



**KOMPARASI ALGORITMA NAÏVE BAYES DAN K-NEAREST
NEIGHBOR UNTUK MENENTUKAN KLASIFIKASI PRODUK
TERLARIS (STUDI KASUS: PERUSAHAAN FROZEN FOOD XYZ)**

TUGAS AKHIR

Sena Melianta Aji
41518010074

UNIVERSITAS
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022**



**Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk
Menentukan Klasifikasi Produk Terlaris
(Studi Kasus: Perusahaan Frozen Food XYZ)**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Sena Melianta Aji
41518010074

UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518010074

Nama : Sena Melianta Aji

Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan di dalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.



Jakarta, 04 Juli 2022



(Sena Melianta Aji)

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Sena Melianta Aji
NIM : 41518010074
Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non Eksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 04 Juli 2022

UNIVERSITAS
MERCU BUANA



(Sena Melianta Aji)

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Sena Melianta Aji
 NIM : 41518010074
 Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi	✓
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal : Jurnal Ilmu Teknik dan Komputer		
	ISSN : 2621-1491		
	Link Jurnal : https://publikasi.mercubuana.ac.id/index.php/jitkom		
	Link File Jurnal Jika Sudah di Publish :		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui
 Dosen Pembimbing TA


 Raka Yusuf, ST, MTI

Jakarta, 04 Juli 2022



Sena Melianta Aji

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010074
Nama : Sena Melianta Aji
Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022



(Misbahul Fajri, M.TI)
Dosen Penguji 1

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010074
Nama : Sena Melianta Aji
Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022



(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Dosen Penguji 2

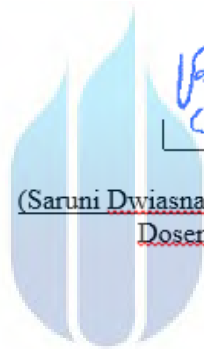
UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010074
Nama : Sena Melianta Aji
Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan Frozen Food XYZ)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022



Pani
(Runi)

(Saruni Dwiasnati, ST, MM, M.Kom)
Dosen Penguji 3

□

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM : 41518010074
Nama : Sena Melianta Aji
Judul Tugas Akhir : Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, Juli 2022

Menyetujui,

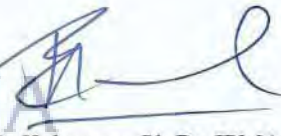


(Raka Yusuf, ST, MTI)
Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan kehadirat Allah swt, yang telah melimpahkan rahmat dan karunia nya, sehingga pada akhirnya penulis dapat menyelesaikan tugas akhir ini dengan baik, Adapun judul tugas akhir yang penulis buat adalah “Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)”

Penulis laporan tugas akhir ini diajukan untuk memenuhi salah satu syarat kelulusan dalam memperoleh gelar strata 1 (satu) pada Universitas Mercu Buana. Dalam penulisan tugas akhir ini tidak lepas dari hambatan dan kesulitan, namun berkat bimbingan, nasihat dan saran serta kerjasama dari berbagai pihak, khusus, pembimbing, segala hambatan tersebut dapat diatasi dengan baik.

Dengan tidak mengurangi apresiasi kepada semua pihak yang telah membantu penulisan secara khusus, baik langsung maupun tidak langsung penulis menyampaikan rasa hormat dan ucapan terima kasih kepada:

1. Kedua Orang Tua yang telah banyak memberikan dukungan baik secara moril maupun materi dan masukan dalam pembuatan tugas akhir ini.
2. Ir.Emil R. Kaburuan, Ph.D.,IPM. selaku kepala program studi Teknik Informatika Universitas Mercu Buana.
3. Wawan Gunawan, S.Kom, M.Kom. selaku Koordinator Tugas Akhir program studi Teknik Informatika Universitas Mercu Buana.
4. Raka Yusuf, ST, MTI. selaku dosen pembimbing saya yang telah membantu dan memberikan arahan kepada saya dalam menyelesaikan tugas akhir ini.
5. Drs. Achmad Kodar, MT selaku dosen pembimbing akademik saya yang telah membimbing selama perkuliahan hingga sampai tugas akhir.
6. Sahabat dan kerabat, yang selalu support dan mendoakan saya dalam menyelesaikan Tugas Akhir ini dan mendapatkan gelar sarjanah.
7. Semua pihak yang membantu dalam menyelesaikan Tugas Akhir ini yang tidak bisa saya sebut satu persatu hingga dapat selesai dengan baik.

Tiada kata lain yang dapat penulis ungkapkan untuk mengucapkan terima kasih terhadap semua pihak yang telah banyak membantu dalam menyelesaikan laporan ini dan semoga Allah swt selalu melimpahkan rahmat dan hidayah-Nya kepada kita semua, Aamiin. Harapn Penulis yaitu semoga Tugas Akhir ini dapat bermanfaat dengan baik dan berguna bagi orang lain.

Jakarta, Juni 2022

Sena Melinta Aji



DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... iii	
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT.....	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
NASKAH JURNAL	1
KERTAS KERJA.....	8
BAB 1. LITERATUR REVIEW.....	9
BAB 2. ANALISA DAN PERANCANGAN.....	16
BAB 3. SOURCE CODE.....	17
BAB 4. DATASET.....	30
BAB 5. TAHAP EKSPERIMEN.....	32
BAB 6. HASIL SEMUA EKPERIMEN	35
DAFTAR PUSTAKA	41
LAMPIRAN DOKUMEN HAKI.....	44
LAMPIRAN KORESPONDENSI	46

NASKAH JURNAL

Jurnal Ilmu Teknik dan Komputer
Vol. X No. X Jan 2017

Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan Frozen Food XYZ)

Sena Melianta Aji¹, Raka Yusuf²

Informatika, Fakultas Ilmu Komputer, Universitas Mercubuana
Jl. Raya Meruya Selatan, Kembangan, Jakarta, Indonesia 11650
Senaaji2000@gmail.com¹, raka@mercubuana.ac.id²

ABSTRACT

Penjualan pada frozen food di kalangan masyarakat semakin meluas karena kehidupan masyarakat saat lebih memilih makanan yang praktis dan cepat saji. Salah satu yang dapat kita manfaatkan adalah teknologi dalam menggali informasi yang bermanfaat dari gudang data perusahaan penjualan frozen food. Strategis yang dapat dilakukan adalah dengan mempelajari pola perilaku konsumen. Pola tersebut dapat diketahui dengan memanfaatkan data transaksi penjualan di perusahaan PT. *Frozen Food XYZ*. Demikian besar transaksi harian yang terus bertambah. Transaksi yang banyak tersebut akan mempersulit pelaku perusahaan dalam mengolah data mereka. Tujuan penelitian ini adalah mencoba menerapkan teknik metode algoritma naïve bayes dan k-nearest neighbors memberikan informasi berupa klasifikasi penjualan produk frozen food yang paling laris dikalangan masyarakat dan tidak laris dikalangan masyarakat (laris dan tidak laris). Dari rata – rata accuracy yang ada dapat dilihat bahwa nilai rata rata accuracy algoritma Naïve Bayes adalah 77%, sedangkan algoritma K-Nearest Neighbor nilai rata – rata accuracy sebesar 99%.

Kata Kunci: Klasifikasi, Naive Bayes, KNN

PENDAHULUAN

Perkembangan frozen food di pusat perbelanjaan bermunculan dengan berbagai macam bentuk dan ukuran yang menyebabkan persaingan dalam dunia ritel semakin ketat. Bisnis atau biasa disebut dengan perusahaan semakin terasa keberadaannya dalam kehidupan masyarakat sehari-hari. Beberapa contoh bentuk pusat perbelanjaan yang meramalkan dunia ritel seperti minimarket, supermarket. Perusahaan yang bergerak dalam bidang ritel dengan konsep menjual produk

makanan (*food*) dari segi pemasaran memiliki peran penting sebagai salah satu lembaga pemasaran dalam pemasaran produk. Dalam fungsi-fungsi inti pemasaran ini dilakukan upaya-upaya yang meliputi peningkatan penjualan hasil produk dan menciptakan semakin banyak pelanggan dan pemasaran sangat penting untuk menjamin perkembangan perusahaan dimasa yang akan datang.

Salah satu strategis yang dapat dilakukan adalah dengan mempelajari pola perilaku konsumen.(Wiyanto, 2019) Pola tersebut dapat diketahui dengan

memanfaatkan data transaksi penjualan di perusahaan PT. *Frozen Food XYZ*. Demikian besar transaksi harian yang terus bertambah. Transaksi yang banyak tersebut akan mempersulit pelaku perusahaan dalam mengolah data mereka.

Prioritas pada peningkatan penjualan dengan menarik konsumen baru atau berorientasi pada upaya mempertahankan pelanggan yang telah ada. Namun, perlu diketahui bahwa konsumen baru tersebut tidak bisa menjanjikan akan terus menggunakan produk perusahaan dengan demikian pelaku bisnis ritel lebih baik mempertahankan pelanggan yang telah ada karena mampu memberikan keuntungan yang lebih besar dari perusahaan. (Umiyati, 2021)

Dalam proses penelitian ini akan menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor. Algoritma Naïve Bayes merupakan suatu metode independen yang membahas mengenai klasifikasi sederhana berdasarkan teori pada bayes. (Wijaya & Dwiasnati, 2020) Algoritma ini tergolong cepat dan sederhana dengan menggunakan fitur yang dapat berdiri sendiri. Sedangkan K-Nearest Neighbor merupakan klasifikasi yang menyimpan semua *data training* dan melakukan klasifikasi dengan cara membandingkan antara atribut yang paling cocok dengan atribut *record* dan *data training*. (Alfani W.P.R. et al., 2021) Dengan keunggulan masing-masing algoritma diatas penelitian ini menggunakan kedua algoritma

untuk implementasi dalam mengklasifikasi data penjualan *frozen food*.

RUMUS MASALAH

Berdasarkan latar belakang diatas, maka rumusan masalah di dalam penelitian ini yaitu:

1. Bagaimana implementasi algoritma Naïve Bayes dan algoritma K-Nearest Neighbor dalam mengklasifikasikan terlaris dan tidak terlaris?
2. Algoritma apa yang lebih cocok antara algoritma Naïve Bayes atau algoritma K-Nearest Neighbor dalam menentukan terlaris atau tidak terlaris nya produk tersebut?

A. Tujuan

Tujuan pada penelitian ini yaitu:

1. Menentukan faktor-faktor yang mempengaruhi dalam menentukan mengklasifikasikan terlaris dan tidak terlaris dengan menggunakan algoritma naïve bayes dan algoritma K-Nearest Neighbor.
2. Mengklasifikasikan terlaris dan tidak terlaris di *frozen food*.
3. Menentukan algoritma yang paling cocok dalam penelitian terlaris dan tidak terlaris.

STUDI LITERATUR

A. Klasifikasi

Klasifikasi merupakan suatu metode untuk mengelompokkan sebuah objek ke dalam kelompok atau kelas. Berdasarkan nilai-nilai yang ditentukan dengan cara mengambil data yang terdiri dari kelas. Teknik ini dapat memberikan klasifikasi pada data baru dengan

memanipulasi data yang telah diklasifikasikan dan dengan menggunakan hasil untuk memberikan jumlah atribut. Dimana terdiri dari sekumpulan data yang masing-masing sudah dikelompokkan kelas tertentu, (Odi Nurdiawan, 2018) masalah klasifikasi yang diperhatikan yaitu dengan penentuan aturan yang memungkinkan adanya klasifikasi walaupun data yang ada belum menunjukkan hal tersebut.

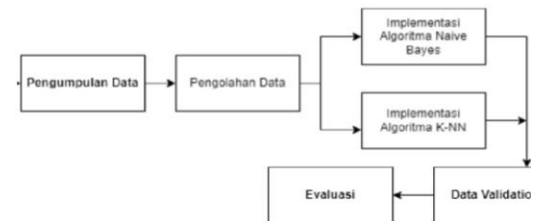
B. Algoritma Naïve Bayes

Algoritma naïve bayes adalah metode probabilistik pengklasifikasian sederhana berdasarkan teorema bayes dimana pengklasifikasian dilakukan melalui training set sejumlah data secara efisien. (Romli et al., 2019) Naïve bayes mengasumsikan bahwa nilai dari sebuah input atribut pada kelas yang diberikan tidak tergantung dengan nilai pada data tersebut (Terlaris & Penjualan, 2020).

C. Algoritma K-NN

Proses yang dilakukan dalam algoritma K-Nearest Neighbor merupakan dataset yang digunakan akan dijadikan sebagai data pembelajaran dimana terdapat label ya atau tidak. (Handoko et al., 2021) Lalu tentukan jumlah tetangga terdekat. Setelah menemukan K maka data testing akan dihitung jarak terhadap data training tersebut. Lalu dilihat jarak paling dekat dengan K yang telah ditentukan ((A. Elsyia Aprilia, 2019), lalu kelompok data testing tersebut berdasarkan label mayoritas pada

METODOLOGI



1. Pengumpulan Data

Peneliti melakukan pengumpulan berupa data di perusahaan PT. *Frozen Food XYZ*. Data yang diambil yaitu data penjualan di perusahaan PT. *Frozen Food XYZ*. Pada awal tahun 2021 yang ditarik melalui database. Database yang diberikan yaitu berjumlah 1.213 record yang dimana 9 atribut Nama Customer, Tanggal, Item, Jml, Satuan, Category, Harga (1 Kg), Qty (Kg), Total Harga.

2. Pengolahan Awal Data (Pre-Processing)

Data yang telah di peroleh di lakukan pre-processing. Preprocessing data adalah hal yang harus dilakukan dalam proses data mining, karena tidak semua data atau atribut data dalam data digunakan dalam proses data mining. Proses ini dilakukan agar data yang akan digunakan sesuai dengan kebutuhan. Terdapat beberapa jenis preprocessing yang akan digunakan dalam penelitian ini, antara lain:

a Data Cleaning

Pada tahap data cleaning dilakukan pembersihan data dari *missing values* dan duplikat yang ada pada data penjualan di PT. *Frozen Food XYZ*. *Missing values* adalah kondisi dimana suatu atribut memiliki nilai

yang tidak valid. *Processing/Cleaning* merupakan proses pembersihan yang membuang duplikat data, memeriksa data yang konsisten, dan memperbaiki kesalahan pada data. Pengecekan *missing values* dilakukan pada keseluruhan dataset, sedangkan pengecekan duplikat dilakukan pada data per tahun.

b. Data Transformation

Pada tahap ini dilakukan untuk mengubah tipe data pada kolom date dari yang semula string (object) menjadi tipe data pada kolom date. Setelah mengubah tipe data pada kolom status produk dari yang semula string (object) menjadi integer. Melakukan tahapan mengubah tipe data pada kolom "Harga (1 Kg)" dari yang semula float menjadi integer dan juga kolom "Status Item" dari yang semula object menjadi integer.

c. Data Reduction

Pada tahap ini penelitian dilakukan reduksi data dari sejumlah dataset dengan menghilangkan kolom pada penelitian ini tidak semua atribut digunakan untuk kolom No. Satuan pada penelitian ini dihapus.

3. Implementasi Algoritma Naïve Bayes Dan K-NN

Implementasikan menggunakan 2 algoritma yaitu algoritma Naïve Bayes dan K-NN. Dalam implementasikan nilai x dan y ditentukan yaitu variabel independen dan dependen. Nilai x variabel yang digunakan adalah Jumlah Item Terjual, Harga (1 Kg), dan

Total Harga dan nilai y yang digunakan Status item.

a. Algoritma Naïve Bayes

Proses yang dilakukan dalam algoritma Naïve Bayes merupakan data yang dimasukan akan di hitung probabilitas dari setiap kelas. Setelah didapat maka hasil kelas akan dibandingkan. Jika nilai yang lebih besar maka itu Terlaris dan jika nilai lebih kecil maka tidak laris

b. Algoritma K-NN

Proses yang dilakukan dalam algoritma K-Nearest Neighbor merupakan dataset yang digunakan akan dijadikan sebagai data pembelajaran dimana terdapat label ya atau tidak. Lalu tentukan jumlah tetangga terdekat. Setelah menemukan K maka data testing akan dihitung jarak terhadap data training tersebut. Lalu dilihat jarak paling dekat dengan K yang telah ditentukan, lalu kelompok data testing tersebut berdasarkan label mayoritas pada

4. Data Validation

validasi model digunakan teknik *Split Validation* yaitu teknik validasi yang membagi data menjadi dua bagian secara acak, sebagai *data training* dan sebagai lainnya secara *data testing*

5. . Evaluasi

Evaluasi diperlukan untuk menganalisis dan mengukur sejauh mana peninjauan model mengenai hasil evaluasi yang telah dikerjakan. Akan di lakukan perbandingan kuantitatif dengan pertimbangan nilai komparasi. Sedang untuk model digunakan

teknik *confusion Matrix*. Lalu hasil evaluasi akan diberikan dengan classification report yang terdiri dari Accuracy, Precision dan Recall.

1. Accuracy

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision

$$precision = \frac{TP}{TP + FP}$$

3. Recall

$$recall = \frac{TP}{TP + FN}$$

HASIL DAN DISKUSI

Pada bagian ini menampilkan hasil semua eksperimen yang telah dilakukan.

Eksperimen yang dilakukan pada penelitian ini menggunakan split validation pada pembagian data yaitu data training 70% dan data testing 30%.

A. Algoritma naïve bayes

hasil accuracy, precision dan recall

Tabel .1 Hasil Algoritma Naïve Bayes

	Accuracy	Precision	Recall
1	77%	76%	100%
2	77%	100%	14%

Confusion Matrix

```
-----Nilai Perbandingan Performa Naive Bayes terhadap
Data Training-----
[[618  0]
 [198 31]]
```

Gambar 1. Confusion Matrix

Berikut adalah penjelasan dari hasil output diatas terkait nilai perbandingan TP, TN, FP, dan FN di dalam confusion matrix dari algoritma Naive Bayes:

- True positive (TP) merupakan data positif yang berhasil diklasifikasikan dengan tepat. Jumlah true positive pada pengujian kali ini adalah 618.
- True negative (TN) merupakan data negatif yang berhasil diklasifikasikan dengan benar. Jumlah true negative pada algoritma naive bayes yang berhasil diklasifikasikan dengan benar mencapai 31.
- False positive (FP) merupakan data negatif namun diklasifikasikan sebagai data positif. Jumlah nilai false positive pada pengujian kali ini adalah 0.
- False negative (FN) merupakan data positif namun diklasifikasikan sebagai data negatif. Jumlah false negative di algoritma naive bayes yaitu mencapai 198.

B. Algoritma K-Nearest neighbor

hasil accuracy, precision dan recall

Tabel 2. Hasil Algoritma KNN

	Accuracy	Precision	Recall
1	99%	99%	100%
2	99%	99%	98%

Confusion Matrix

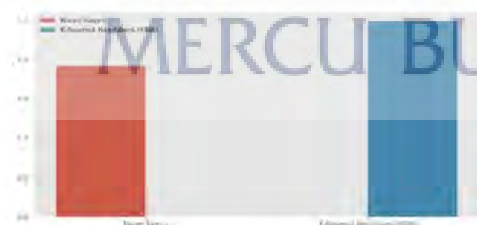
```
-----Nilai Perbandingan Performa KNN terhadap Data
Training-----
[[615  3]
 [ 4 225]]
```

Gambar 2. Confusion Matrix

Berikut adalah penjelasan dari hasil output diatas terkait nilai perbandingan TP, TN, FP, dan FN di dalam confusion matrix dari algoritma Naive Bayes:

- True positive (TP) merupakan data positif yang berhasil diklasifikasikan dengan tepat. Jumlah true positive pada pengujian ini adalah 615.
- True negative (TN) merupakan data negatif yang berhasil diklasifikasikan dengan benar. Jumlah true negative pada algoritma KNN yang berhasil diklasifikasikan dengan benar mencapai 225.
- False positive (FP) merupakan data negatif namun diklasifikasikan sebagai data positif. Jumlah nilai false positive pada pengujian ini adalah 3.
- False negative (FN) merupakan data positif namun diklasifikasikan sebagai data negatif. Jumlah false negative di algoritma KNN yaitu mencapai 4.

C Grafik kedua Algoritma Naive Bayes dan K-Nearest Neighbors



Gambar .3 grafik algoritma Naive Bayes dan KNN

Berdasarkan perbandingan tingkat akurasi data training melalui grafik batang diatas, saya dapat mengetahui dan mengambil kesimpulan diantaranya adalah sebagai berikut:

1. Nilai akurasi di dalam data training terhadap algoritma Naive Bayes adalah sebesar 0.77 atau 77% yang menjadikan algoritma Naive Bayes di data training ini memiliki hasil klasifikasi/prediksi yang tepat tetapi jika dibandingkan dengan grafik batang di hasil akurasi algoritma K-Nearest Neighbors, hasil akurasi naive bayes terkhusus untuk data training masih tidak cukup tinggi dan kurang tepat.
2. Sedangkan nilai akurasi di dalam data training terhadap algoritma K-Nearest Neighbors (KNN) sebesar 0.99 atau 99%. Menurut saya akurasi tersebut sudah sangat tinggi dan tepat dalam mengklasifikasikan item laris dan tidak laris selama bulan Januari - September 2021

KESIMPULAN

Berdasarkan hasil penelitian ini semua percobaan penelitian yang sudah dilakukan mengklasifikasn dari dua algoritma. Menggunakan data yang memanfaatkan data transaksi penjualan di perusahaan PT. *Frozen Food XYZ*. Demikian besar transaksi harian yang terus bertambah, Sedangkan algoritma naive bayes nilai rata-rata accuracynya 77% lebih rendah membuat performa naive bayes di bawah algoritma KNN. Naive bayes akan

berjalan baik pada data yang bersifat independen. Sedangkan data yang diolah memiliki variabel yang dependent. Sedangkan, Algoritma K-NN nilai rata – rata accuracynya 99% lebih tinggi dibandingkan algoritma Naive Bayes hal ini karena dari karakteristik data, algoritma K-NN dapat bekerja dengan baik setelah data melalui preprocessing dan menghasilkan data nominal. Keterkaitan data satu sama lain membuat performa K-NN meningkat.

Dengan dilakukan penelitian ini dapat memberikan informasi bermanfaat mengenai penjualan frozen food dan algoritma yang lebih unggul KNN dari algoritma Naive Bayes.

DAFTAR PUSTAKA

- A. Elysia Aprilia. (2019). *Klasifikasi Penjualan Sepeda Motor Honda Menggunakan Metode K-Nearest Neighbor*.
- Alfani W.P.R., A., Rozi, F., & Sukmana, F. (2021). Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor. *JIP1 (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 6(1), 155–160. <https://doi.org/10.29100/jipi.v6i1.1910>
- Handoko, D., Tambunan, H. S., & Hardinata, J. T. (2021). Analisis Penjualan Produk Paket Kuota Internet Dengan Metode K-Nearest Neighbor. *Jurassic (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 6(1), 111. <https://doi.org/10.30645/jurasik.v6i1.275>
- Odi Nurdiawan, S. N. (2018). Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran. *Jurnal Teknologi Informasi Dan Komunikasi, April*, 84–95.
- Romli, I., Pusnawati, E., & Bangsa, U. P. (2019). Penentuan Tingkat Penjualan Mobil Di Indonesia Dengan Menggunakan Algoritma Naive Bayes. *Prosiding Seminar Nasional Teknologi Dan Sains (SNasTekS)*, x(x).
- Terlaris, P., & Penjualan, P. (2020). Penerapan Algoritma Naive Bayes Untuk Menentukan Klasifikasi Produk Terlaris Pada Penjualan Pulsa. *10*, 195–207.
- Umiyati. (2021). Implementasi Metode Klasifikasi Algoritma Naive Bayes Dalam Menganalisis Transaksi Penjualan Pada 212 Mart Kuto Palembang. *4*(1), 6.
- Wijaya, H. D., & Dwiasnati, S. (2020). Implementasi Data Mining dengan Algoritma Naive Bayes pada Penjualan Obat. *Jurnal Informatika*, 7(1), 1–7. <https://doi.org/10.31311/ji.v7i1.6203>
- Wiyanto. (2019). Analisa Tingkat Kepuasan Pelanggan Terhadap Pelayanan Perusahaan Otobus XYZ Menggunakan Metode Naive Bayes. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699. <http://krsti.stmik-amik-riau.ac.id/index.php/satin/article/view/666>

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “Komparasi Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Menentukan Klasifikasi Produk Terlaris (Studi Kasus: Perusahaan *Frozen Food XYZ*)”. Kertas kerja berisi semua material hasil penelitian Tugas Akhir yang tidak dimuat/atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: literature review, dataset yang digunakan, source code, dan hasil eksperimen secara keseluruhan.

Bagian 1 : Literature Review menjabarkan mengenai beberapa jurnal yang terkait dengan penelitian.

Bagian 2 : Analisis dan Perancangan menjabarkan analisis masalah terhadap penelitian

Bagian 3 : Source Code menjabarkan kumpulan kode disetiap proses yang dilakukan peneliti terdiri dari membaca data, preprocessing data, data visualisasi, data modelling, implementasi algoritma.

Bagian 4 : Pengambilan dataset yang akan dilakukan.

Bagian 5 : Tahapan Eksperimen berisi pengumpulan data, preprocessing data, implementasi algoritma, data validasi, dan evaluasi.

Bagian 6 : Hasil semua eksperimen secara keseluruhan yang mencakup skenario pengujian yang dilakukan oleh peneliti.

