



## EKASAKTI JURNAL PENELITIAN & PENGABDIAN (EJPP)

DOI: <https://doi.org/10.31933/ejpp.v4i2>Lisensi: <https://creativecommons.org/licenses/by/4.0/>

Diterima: 13 Mei 2024, Diperbaiki: 29 Mei 2024, Diterbitkan: 1 Juni 2024



### KLASIFIKASI TINGKAT KEBERHASILAN SURVIVAL RATE (SR) PADA PRODUKSI UDANG VANAME MENGGUNAKAN ALGORITMA NAÏVE BAYES

**Ar Razi<sup>1</sup>, Desvina Yulisda<sup>2</sup>**<sup>1</sup>Prodi Teknik Informatika, Fakultas Teknik, Universitas Malikussaleh, [ar.razi@unimal.ac.id](mailto:ar.razi@unimal.ac.id)<sup>2</sup>Prodi Sistem Informasi, Fakultas Teknik, Universitas Malikussaleh, [desvina.yulisda@unimal.ac.id](mailto:desvina.yulisda@unimal.ac.id)Corresponding Author: [ar.razi@unimal.ac.id](mailto:ar.razi@unimal.ac.id)

**Abstract:** Data mining is the process of collecting and processing data with the aim of extracting important information from the data. This process can be done using software that uses mathematical calculations, statistics, or AI. Naive Bayes is the most common classification technique and has a high level of accuracy. Many studies on classification have used the Naive Bayes algorithm. Naive Bayes is a simple probability classification technique used to assume that the explanatory variables are independent. The focus of learning this algorithm is probability estimation. One of the advantages of the naive Bayes algorithm is that the resulting error rate is lower. In addition, this algorithm has a higher level of accuracy and speed when used on larger datasets. This research uses the Naïve Bayes algorithm to classify the Survival Rate (SR) of Vaname shrimp into three classes, namely high, medium and low. The number of sample data used was 200 data which was divided into 2 categories, namely 170 training data and 30 testing data. The variables used in this research are temperature, PH, DO (dissolved oxygen) and salinity. The classification was validated using a confusion matrix and produced an accuracy of 70.4%, precision of 98%, and recall of 79.7%.

**Keywords:** Algoritma Naïve Bayes, Data Mining, AI

**Abstrak:** Data mining adalah proses pengumpulan dan pengolahan data dengan tujuan mengekstrak informasi penting dari data. Proses ini dapat dilakukan dengan menggunakan perangkat lunak yang menggunakan perhitungan matematika, statistika, atau AI. Naive Bayes adalah teknik pengklasifikasian yang paling umum dan memiliki tingkat keakuratan yang tinggi. Banyak penelitian tentang pengklasifikasian telah menggunakan algoritma naive bayes. Naive Bayes adalah teknik pengklasifikasian probabilitas sederhana yang digunakan untuk asumsi bahwa antar variabel penjelas saling bebas (independen). Fokus pembelajaran algoritma ini adalah pengestimasian probabilitas. Salah satu keuntungan algoritma naive bayes yaitu tingkat nilai error yang dihasilkan lebih rendah. Selain itu, algoritma ini memiliki tingkat akurasi dan kecepatan yang lebih tinggi ketika digunakan pada dataset yang lebih besar. Penelitian ini menggunakan algoritma Naïve Bayes dalam mengklasifikasikan Survival Rate (SR) udang Vaname kedalam tiga kelas yaitu tinggi, sedang dan rendah. Jumlah data sample yang digunakan sebanyak 200 data yang dibagi dalam 2 kategori yaitu 170 data training dan 30 data testing. Variable yang digunakan dalam penelitian ini adalah suhu, PH,

DO (dissolved oxygen) dan salinitas. Klasifikasi divalidasi menggunakan confusion matrix dan menghasilkan akurasi sebesar 70,4 %, presisi 98%, dan recall 79,7%.

**Kata Kunci:** Udang Vaname; Data Mining; Algoritma Naïve Bayes

## PENDAHULUAN

Indonesia memiliki potensi perikanan, khususnya budidaya tambak, karena negara ini memiliki banyak pulau dan kekayaan yang sangat berbeda, dan salah satu sektor yang memberikan kontribusi signifikan bagi negara ini adalah perikanan. Pada tahun 2012, sektor perikanan memberikan kontribusi sebesar 2,29%, diikuti oleh peternakan (1,63%) dan perkebunan (1,99%) (Anijar, 2014). Meskipun Indonesia memiliki banyak tambak, pembudidaya menghadapi kesulitan dalam mengelola tambak karena kurangnya pengetahuan tentang cara budidaya yang tepat. Udang kaya akan protein hewani berkualitas tinggi, bebas kolesterol, dan baik untuk pertumbuhan. Salah satu produk ekspor Indonesia adalah udang. Amerika Serikat adalah salah satu negara terbesar di mana udang Indonesia dikirim. Udang vaname adalah salah satu jenis udang yang sangat populer di Indonesia.

Udang putih adalah nama yang digunakan masyarakat untuk udang vaname. Hewan laut yang disebut vaname, yang berasal dari daerah sub tropis pantai Barat Amerika, mulai dari teluk California di Mexico bagian Utara sampai ke pantai Barat Guatemala, El Salvador, Nicaragua, Kosta Rica di Amerika Tengah hingga ke Peru di Amerika Selatan, Udang vaname menjadi populer di tanah air sekitar tahun 2000-an. Karena prospeknya yang menjanjikan dan peluang penjualan yang besar, masyarakat mulai banyak berinvestasi dalam budidaya udang vaname. Jumlah produksi udang menjadi perhatian utama. Sektor perikanan merupakan faktor yang sangat penting terkait ketahanan pangan. Jumlah produksi udang menjadi hal yang paling diperhatikan mengingat udang merupakan komoditas unggulan ekspor saat ini. Tingkat produktivitas udang dipengaruhi beberapa faktor penghambat salah satunya yaitu penyakit pada udang serta kondisi ideal perairan tambak yang mempengaruhi survival rate udang vaname. Penanganan yang kurang cepat dan tidak tepat dapat mengakibatkan gagal panen sehingga produksi udang menjadi menurun serta penghasilan petambak juga berkurang. Karena itu yang perlu dilakukan adalah melakukan inovasi untuk mengatasi permasalahan yang disebabkan oleh kondisi yang mengakibatkan tidak idealnya kondisi tambak menjadi kondisi yang lebih stabil. Setiap jenis udang memiliki ekosistem atau kondisi tambak yang berbeda. Petani perlu mengetahui kondisi tambak yang ideal untuk meningkatkan tingkat survival rate udang vaname, semakin tinggi survival rate udang vaname maka semakin banyak keuntungan yang akan didapat oleh petambak udang. Sehingga peningkatan produksi tetap terjaga. Oleh karena itu, diperlukan pengklasifikasian tingkat keberhasilan Survival Rate (SR) pada produksi udang vaname.

## METODE PENELITIAN

Naïve Bayes pertama kali diusulkan oleh Revered Thomas Bayes, antara tahun 1702 dan 1761 penggunaan Naïve Bayes mulai dikenalkan. Naïve Bayes Classifier adalah suatu algoritma di dalam data mining yang menerapkan teorema Bayes untuk klasifikasi. Penggunaan metode Naïve Bayes dalam pengklasifikasian memiliki keuntungan dimana penggunaan metode ini hanya membutuhkan jumlah data training yang kecil dalam menentukan estimasi parameter selama proses klasifikasi (Budi Santosa, 2007). Menurut definisi lain, Naive Bayes adalah pengklasifikasian probabilitik sederhana yang menghitung kemungkinan dengan menambahkan frekuensi dan menggabungkan nilai dari kumpulan data. Metode ini didasarkan pada asumsi bahwa dalam situasi di mana nilai output diberikan, nilai variabel akan saling bebas. Sebagai contoh, persamaan teorema Bayes dapat

dituliskan sebagai berikut:

$$P(A|B) = \frac{P(B|A)*P(A)}{P(B)} \dots\dots\dots (2.1)$$

Dimana :

$P(A|B)$  : Peluang terjadinya A berdasarkan kondisi B (posteriori prob)

$P(A)$  : Peluang terjadinya A (prior prob)

$P(B|A)$  : Peluang terjadinya B berdasarkan kondisi pada hipotesis A

$P(B)$  : Peluang Terjadinya B

Pada penelitian ini algoritma Naïve Bayes Classifier digunakan untuk mengklasifikasikan tingkat keberhasilan Survival Rate (SR) pada produksi udang vaname, sample data diperoleh dari petambak udang vaname yang berada di desa ujung blang kecamatan banda sakti kota lhokseumawe, dimana sample yang diambil sebanyak 200 data yang nantinya data tersebut akan digunakan sebagai data training dan testing, data yang digunakan sebagai data training sebanyak 170 data dan 30 data digunakan sebagai data testing.

Tahap pengujian model dilakukan dengan menggunakan metode confusion matrix yang menampilkan hasil dari penilaian model melalui tabel matrik. Jika dataset memiliki dua kelas maka kelas yang pertama disebut positif dan kelas lainnya disebut negatif (Putra & Wibowo, 2020). Confusion matrix ini menggunakan tabel matriks seperti pada Tabel 1 dimana terdapat record perbandingan hasil klasifikasi data testing atau data uji berdasarkan data training atau data latih dengan data sebenarnya (Utomo & Mesran, 2020).

True positive adalah jumlah dari record positif yang diklasifikasikan sebagai positif, false positive adalah jumlah dari record negatif yang diklasifikasikan sebagai positif, false negative adalah jumlah dari record positif yang diklasifikasikan sebagai negatif, true negative adalah jumlah dari record negatif yang diklasifikasikan sebagai negatif (Putra & Wibowo, 2020). Untuk mengukur performa suatu model dapat menggunakan empat metode sebagai berikut (Widhi Saputro & Wulan Sari, 2019):

#### 1 Akurasi

Metode pertama ialah akurasi, akurasi adalah metode yang didasari tingkat kedekatan antara nilai prediksi dengan nilai sebenarnya. Akurasi adalah hasil dari penjumlahan nilai diagonal dibagi dengan jumlah total keseluruhan data dan selanjutnya dikalikan 100%. Rumus akurasi dijabarkan pada persamaan berikut.

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \times 100$$

2 Presisi adalah metode yang dipakai untuk menghitung nilai proporsi kelas positif yang berhasil diprediksi dengan benar dari keseluruhan hasil prediksi kelas positif. Presisi menunjukkan jumlah data kategori positif yang diklasifikasi secara benar dibagi dengan total data yang diklasifikasi positif.

#### 3 Recall

Recall adalah metode yang digunakan untuk menghitung presentase kelas data positif yang berhasil diprediksi benar dari keseluruhan data kelas positif.

**Tabel 1.**  
**Confusion Matrix**

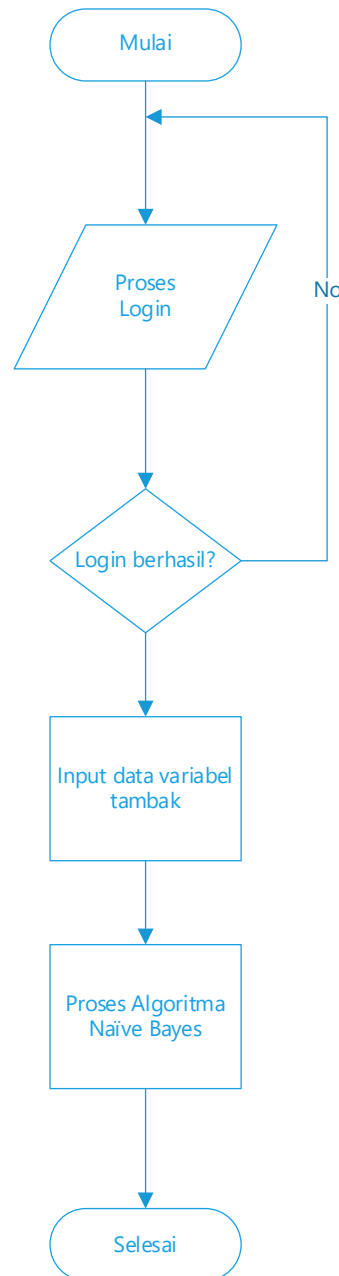
Kelas Asli	Kelas Prediksi	
	Positif	Negatif
Positif	True positif (TP)	True Negatif (FN)
Negatif	False Positif (FP)	False Negatif (TN)

True positive adalah jumlah dari record positif yang diklasifikasikan sebagai positif, false positive adalah jumlah dari record negatif yang diklasifikasikan sebagai positif, false negative adalah jumlah dari record positif yang diklasifikasikan sebagai negatif, true negative adalah jumlah dari record negatif yang diklasifikasikan sebagai negatif (Putra & Wibowo, 2020).

## HASIL DAN PEMBAHASAN

Skema klasifikasi tingkat keberhasilan Survival Rate (SR) pada produksi udang vaname menggunakan algoritma naïve bayes rincian dapat dilihat pada gambar 1. Skema sistem dimulai dari proses login sampai dengan selesai, di mana setiap tahapan dijelaskan berikut ini.

- a. Mulai  
Mulai adalah proses sistem dimulai.
- b. Masukkan username dan password untuk login  
User menginputkan username dan password untuk login ke dalam sistem
- c. Input data variabel tambak  
Pada tahap ini, user menginputkan data testing untuk di klasifikasikan oleh sistem yaitu, suhu, PH, DO (disolved oxygen) dan salinitas.
- d. Sistem melakukan klasifikasi dengan menggunakan algoritma naïve bayes.
- e. Sistem menampilkan hasil klasifikasi dengan tiga kelas yaitu tinggi, sedang dan rendah



**Gambar 1. Skema sistem Klasifikasi menggunakan algoritma naïve bayes**

Pada tahap klasifikasi ini, Algoritma Naïve Baiyes digunakan. Algoritma ini dibagi menjadi dua proses: proses pelatihan untuk pelatihan dan proses pengujian dengan mempertimbangkan kemungkinan dari dataset pelatihan.

#### 1. Proses Training

Pada tahap ini, data training dijadikan sebagai acuan untuk mencari nilai probabilitas, data training di ambil langsung ke lapangan berdasarkan pengalaman dari penambak udang dan data dari dinas perikanan kota lhokseumawe.

**Tabel 2.**  
**Data Training**

No	Suhu	PH	DO (ml/l)	Kelas
1.	28 celcius	7,5	5	Tinggi

2.	26 celcius	6	4	Sedang
3.	24 celcius	9	3	Rendah

## 2. Proses Testing

Tahap ini merupakan tahapan setelah proses training yang telah dilakukan diatas. Pada tahapan ini dilakukan pengujian dengan cara menggunakan data uji kedalam model yang telah dibentuk pada tahapan training diatas.

**Tabel 3.**  
**Data Testing**

No	Suhu	PH	DO (ml/l)	Kelas
1.	29 celcius	8	3	?
2.	28 celcius	7,5	4	?
3.	25 celcius	6	5	?

## Evaluasi Klasifikasi *Naïve Baiyes*

Setelah proses klasifikasi dilakukan diatas maka tahap selanjutnya yaitu tahap evaluasi klasifikasi. Dimana pada tahap ini dilakukan pengujian dengan menggunakan *Confusion Matrix*. Berikut tabel *Confusion Matrix* nya:

**Tabel 4.**  
**Perhitungan Confusion Matrix**

Aktual	Kelas Prediksi		
	Tinggi	Sedang	Rendah
Aktual Tinggi	26	0	0
Aktual Sedang	10	15	0
Aktual Rendah	5	0	10

Kinerja dari algoritma naïve bayes pada penelitian ini sangat baik dimana hasil akurasi yang diperoleh sebesar 70,4 %, presisi 98%, dan recall 79,7% . .

## KESIMPULAN

Berdasarkan serangkaian penelitian dan hasil pembahasan, maka dapat diambil kesimpulan sebagai berikut:

1. Sistem ini dibuat guna bisa mempermudah masyarakat dalam menentukan parameter variabel terbaik untuk kondisi perairan udang vaname.
2. Metode *Naïve Bayes* mempunyai keakuratan yang baik dalam menentukan kondisi optimal parameter suhu,ph dan sanilitas air tambak dalam meningkatkan produksi udang.

## REFERENSI

- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4(2), 437. <https://doi.org/10.30865/mib.v4i2.2080>
- Putra, D., & Wibowo, A. (2020). Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma *Naïve Bayes*. *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2, 84–92.
- Budi Santosa. (2007). *Data Mining : Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Garah Ilmu.
- Widhi Saputro, I., & Wulan Sari, B. (2019). Uji Performa Algoritma *Naïve Bayes*

- untuk Prediksi Masa Studi Mahasiswa Naïve Bayes Algorithm Performance Test for Student Study Prediction. *Citec Journal*, 6(1).
- Nurdin, N., Suhendri, M., Afrilia, Y., & Rizal, R. (2021). Klasifikasi Karya Ilmiah (Tugas Akhir) Mahasiswa Menggunakan Metode Naive Bayes Classifier (NBC). *Sistemasi*, 10(2), 268. <https://doi.org/10.32520/stmsi.v10i2.1193>
- Nurjanah, W. E., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1750–1757.
- Sulistyastuti, I. A. A. (2020). STRATEGI KOMUNIKASI PERSUASIF GOJEK DALAM MENARIK MINAT PELANGGAN (Studi Kasus Gojek di Wilayah Ponorogo). *Skripsi*.
- Syadid, F. (2019). Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency Invers Document Frequency (Tf-Idf) Dan Metode Multi Layer Perceptron (Mlp) Neural Network. *Skripsi Universitas Islam Negeri Syarif Hidayatullah Jakarta*, 72.
- Wahyudi, R., & Kusumawardana, G. (2021). Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *Jurnal Informatika*, 8(2), 200–207. <https://doi.org/10.31294/ji.v8i2.9681>