



ISSN : 2339 - 1871

JURNAL ILMIAH BETRIK

Besemah Teknologi Informasi dan Komputer

Editor Office : LPPM Sekolah Tinggi Teknologi Pagar Alam, Jln. Masik Siagim No. 75
Simpang Mbacang, Pagar Alam, SUM-SEL, Indonesia
Phone : +62 852-7901-1390.
Email : betrik@sttpagaralam.ac.id | admin.jurnal@sttpagaralam.ac.id
Website : <https://ejournal.sttpagaralam.ac.id/index.php/betrik/index>

KLASIFIKASI MINAT BELI PELANGGAN TERHADAP UDANG VANAME MENGGUNAKAN ALGORITMA NAÏVE BAYES

Aqillah Fadia Haya¹, Reynaldi Azhar², Muhamad Khandava Mulyadien³
Betha Nurina Sari⁴

^{1,2,3,4}Program Studi Teknik Informatika Universitas Singaperbangsa Karawang
Jl. HS.Ronggo Waluyo, Puseurjaya, Kec. Telukjambe Tim.,
Kabupaten Karawang, Jawa Barat 41361

Sur-el : aqillah.fadia18029@student.unsika.ac.id¹, reynaldi.azhar18064@student.unsika.ac.id²,
muhamad.khandava18062@student.unsika.ac.id³,
betha.nurina@staff.unsika.ac.id⁴

Abstrak: Udang *vaname* merupakan jenis udang yang kebal dalam serangan penyakit dan memiliki peluang keuntungan yang menjanjikan sehingga banyak pengusaha budidaya udang menyukai udang jenis *vaname* ini. Tahun 2019 Indonesia mampu mengeksport udang sebanyak 90% ke negara Amerika Serikat, Malaysia, Thailand, Belanda, Cina, Jepang, Singapura dan Britania Raya. Banyaknya jumlah ekspor mengakibatkan pengusaha udang ini kewalahan dalam mengatasi penjualan. Penelitian yang dilakukan oleh Wijaya & Dwiasnati (2020) membuktikan bahwa algoritma *Naive Bayes* berguna untuk memudahkan strategi penjualan obat, sehingga pada penelitian ini akan melakukan strategi penjualan minat beli pelanggan terhadap udang *vaname* dengan menggunakan 206 data penjualan pada bulan Oktober sampai Desember 2021. Atribut yang digunakan adalah usia, waktu pembelian dan status. Pengolahan data akan dibantu oleh bahasa pemrograman *python* dan menggunakan algoritma *Naive Bayes* dengan metode *Knowledge Discovery In Databases*. Hasil akhir dalam penelitian ini diperoleh akurasi 83% dengan *precision 100%* dan *recall 77%* sehingga penelitian penjualan udang menggunakan algoritma *Naive Bayes* dapat dilakukan.

Kunci Utama: *Confusion Matrix*; *Data Mining*; Klasifikasi; *Knowledge Discovery In Databases*; *Naive Bayes*.

Abstract: *Vannamei shrimp* is a type of shrimp that is immune to disease attacks and has promising profit opportunities so that many shrimp farming entrepreneurs like this type of *vaname* shrimp. In 2019 Indonesia was able to export 90% of shrimp to the United States, Malaysia, Thailand, the Netherlands, China, Japan, Singapore and the United Kingdom. The large number of exports has resulted in this shrimp entrepreneur being overwhelmed in dealing with sales. Research conducted by Wijaya & Dwiasnati (2020) proves that the *Naive Bayes* algorithm is useful for facilitating drug sales strategies, so that in this study we will conduct a sales strategy for customer buying interest in *vaname* shrimp by using 206 sales data from October to December 2021. used are age, time of purchase and status. Data processing will be assisted by the *python* programming language and using the *Naive Bayes* algorithm with the *Knowledge Discovery In Databases* method. The final result in this study obtained 83% accuracy

with 100% precision and 77% recall so that research on shrimp sales using the Naive Bayes algorithm can be carried out.

Keywords : Confusion Matrix; Data Mining; Classification; Knowledge Discovery In Databases; Naïve Bayes.

1. PENDAHULUAN

Litopenaeus vaname atau udang *vaname* ini masuk ke Indonesia pada Tahun 2001 dan merupakan jenis udang yang berasal dari pantai bagian barat pasifik Amerika Latin [2]. Indonesia merupakan salah satu negara yang pengusahanya banyak membudidayakan udang jenis ini, hal tersebut dikarenakan udang *vaname* ini mempunyai peluang keuntungan yang menjanjikan, memiliki selera makan yang baik dan kebal dari serangan penyakit [1]. Tahun 2019 Indonesia mampu mengeksport udang sebanyak 90% ke negara Amerika Serikat, Malaysia, Thailand, Belanda, Cina, Jepang, Singapura dan Britania Raya (Kusuma 2021). Tingginya pengiriman udang ini menandakan permintaan akan udang *vaname* yang semakin pesat dan berkembang.

Dalam dunia bisnis persaingan antar pengusaha adalah suatu hal yang wajar, maka dari itu seorang pengusaha harus memutar otak untuk mencari bagaimana cara menjamin kelangsungan bisnis mereka dan selalu membuat pelanggan nyaman untuk terus membeli barang dan jasa di tokonya [6]. Pengusaha diusahakan memiliki ketersediaan data yang cukup, yang nantinya data tersebut akan diolah untuk memperoleh beberapa pengetahuan baru yang dapat membantu dalam menjalankan strategi bisnis [9]. Pandangan mengenai minat konsumen penting bagi pengusaha dalam menjalankan bisnisnya, salah satunya adalah masalah strategi waktu [5]. Dengan mengatur strategi waktu seorang pengusaha bisa memberikan informasi kepada pelanggan untuk mengetahui waktu kapan yang cocok untuk mereka datang ke toko tersebut dengan tujuan agar pelanggan tidak kecewa dan

mengetahui waktu kapan yang cocok untuk mereka datang ke toko tersebut.

Metode *data mining* bisa diterapkan dalam melakukan suatu strategi terhadap minat beli pelanggan, salah satu algoritma *data mining* adalah *Naive Bayes*. Teorema bayes menjadi dasar dari algoritma *Naive Bayes* yang merupakan jenis algoritma klasifikasi yang memiliki akurasi tinggi jika menggunakan data yang besar dan *Naive Bayes* ini menggunakan statistika yang berfungsi untuk melakukan pandangan probabilitas dari anggota suatu kelas [7]. Sinta Maulina Dewi, Agus Perdana Windarto, Dedy Hartama menerapkan *Naive Bayes* untuk Strategi Penjualan Produk di UD.Selamat Selular [3]. Terdapat penelitian lain yang juga menerapkan algoritma *Naive Bayes* untuk memudahkan strategi penjualan obat [12].

Dari latar belakang yang sudah dijabarkan, maka penelitian ini dilakukan untuk melakukan strategi penjualan terhadap minat beli pelanggan. Data yang digunakan adalah data sampel penjualan udang di salah satu tambak yang ada di Pulau Nusakambangan. Tahapan penelitian ini dilakukan dengan metodologi KDD (*Knowledge Discovery In Databases*) dengan melakukan proses klasifikasi menggunakan algoritma *Naive Bayes*. Penelitian ini bertujuan untuk mengetahui apakah algoritma *Naive Bayes* dapat digunakan untuk mengetahui strategi penjualan terhadap minat beli pelanggan di tambak udang yang ada di Pulau Nusakambangan.

2. METODE PENELITIAN

Penelitian ini menggunakan algoritma *Naive Bayes* dengan menerapkan metode KDD (*Knowledge Discovery In Databases*). KDD (*Knowledge Discovery*

In Databases) adalah salah satu metodologi yang diterapkan pada data mining dimana proses KDD ini akan menghasilkan suatu *knowledge* yang nantinya dapat digunakan untuk mengambil suatu keputusan [4].

Penelitian ini dalam melakukan pengolahan data akan dilakukan dengan bahasa pemrograman *python* yang dibantu dengan *Google Collabs*. Supaya metode ini mudah dipahami akan dijabarkan mengenai tahapan KDD, sebagai berikut :

2.1 Data Selection

Data Selection perlu dilakukan sebelum tahap selanjutnya dilakukan. *Data selection* merupakan tahap dimana akan melakukan proses pemilihan data atau atribut yang akan digunakan dalam penelitian.

2.2 Data Preprocessing

Data Preprocessing perlu dilakukan sebelum mulai ke proses KDD selanjutnya. *Data Preprocessing* merupakan tahap membersihkan data atau biasa disebut dengan *cleansing*, dimana pada tahap ini akan menghapus data yang tidak digunakan, memperbaiki data yang akan digunakan dalam proses *data mining*.

2.3 Data Transformation

Tahap *transformation* dilakukan untuk menormalisasikan variabel yang digunakan, dimana akan menggunakan *standardscaler* pada *python*.

2.4 Data Mining

Data Mining adalah suatu proses mencari informasi yang unik pada suatu data yang digunakan dengan menggunakan suatu metode dan algoritma tertentu [11]. Tahap ini akan dilakukan pemodelan untuk strategi penjualan minat beli pelanggan menggunakan algoritma dalam *data mining* yaitu *Naïve Bayes*

2.5 Algoritma Naïve Bayes

Algoritma *Naïve Bayes* ditemukan oleh Thomas Bayes dimana algoritma ini merupakan metode yang prosesnya menilai dari probabilitas sederhana dan data statistik yang digunakan dalam melakukan prediksi peluang dilakukan sebelumnya [10].

Algoritma *Naïve Bayes* menyatakan bahwa suatu atribut atau variabel tidak ada kaitannya dengan atribut atau variabel yang lain. Rumus Algoritma *Naïve Bayes* dapat dilihat pada (1) [8].

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

X	:	Data yang belum dik
H	:	Hipotesis <i>class</i> nya
P(H X)	:	Probabilit berdasark
P(H)	:	Probabilit
P(X H)	:	Probabilit berdasark
P(X)	:	Probabilit

2.6 Evaluation

Tahap *evaluation* merupakan kegiatan evaluasi kinerja dari suatu algoritma yang digunakan, dimana menggunakan *confusion matrix* untuk mengukur kinerjanya.

Tabel 1. Tabel *Confusion Matrix*

Class	Diklasifikasikan Sebagai	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	<i>TP</i> (<i>True Positive</i>)	<i>FN</i> (<i>False Negative</i>)
<i>Negative</i>	<i>FP</i> (<i>False Positive</i>)	<i>TN</i> (<i>True Negative</i>)

Keterangan :

TP = Jumlah *True Positives*

TN = Jumlah *True Negatives*

P = Jumlah Tupel Positif

N = Jumlah Tupel Negatif

FP = Jumlah *False Positive*.

Confusion matrix adalah *matrix* yang berfungsi untuk mengukur kinerja model klasifikasi terhadap data *training* dengan nilai yang sudah diketahui [9]. Tabel *confusion matrix* dapat dilihat pada Tabel 2.1. *Confusion matrix* dapat mengukur *Accuracy*, *Precision*, dan *Recall* [10].

Nilai *Accuracy* dapat diukur menggunakan rumus pada nomor (2).

$$\frac{TP + TN}{TP + TN + FP} \quad (2)$$

Nilai *Recall* dapat diukur menggunakan rumus pada nomor (3).

$$\frac{TP}{TP + FN} \quad (3)$$

Nilai *Recall* dapat diukur menggunakan rumus pada nomor (4).

$$\frac{TP}{TP + FP} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Penelitian ini hanya menggunakan 3 atribut dari 5 atribut dan akan mengubah semua data teks menjadi data numerik, dimana pada status tidak beli akan di ubah menjadi (0) dan beli menjadi (1). Terdapat 206 data pada bulan Oktober sampai Desember 2021 dan selama proses mengolah data ini akan dibantu oleh bahasa pemrograman *python* dan *Google Collabs*. Selanjutnya penjelasan dalam proses pengolahan data yang menggunakan metode KDD akan dijabarkan sebagai berikut :

3.1 Data Selection

Tahap awal yang harus dilakukan adalah import data pembelian udang berbentuk .csv ke dalam *Google Collabs*, dimana terdapat atribut id, kelamin, usia, dan waktu. Dapat dilihat pada Tabel 2.

Tabel 2. Data Asli Transaksi Udang

id	kelamin	usia	waktu pembelian	status
1	Female	19	12	Tidak beli
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
206	male	27	7	Beli

3.2 Data Preprocessing

Tahap awal dalam *data preprocessing* disini adalah dengan cara membersihkan atribut yang tidak digunakan dalam proses pengolahan data, dimana atribut awal dapat dilihat pada Tabel 2., sehingga pada Tabel 3 adalah atribut yang sudah di *cleaning* yaitu usia, waktu pembelian dan status.

Tabel 3. Atribut setelah di *cleaning*

usia	waktu pembelian	status
19	12	0
.	.	.
.	.	.
.	.	.
30	8	1

Kemudian memisahkan antara data *training* dengan *testing* menggunakan data test 20% dari 100% data yang ada. Terdapat *output* 165 data *x_train* pada Tabel 3.

Tabel 4. Hasil x_{train} 80%

[[19 5]
[35 8]
[19 9]
.
.
.
[36 7]
[27 7]
[29 12]]

Tabel 4. terdapat 42 data hasil *output* dari x_{test} 20%.

Tabel 5. Hasil x_{test} 20%

[[20 9]
[30 6]
[28 4]
.
.
.
[28 9]
[27 10]
[48 9]]

3.3 Data Transformation

Tahap ini adalah dimana menggunakan *standardscaler* untuk menormalisasi x_{train} dan x_{test} , sehingga Tabel 6 adalah perbedaan nilai sebelum dan sesudah di lakukan proses *standardscaler* pada x_{train} .

Tabel 6. Nilai *standardscaler* x_{train}

Sebelum	Sesudah
[[19 5]	[[-0.93406066 1.43462336]
.	.
.	.
.	.

[29 12]]	[-1.74962108 0.29821983]]
------------	--------------------------------

Tabel 6 adalah perbedaan nilai sebelum dan sesudah di lakukan proses *standardscaler* pada x_{test} .

Tabel 7. Nilai *standardscaler* x_{test}

Sebelum	Sesudah
[[20 9]	[[0.59947176 1.32098301
.	.
.	.
.	.
[48 9]]	[1.86094686 0.59947176]]

3.4 Data Mining

Selanjutnya memanggil perintah pada Tabel 8 yang diproses menggunakan bahasa pemrograman *python* dengan bantuan *Google Collabs* dimana perintah tersebut berfungsi untuk mengaktifkan atau memanggil klasifikasi *Naïve Bayes* dengan menentukan hasil dari perkiraan data *testing*.

Tabel 8. Perintah *Naïve Bayes*

```
from sklearn.naive_bayes import GaussianNB
modelnb = GaussianNB()
nbtrain = modelnb.fit(x_train, y_train)
nbtrain.class_count_
```

Sehingga menghasilkan klasifikasi pada Tabel 8 yang menghasilkan data *testing* prediksi yang kemudian akan dibandingkan dengan Tabel 8 yang merupakan y_{test} atau data *testing* asli.

Tabel 9. Nilai y_{test} prediksi

```
[[0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0,
0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 1, 1]]
```

Tabel 10. Nilai y_{test} asli

([1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0,
0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
0, 0, 1, 0, 1, 1])

Bisa dilihat bahwa nilai pada Tabel 9 dan Tabel 10 memiliki perbedaan, dimana ada 7 data yang berbeda.

3.5 Evaluation

Tahap *evaluation* akan menggunakan bantuan *Confusion matrix* dimana akan mengukur suatu kinerja algoritma berupa akurasi, *recall*, *f1score*, *support* dan *precision* dari algoritma *Naïve Bayes*. Tabel 11 merupakan nilai yang didapatkan dari proses *Confusion matrix*.

Tabel 11. Hasil Confusion Matrix

Diklasifikasikan Sebagai		
	TRUE	FALSE
TRUE	27	0
FALSE	7	8

Hasil penelitian dari Klasifikasi Minat Beli Pelanggan Terhadap Udang *Vaname* Menggunakan Algoritma *Naïve Bayes* yang diukur menggunakan *confusion matrix*.

- Rumus menghitung nilai *precision* :

$$27/(27+0) = 100\% \quad (2)$$
- Rumus menghitung *recall* =

$$27/(27+8) = 77\% \quad (3)$$
- Rumus menghitung nilai akurasi =

$$27+8/(27+8+0+7) = 83\% \quad (4)$$

Dapat dilihat pada Gambar 1 adalah hasil dari *classification report*.

	precision	recall	f1-score	support
0	0.79	1.00	0.89	27
1	1.00	0.53	0.70	15
accuracy			0.83	42
macro avg	0.90	0.77	0.79	42
weighted avg	0.87	0.83	0.82	42

Gambar 1. *Classification Report*

3. SIMPULAN

Berdasarkan dari penjabaran yang sudah dijelaskan diatas dapat disimpulkan bahwa penggunaan algoritma *Naïve Bayes* dapat digunakan untuk mengetahui Minat Beli Pelanggan Terhadap Udang *Vaname*. Berdasarkan analisis data yang sudah dilakukan menggunakan 20% data *testing* yang sudah diberi label positif dan negatif dan dapat dibuktikan bahwa menghasilkan nilai akhir akurasi 83% dengan *precision* 100% dan *recall* 77%.

DAFTAR RUJUKAN

- [1] Adipu, Y. (2019). Profil Kualitas Air Pada Budidaya Udang *Vaname* (*Litopenaeus vannamei*) Sistem Bioflok Dengan Sumber Karbohidrat Gula Aren. *Jurnal MIPA*, 8(3), 122. <https://doi.org/10.35799/jmuo.8.3.2019.25967>
- [2] Anwar, S., & Abdurrohman, A. (2020). Pemanfaatan Teknologi Internet of Things Untuk Monitoring Tambak Udang *Vaname* Berbasis Smartphone Android Menggunakan Nodemcu Wemos D1 Mini. *Infotronik: Jurnal Teknologi Informasi Dan Elektronika*, 5(2), 77. <https://doi.org/10.32897/infotronik.2020.5.2.484>
- [3] Dewi, S. M., Windarto, A. P., & Hartama, D. (2019). Penerapan Datamining Dengan Metode Klasifikasi Untuk Strategi Penjualan Produk Di Ud.Selamat Selular. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 617–621. <https://doi.org/10.30865/komik.v3i1.1669>
- [4] Eshardiansyah, R. P., Sulistiyowati, N., & Jajuli, M. (2021). Algoritma C4 . 5 Untuk Klasifikasi Jenis Kekerasan pada Anak (

- Kasus DP3A Kabupaten Karawang).
Jurnal Sains Komputer & Informatika (J-SAKTI), 5(September), 687–696.
<http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/368>
- [5] Lukhayu Pritalia, G. (2018). Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce. *Indonesian Journal of Information Systems*, 1(1), 47–56.
<https://doi.org/10.24002/ijis.v1i1.1727>
- [6] Naibaho, H., Suyoto, Y. T., & Barata, D. D. (2020). Strategi Penjualan Pedagang Pasar Modern Berbasis Custome Data Mining. *Jurnal Terapan Abdimas*, 5(1), 1.
<https://doi.org/10.25273/jta.v5i1.4479>
- [7] Pratiwi, R. P., Tazro, I., & Juliane, C. (2022). Penerapan Algoritma Naïve Bayes untuk Mengidentifikasi Strategi Marketing dalam Penjualan Deposit E-Money. *XIII(1)*, 65–72.
- [8] Ridwan, A. (2020). Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 4(1), 15–21.
<https://doi.org/10.47970/siskom-kb.v4i1.169>
- [9] Simanjuntak, H. E., & Windarto, W. (2020). Analisa Data Mining Menggunakan Frequent Pattern Growth pada Data Transaksi Penjualan PT Mora Telematika Indonesia untuk Rekomendasi Strategi Pemasaran Produk Internet. *Jurnal Media Informatika Budidarma*, 4(4), 914–923.
<https://doi.org/10.30865/mib.v4i4.2300>
- [10] Utami, D. R., Kurniawan, A., & Enri, U. (2021). Naïve Bayes Klasifikasi untuk Rekomendasi Strategi Promosi Tindakan pada Aesthetic Dental Clinic Karawang. 5(September), 601–610.
- [11] Wahyudi, I., Bahri, S., & Handayani, P. (2019). Aplikasi Pembelajaran Pengenalan Budaya Indonesia. *V(1)*, 135–138.
<https://doi.org/10.31294/jtk.v4i2>
- [12] Wijaya, H. D., & Dwiasnati, S. (2020). Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat. *Jurnal Informatika*, 7(1), 1–7.
<https://doi.org/10.31311/ji.v7i1.6203>