data baru karena diantara data tersebut  $7,495$  memiliki gini index terkecil yaitu  $0,346$

5. Hasil perhitungan gini index untuk split  $28,755$  dan  $29,16$

$Gini (< 28,755) = 1 - (2/16)^2 - (14/16)^2 = 1 - 0,016 - 0,765 = 0,219$   
 $Gini (>= 28,755) = 1 - (38/60)^2 - (22/60)^2 = 1 - 0,401 - 0,134 = 0,465$   
 Total Gini index adalah:  $Gini = (16/76)*0,219 + (60/76)*0,465 = 0,413$   
 $Gini (< 29,16) = 1 - (36/70)^2 - (34/70)^2 = 1 - 0,264 - 0,236 = 0,5$   
 $Gini (>= 29,16) = 1 - (4/6)^2 - (2/6)^2 = 1 - 0,444 - 0,111 = 0,445$   
 Total Gini index adalah:  $Gini = (70/76)*0,5 + (6/76)*0,445 = 0,495$

Setelah melakukan perhitungan keseluruhan data Suhu maka split position  $< 28,755$  dan  $>=28,755$  menggantikan data baru karena diantara data tersebut  $28,755$  memiliki gini index terkecil yaitu  $0,413$

6. Hasil perhitungan gini index untuk split  $118,075$  dan  $122,5$

$Gini (< 118,075) = 1 - (22/54)^2 - (32/54)^2 = 1 - 0,166 - 0,351 = 0,483$   
 $Gini (>= 118,075) = 1 - (18/22)^2 - (4/22)^2 = 1 - 0,669 - 0,033 = 0,298$   
 Total Gini index adalah:  $Gini = (54/76)*0,483 + (22/76)*0,298 = 0,429$   
 $Gini (< 122,5) = 1 - (30/64)^2 - (34/64)^2 = 1 - 0,219 - 0,281 = 0,5$   
 $Gini (>= 122,5) = 1 - (10/12)^2 - (2/12)^2 = 1 - 0,694 - 0,028 = 0,278$   
 Total Gini index adalah:  $Gini = (64/76)*0,5 + (12/76)*0,278 = 0,463$

Setelah melakukan perhitungan keseluruhan data Alkalanitas maka split position  $< 118,075$  dan  $>=118,075$  menggantikan data baru karena diantara data tersebut  $118,075$  memiliki gini index terkecil yaitu  $0,429$ .

Maka dari perhitungan gini index di atas didapat data-data yang sudah ditransformasi ke dalam bentuk kategori yaitu:

Tabel 6. Nilai Atribut yang sudah dikategorikan

No	Atribut	Kategori
1	Tinggi Air	$< 126,75$
		$>= 126,75$
2	Salinitas	$< 25,635$
		$>= 25,635$
3	DO	$< 4,695$
		$>= 4,695$
4	pH	$< 7,495$
		$>= 7,495$
5	Suhu	$< 28,755$
		$>= 28,755$
6	Alkalinitas	$< 118,075$
		$>= 118,075$

B. Hasil Pengujian

Dalam kasus ini akan dibuat pohon keputusan untuk analisis data menggunakan metode algoritma C4.5 adalah dengan menggunakan  $76$  data dengan  $6$  atribut dan  $1$  atribut target. Tabel data yang siap di analisis sesuai dengan tabel 6. di atas.

Dalam kasus ini akan dibuat pohon keputusan untuk menentukan hasil budidaya udang bagus atau tidak dengan melihat atribut tinggi air, salinitas, DO, pH, suhu, alkalinitas dan output.

Secara umum algoritma C4.5 dapat digunakan untuk membangun keputusan dengan cara sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses tersebut pada setiap cabangnya, sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk menjadikan atribut sebagai akar, harus berdasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti dalam Persamaan (2).

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Sedangkan, perhitungan nilai entropy dapat dilihat pada Persamaan (3).



Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5  
(Pt Anugerah Sumber Laut Jaya)

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (3)$$

Langkah pertama adalah menghitung jumlah kasus untuk keputusan bagus dan tidak bagus, termasuk jumlah kasus untuk atribut tinggi air, salinitas, DO, pH, suhu dan alkalinitas. Selanjutnya menghitung nilai entropy dari semua kasus setelah itu melakukan perhitungan Gain untuk masing-masing atribut.

Tabel 7. Tabel perhitungan Node 1

Nod e	Atribut	Nilai	Kasu s	B	T B	Entro pi	Gain
1		Total	76	40	36	0,998	
	Tinggi Air	< 126,75	28	22	6	0,750	0,119
		>= 126,75	48	18	30	0,954	
	Salinitas	< 25,635	3	3	0	0,000	0,038
		>= 25,635	73	37	36	1,000	
	DO	< 4,695	27	4	23	0,605	0,245
		>= 4,695	49	36	13	0,835	
	pH	< 7,495	31	65	2	0,709	0,234
		>= 7,495	45	34	11	0,802	
	Suhu	< 28,755	16	24	1	0,544	0,135
		>= 28,755	60	38	22	0,948	
	Alkalinitas	< 118,075	54	22	32	0,975	0,107
		>= 118,075	22	18	4	0,684	

Baris Total kasus dan seluruh split position setiap atribut pada tabel di atas hitung nilai entropy dengan Persamaan 2 sebagai berikut:

$$Entropy(kasus) = \left(-\left(\frac{40}{76}\right) * \log_2 \left(\frac{40}{76}\right)\right) + \left(-\left(\frac{36}{76}\right) * \log_2 \left(\frac{36}{76}\right)\right) = 0,998$$

Atribut Tinggi Air

$$Entropy(<126,75) = \left(-\left(\frac{22}{28}\right) * \log_2 \left(\frac{22}{28}\right)\right) + \left(-\left(\frac{6}{28}\right) * \log_2 \left(\frac{6}{28}\right)\right) = 0,750$$

$$Entropy(>= 126,75) = \left(-\left(\frac{18}{48}\right) * \log_2 \left(\frac{18}{48}\right)\right) + \left(-\left(\frac{30}{48}\right) * \log_2 \left(\frac{30}{48}\right)\right) = 0,954$$

Atribut Salinitas

$$Entropy(< 25,635) = \left(-\left(\frac{3}{3}\right) * \log_2 \left(\frac{3}{3}\right)\right) + \left(-\left(\frac{0}{3}\right) * \log_2 \left(\frac{0}{3}\right)\right) = 1$$

Atribut DO

$$Entropy(< 4,695) = \left(-\left(\frac{4}{27}\right) * \log_2 \left(\frac{4}{27}\right)\right) + \left(-\left(\frac{23}{27}\right) * \log_2 \left(\frac{23}{27}\right)\right) = 0,605$$

$$Entropy(>= 4,695) = \left(-\left(\frac{36}{49}\right) * \log_2 \left(\frac{36}{49}\right)\right) + \left(-\left(\frac{13}{49}\right) * \log_2 \left(\frac{13}{49}\right)\right) = 0,835$$

Atribut pH

$$Entropy(< 7,495) = \left(-\left(\frac{6}{31}\right) * \log_2 \left(\frac{6}{31}\right)\right) + \left(-\left(\frac{25}{31}\right) * \log_2 \left(\frac{25}{31}\right)\right) = 0,709$$

$$Entropy(>= 7,495) = \left(-\left(\frac{34}{45}\right) * \log_2 \left(\frac{34}{45}\right)\right) + \left(-\left(\frac{11}{45}\right) * \log_2 \left(\frac{11}{45}\right)\right) = 0,802$$

Atribut Suhu

$$Entropy(< 28,755) = \left(-\left(\frac{2}{16}\right) * \log_2 \left(\frac{2}{16}\right)\right) + \left(-\left(\frac{14}{16}\right) * \log_2 \left(\frac{14}{16}\right)\right) = 0,544$$

$$Entropy(>= 28,755) = \left(-\left(\frac{38}{60}\right) * \log_2 \left(\frac{38}{60}\right)\right) + \left(-\left(\frac{22}{60}\right) * \log_2 \left(\frac{22}{60}\right)\right) = 0,948$$

Atribut Alkalinitas

$$Entropy(< 118,075) = \left(-\left(\frac{22}{54}\right) * \log_2 \left(\frac{22}{54}\right)\right) + \left(-\left(\frac{32}{54}\right) * \log_2 \left(\frac{32}{54}\right)\right) = 0,975$$

$$Entropy(>= 118,075) = \left(-\left(\frac{18}{22}\right) * \log_2 \left(\frac{18}{22}\right)\right) + \left(-\left(\frac{4}{22}\right) * \log_2 \left(\frac{4}{22}\right)\right) = 0,684$$

Selanjutnya hitung nilai Gain setiap atribut dengan Persamaan 1 sebagai berikut:

$$\text{Gain (Tinggi Air)} = 0,998 - \left(\left(\frac{28}{76}\right) * 0,75 + \left(\frac{48}{76}\right) * 0,954\right) = 0,119$$

$$\text{Gain (Salinitas)} = 0,998 - \left(\left(\frac{3}{76}\right) * 0 + \left(\frac{73}{76}\right) * 1\right) = 0,038$$

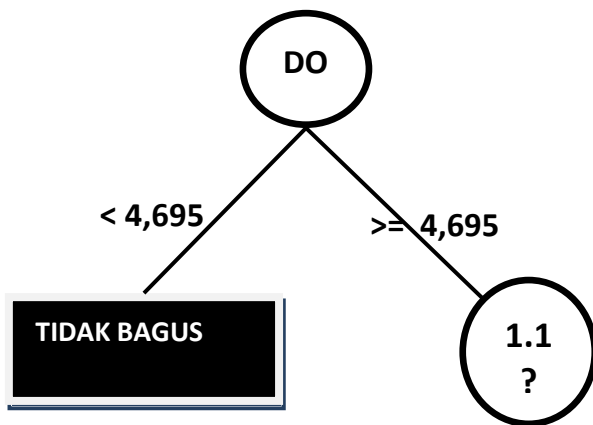
$$\text{Gain (DO)} = 0,998 - \left(\left(\frac{27}{76}\right) * 0,605 + \left(\frac{49}{76}\right) * 0,835\right) = 0,245$$

$$\text{Gain (pH)} = 0,998 - \left( \left( \frac{31}{76} \right) * 0,709 + \left( \frac{45}{76} \right) * 0,802 \right) = 0,234$$

$$\text{Gain (Suhu)} = 0,998 - \left( \left( \frac{16}{76} \right) * 0,544 + \left( \frac{60}{76} \right) * 0,948 \right) = 0,135$$

$$\text{Gain (Alkalinitas)} = 0,998 - \left( \left( \frac{54}{76} \right) * 0,975 + \left( \frac{22}{76} \right) * 0,684 \right) = 0,107$$

Dari hasil perhitungan, diketahui bahwa atribut dengan gain tertinggi adalah DO dengan nilai 0,254, maka atribut DO menjadi node akar (*root node*). Kemudian pada atribut DO dipilih entropy terkecil untuk dijadikan daun atau *leaf* adalah <4,695 dengan nilai entropy 0,605 memiliki 27 kasus dengan nilai 4 BAGUS dan nilai 23 TIDAK BAGUS. Sedangkan atribut DO >= 4,695 perlu dilakukan perhitungan kembali. Dari hasil perhitungan tersebut dapat digambarkan pohon keputusan sementara seperti gambar 4 di bawah ini.



Gambar 4. Hasil Pohon Keputusan Node 1

Langkah kedua, berdasarkan dari pohon keputusan node 1, data-data di atribut DO >= 4,695 kita analisa dengan menghitung jumlah kasus untuk output BAGUS dan TIDAK BAGUS untuk atribut tinggi air, salinitas, pH, suhu, dan alkalinitas. Selanjutnya kita menghitung entropy setiap atribut dan gain tertinggi untuk dijadikan internal node.

Baris Total kasus dan seluruh split position setiap atribut di hitung nilai entropy dengan Persamaan 2 sebagai berikut:

$$\text{Entropy (kasus)} = - \left( \frac{36}{49} \right) * \log_2 \left( \frac{36}{49} \right) + - \left( \frac{13}{49} \right) * \log_2 \left( \frac{13}{49} \right) = 0,835$$

Atribut Tinggi Air

$$\text{Entropy (< 126,75)} = - \left( \frac{21}{22} \right) * \log_2 \left( \frac{21}{22} \right) + - \left( \frac{1}{22} \right) * \log_2 \left( \frac{1}{22} \right) = 0,267$$

Tabel 8. Tabel Perhitungan Node 1.1

Node	Atribut	Nilai	Kasus	B	T B	Entropy	Gain
1.1		TOTAL	49	36	13	0,835	
	Tinggi Air	< 126,75	22	21	1	0,267	0,169
		>= 126,75	27	15	12	0,991	
	Salinitas	< 25,635	3	3	0	0	0,029
		>= 25,635	46	33	13	0,859	
	pH	< 7,495	16	4	12	0,811	0,438
		>= 7,495	33	32	1	0,196	
	Suhu	< 28,755	2	1	1	1	0,008
		>= 28,755	47	35	12	0,820	
	Alkalinitas	< 118,075	31	9	22	0,963	0,112
		>= 118,075	18	17	1	0,310	

$$\text{Entropy (>= 126,75)} = - \left( \frac{15}{27} \right) * \log_2 \left( \frac{15}{27} \right) + - \left( \frac{12}{27} \right) * \log_2 \left( \frac{12}{27} \right) = 0,991$$

Atribut Salinitas

$$\text{Entropy (< 25,635)} = - \left( \frac{3}{3} \right) * \log_2 \left( \frac{3}{3} \right) + - \left( \frac{0}{3} \right) * \log_2 \left( \frac{0}{3} \right) = 0$$

$$\text{Entropy (>= 25,635)} = - \left( \frac{33}{46} \right) * \log_2 \left( \frac{33}{46} \right) + - \left( \frac{13}{46} \right) * \log_2 \left( \frac{13}{46} \right) = 0,859$$

Atribut pH

$$\text{Entropy (< 7,495)} = - \left( \frac{4}{16} \right) * \log_2 \left( \frac{4}{16} \right) + - \left( \frac{12}{16} \right) * \log_2 \left( \frac{12}{16} \right) = 0,811$$

$$\text{Entropy (>= 7,495)} = - \left( \frac{32}{33} \right) * \log_2 \left( \frac{32}{33} \right) + - \left( \frac{1}{33} \right) * \log_2 \left( \frac{1}{33} \right) = 0,196$$

Atribut Suhu

$$\text{Entropy (< 28,755)} = - \left( \frac{1}{2} \right) * \log_2 \left( \frac{1}{2} \right) + - \left( \frac{1}{2} \right) * \log_2 \left( \frac{1}{2} \right) = 1$$

$$\text{Entropy (>= 28,755)} = - \left( \frac{35}{47} \right) * \log_2 \left( \frac{35}{47} \right) + - \left( \frac{12}{47} \right) * \log_2 \left( \frac{12}{47} \right) = 0,82$$

Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname Dengan Metode Algoritma C4.5  
(Pt Anugerah Sumber Laut Jaya)

Atribut Alkalinitas

$$\text{Entropy} (< 118,075) = \left(-\left(\frac{19}{31}\right) * \log_2 \left(\frac{19}{31}\right)\right) + \left(-\left(\frac{12}{31}\right) * \log_2 \left(\frac{12}{31}\right)\right) = 0,963$$

$$\text{Entropy} (>= 118,075) = \left(-\left(\frac{17}{18}\right) * \log_2 \left(\frac{17}{18}\right)\right) + \left(-\left(\frac{1}{18}\right) * \log_2 \left(\frac{1}{18}\right)\right) = 0,31$$

Selanjutnya hitung nilai Gain setiap atribut dengan Persamaan 1 sebagai berikut:

$$\text{Gain (Tinggi Air)} = 0,835 - \left(\left(\frac{22}{49}\right) * 0,267 + \left(\frac{27}{49}\right) * 0,991\right) = 0,169$$

$$\text{Gain (Salinitas)} = 0,835 - \left(\left(\frac{3}{49}\right) * 0 + \left(\frac{46}{49}\right) * 0,859\right) = 0,028$$

$$\text{Gain (pH)} = 0,835 - \left(\left(\frac{16}{49}\right) * 0,811 + \left(\frac{33}{49}\right) * 0,196\right) = 0,438$$

$$\text{Gain (suhu)} = 0,835 - \left(\left(\frac{2}{49}\right) * 1 + \left(\frac{47}{49}\right) * 0,82\right) = 0,008$$

$$\text{Gain (Alkalinitas)} = 0,835 - \left(\left(\frac{31}{49}\right) * 0,963 + \left(\frac{18}{49}\right) * 0,31\right) = 0,112$$

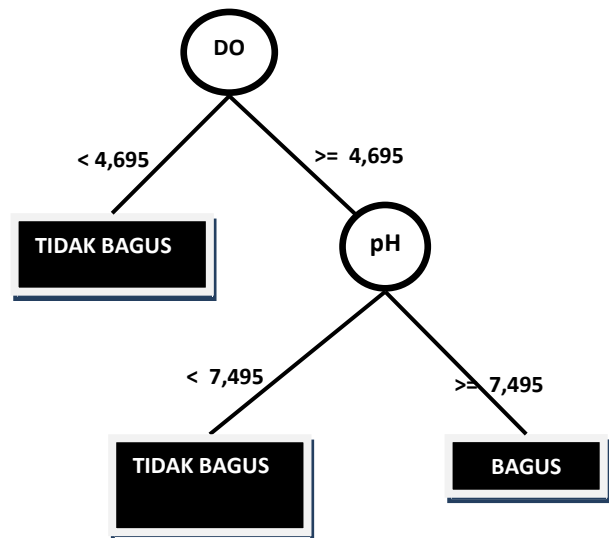
Dari hasil perhitungan diatas dapat terlihat bahwa atribut dengan nilai gain tertinggi adalah pH dengan nilai 0,438, maka atribut pH menjadi node akar (*root node*). Maka diketahui untuk atribut pH  $\geq 7,495$  memiliki 33 kasus dengan nilai 32 BAGUS dan nilai 1 TIDAK BAGUS. Sedangkan atribut pH  $< 7,495$  memiliki 16 kasus dengan nilai 4 BAGUS dan nilai 12 TIDAK BAGUS. Selanjutnya kita analisa lagi dari data yang tersisa dan dapat dilihat hasil data yang tersisa di bawah ini :

Tabel 9. Kasus Data yang Tersisa

Node	Atribut	Nilai	Kasus	B	TB
2.1	-	TOTAL	16	4	12
	Tinggi Air	$< 126,75$	0	0	0
		$\geq 126,75$	16	4	12
	Salinitas	$< 25,635$	0	0	0
		$\geq 25,635$	16	4	12
	Suhu	$< 28,755$	2	1	1
		$\geq 28,755$	14	3	11
	Alkalinitas	$< 118,075$	16	4	12
		$\geq 118,075$	0	0	0

Setelah dilakukan analisa terhadap data yang tersisa maka diambil kesimpulan bahwa pohon keputusan berhenti sampai disini karena tidak ada sampel yang memenuhi test attribute. Dalam hal ini sebuah daun

(*leaf*) dibuat dan dilabeli dengan kelas yang memiliki sampel terbanyak. Pemodelan pohon keputusan berakhir dan dapat digambarkan pohon keputusan seperti gambar 5



Gambar 5. Hasil Pohon Keputusan Akhir

Setelah pohon keputusan terbentuk sesuai pada gambar di atas, maka akan dihasilkan rule dari pohon keputusan adalah sebagai berikut:

- If DO  $< 4,695$  Then hasil = TIDAK BAGUS
- If DO  $\geq 4,695$  and pH  $< 7,49$  Then hasil = TIDAK BAGUS
- If DO  $\geq 4,695$  and pH  $\geq 7,495$  Then hasil = BAGUS

C. Evaluasi

Setelah didapat model pohon keputusan dan menghasilkan rule seperti tersebut di atas maka akan dilakukan pengujian rules dengan data testing sebanyak 17 data.

Tabel 10. Data Testing

Tinggi Air	Salinitas	DO	pH	Suhu	Alkalinitas	Output	Prediksi
128,14	25,85	4,64	7,44	28,73	106,72	BAGUS	TIDAK BAGUS
129,61	26,35	4,72	7,34	28,7	108,85	BAGUS	TIDAK BAGUS
130,5	26,92	4,54	7,48	28,76	108,85	TIDAK BAGUS	TIDAK BAGUS
130,27	26,25	4,56	7,46	28,7	109,23	TIDAK BAGUS	TIDAK BAGUS
128,19	26,5	4,72	7,49	28,73	106,73	BAGUS	TIDAK BAGUS
130,12	26,23	4,64	7,48	28,73	112,5	TIDAK BAGUS	TIDAK BAGUS
127,45	26,62	5,09	7,57	29,03	107,23	TIDAK BAGUS	BAGUS

126,54	26,58	4,88	7,52	29,06	110,19	BAGUS	BAGUS
128,85	26,35	4,76	7,53	28,95	111,15	BAGUS	BAGUS
127,27	26,58	4,75	7,48	28,77	110,77	TIDAK BAGUS	TIDAK BAGUS
127,88	26,57	4,83	7,45	29	111,35	TIDAK BAGUS	TIDAK BAGUS
122,58	26,73	4,72	7,6	28,98	118,65	BAGUS	BAGUS
124,23	26	4,88	7,63	28,84	121,35	BAGUS	BAGUS
119,54	26,42	4,74	7,59	28,82	125,78	BAGUS	BAGUS
126,65	27,15	4,71	7,57	29,38	121,54	BAGUS	BAGUS
125,5	29,5	4,75	7,57	29,38	119,52	BAGUS	BAGUS
127,73	25,88	4,82	7,5	28,63	110,3	BAGUS	BAGUS

Pada tabel di atas menunjukkan ada 4 data yang hasil prediksinya tidak tepat, ditunjukkan dengan kolom yang berwarna biru. Untuk melakukan pengujian maka dihitung akurasi dengan menggunakan metode *Confusion matrix*, *Confusion matrix* merupakan suatu cara pengujian untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan. *Confusion matrix* menggambarkan hasil yang benar dan salah dari suatu model klasifikasi.

Tabel 11. *Counfusion Matrix* data yang diuji

		Predicted Class	
		Bagus (11)	Tidak Bagus (6)
Actual Class (17)	Bagus (11)	True Positive (TP) 8	False Negative (FN) 1
	Tidak Bagus (6)	False Positive (FP) 3	True Negative (TN) 5

$$\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total Sampel} \times 100\%$$

$$= \frac{(8+5)}{17} \times 100\% = 76,47\%$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \times 100\%$$

$$= \frac{(8)}{(8+3)} \times 100\% = 72,72\%$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \times 100\%$$

$$= \frac{(8)}{(8+1)} \times 100\% = 88,88\%$$

Dari perhitungan dengan *counfusion matrix* maka dapat diketahui nilai *accuracy* sebesar **76,47%** untuk nilai *precision* sebesar **72,72%** dan untuk nilai *recall* sebesar **88,88%**. Dari hasil perhitungan yang dilakukan dapat diketahui hasilnya bahwa klasifikasi hasil budidaya udang vaname dengan metode Algoritma C4.5 termasuk dalam kategori **Fair classification**.

## 5. KESIMPULAN

Berdasarkan penelitian Sitem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname dengan Metode Algoritma C4.5, dapat disimpulkan bahwa:

1. Dari hasil pemodelan pohon keputusan yang terbentuk diketahui bahwa atribut atau kriteria DO (*Disscolved Oxygen*) yang paling berpengaruh terhadap budidaya udang vaname.
2. Data-data hasil kualitas air yang dianalisa dengan teknik *data mining* dapat dimanfaatkan untuk mendapatkan informasi yang berguna bagi pihak perusahaan dan dari pemodelan pohon keputusan dengan menggunakan metode Algoritma C4.5 didapatkan hasil *accuracy* sebesar 76,47%, *precision* sebesar 72,72% dan *recall* sebesar 88,88%. Dari hasil pengujian tersebut diketahui bahwa klasifikasi termasuk dalam kategori *Fair Classification*.
3. Klasifikasi dengan Algoritma C4.5 dapat digunakan untuk pendukung keputusan yang dimanfaatkan oleh pihak perusahaan.

## 6. DAFTAR PUSTAKA

Aditya Wirawan Dan Azhari Azhari. (2014). Implementasi Metode Fuzzy-Mamdani Untuk Menentukan Jenis Ikan Konsumsi Air Tawar Berdasarkan Karakteristik Lahan Budidaya Perikanan. *Journal Of Mathematics And Natural Sciences*, 24(1), 29–38.

Badrudin, (Tim Perikanan Wwf-Indonesia). (2014). *Budidaya Udang Vaname Tambak Semi Intnsif Dengan Instalasi Pengolahan Air Limbah (Ipal)*. Wwf-Indonesia.

Dicky Nofriansyah. (2015). *Konsep Data Mining Vs Sistem Pendukung Keputusan* (1st Ed.). Yogyakarta: Depublish.

Hermawati, F. A. (2017). *Data Mining*. Yogyakarta: Putri Chritian.

Lumentut, H. B., & Prof. Sri Hartati, M.Sc, P. . (2015). Sistem Pendukung Keputusan Untuk Memilih Budidaya Ikan Air Tawar Menggunakan Af-Topsis. *Etd Ugm*, 9(2).

Nono Sudarsono, N. S. Dan R. H. (2016). Sistem Penunjang Keputusan Budidaya Ikan Air Tawar Di Giri Tirta Cikalang.

Prabowo Pudjo Widodo, Rahmadya Trias Handayanto, H. (2013). *Penerapan Data Mining Dengan Matlab*. Bandung: Tekayasa Sains.

Setyorini Dan Ratnawati. (2017). Optimasi Hasil Panen Udang Vanamei Di Tambak Menggunakan Metode Fuzzy Mamdani. *Smatika Jurnal*, 7.