

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK
MENGKLASIFIKASIKAN POTENSI KETERLAMBATAN
PENGIRIMAN PADA EKSPEDISI JNE SARIHARJO**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

FAJAR ANGGA MAULANA

21.12.2008

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK
MENGKLASIFIKASIKAN POTENSI KETERLAMBATAN
PENGIRIMAN PADA EKSPEDISI JNE SARIHARJO**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

FAJAR ANGGA MAULANA

21.12.2008

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK
MENGKLASIFIKASIKAN POTENSI KETERLAMBATAN PENGIRIMAN
PADA EKSPEDISI JNE SARIHARJO**

yang disusun dan diajukan oleh

Fajar Angga Maulana

21.12.2008

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 20 Juni 2025

Dosen Pembimbing,



Yoga Pristyanto, S.Kom., M.Eng
NIK. 190302412

HALAMAN PENGESAHAN
SKRIPSI
PENERAPAN ALGORITMA NAÏVE BAYES UNTUK
MENGLASIFIKASIKAN POTENSI KETERLAMBATAN PENGIRIMAN
PADA EKSPEDISI JNE SARIHARJO

yang disusun dan diajukan oleh

Fajar Angga Maulana

21.12.2008

Telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Juni 2025

Susunan Dewan Penguji

Nama Penguji

Bety Wulan Sari, S.Kom., M.Kom
NIK. 190302254

Atik Nurmasani, S.Kom., M.Kom
NIK. 190302354

Yoga Pristyanto, S.Kom., M.Eng.
NIK. 190302412

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 20 Juni 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fajar Angga Maulana
NIM : 21.12.2008

Menyatakan bahwa Skripsi dengan judul berikut:

Penerapan Algoritma Naive Bayes untuk Mengklasifikasikan Potensi Keterlambatan Pengiriman pada Ekspedisi JNE Sariharjo

Dosen Pembimbing : Yoga Pristyanto, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 20 Juni 2025

Yang Menyatakan,



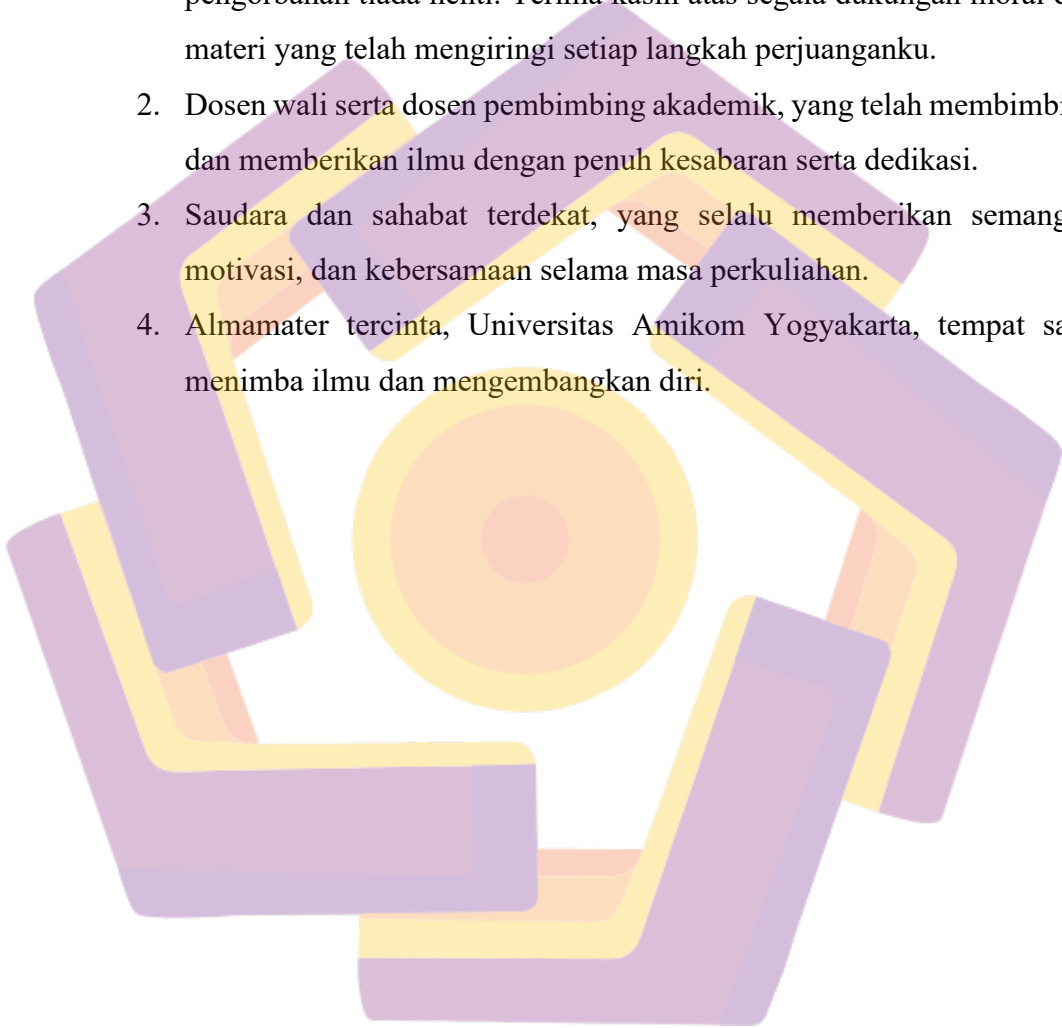
METERAI
TEMPER
01AMX363977250

Fajar Angga Maulana

HALAMAN PERSEMBAHAN

Dengan segala rasa syukur dan kerendahan hati, karya tulis ini penulis persembahkan kepada:

1. Ayah dan Ibu tercinta, yang telah memberikan cinta, doa, semangat, dan pengorbanan tiada henti. Terima kasih atas segala dukungan moral dan materi yang telah mengiringi setiap langkah perjuanganku.
2. Dosen wali serta dosen pembimbing akademik, yang telah membimbing dan memberikan ilmu dengan penuh kesabaran serta dedikasi.
3. Saudara dan sahabat terdekat, yang selalu memberikan semangat, motivasi, dan kebersamaan selama masa perkuliahan.
4. Almamater tercinta, Universitas Amikom Yogyakarta, tempat saya menimba ilmu dan mengembangkan diri.



KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik. Skripsi yang berjudul "Penerapan Algoritma Naïve Bayes untuk Mengklasifikasikan Potensi Keterlambatan Pengiriman pada Ekspedisi JNE Sariharjo" disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penulis menyadari bahwa proses penyusunan skripsi ini tidak terlepas dari dukungan, bimbingan, serta bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa hormat dan terima kasih yang sebesar-besarnya kepada:

1. *Prof. Dr. M. Suyanto, M.M* selaku Rektor Universitas Amikom Yogyakarta
2. *Hanif Al Fatta, S.Kom., M.Kom., Ph.D.,* selaku Dekan Program Fakultas Ilmu Komputer
3. *Anggit Dwi Hartanto, S.Kom., M.Kom.* selaku Ketua Program Studi Sistem Informasi
4. *Yoga Pristyanto, S.Kom., M.Eng.,* selaku dosen pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. *Kedua orang tua dan keluarga besar penulis,* yang selalu mendoakan dan mendukung secara moral dan spiritual dalam setiap langkah perjuangan akademik ini.
6. Seluruh teman-teman dan sahabat seperjuangan yang telah memberikan semangat, inspirasi, dan kebersamaan selama masa perkuliahan dan penyusunan tugas akhir ini.

Yogyakarta, 20 Juni 2025

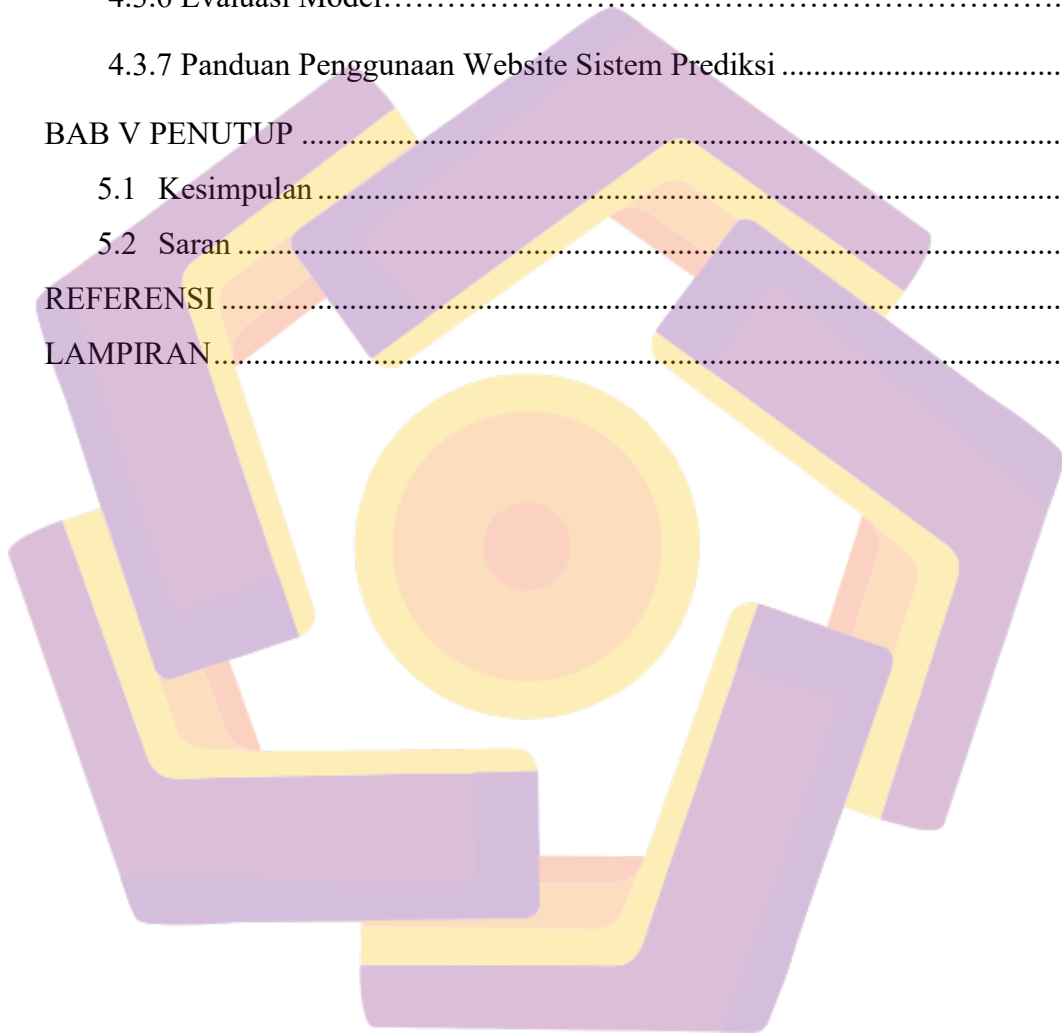
Fajar Angga Maulana

DAFTAR ISI

PENERAPAN ALGORITMA NAÏVE BAYES UNTUK MENGKLASIFIKASIKAN POTENSI KETERLAMBATAN PENGIRIMAN PADA EKSPEDISI JNE SARIHARJO HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur.....	6
2.2 Dasar Teori	15
2.2.1 Klasifikasi.....	15
2.2.2 Data mining.....	15

2.2.3 Preprocessing data.....	16
2.2.4 Naïve bayes.....	17
2.2.5 Confusion Matrix.....	18
2.2.6 Precision & Recall.....	19
2.2.7 F1-score.....	19
2.2.8 Accuracy (Akurasi).....	19
BAB III METODE PENELITIAN	20
3.1 Objek Penelitian.....	20
3.2 Alur Penelitian	21
3.2.1 Pengumpulan Data.....	23
3.2.2 Preprocessing Data.....	23
3.2.3 Pemisahan Data (Data Splitting: Training & Testing).....	24
3.2.4 Penerapan Metode (Model Implementation)	24
3.2.5 Evaluasi Model.....	25
3.3 Alat dan Bahan.....	26
3.3.1 Data Penelitian.....	26
3.3.2 Alat Penelitian.....	27
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Pengumpulan Data	28
4.2 Text Preprocessing	29
4.2.1 Pelabelan Data.....	29
4.2.2 Cleaning Data.....	31
4.2.3 Handling Missing Values.....	32
4.2.4 Categorical Encoding.....	34
4.3 Penerapan Metode (Model Implementation).....	35
4.3.1 Login dan Register Akun	35

4.3.2 Tampilan Utama Sistem.....	37
4.3.3 Fitur Upload Dataset.....	38
4.3.4 Proses Text Preprocessing Sistem	39
4.3.5 Model Probabilistik Naïve Bayes	45
4.3.6 Evaluasi Model.....	50
4.3.7 Panduan Penggunaan Website Sistem Prediksi	54
BAB V PENUTUP	57
5.1 Kesimpulan	57
5.2 Saran	58
REFERENSI	59
LAMPIRAN.....	62



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	11
Tabel 2. 2 Keaslian Penelitian	12
Tabel 2. 3 Keaslian Penelitian	13
Tabel 2. 4 Keaslian Penelitian	14
Tabel 3. 1 Data Penelitian	26
Tabel 3. 2 Perangkat Keras dan Perangkat Lunak	27
Tabel 4. 1 Hasil Pengumpulan Data	28
Tabel 4. 2 Hasil Pelabelan Data	29
Tabel 4. 3 Hasil Cleaning Data	32
Tabel 4. 4 Data yang memiliki nilai tidak valid	33
Tabel 4. 5 Data yang memiliki nilai valid	33
Tabel 4. 6 Categorical Encoding	34
Tabel 4. 7 Data Setelah Encoding	34
Tabel 4. 8 Evaluasi model	51

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	22
Gambar 4. 1 Distribusi kelas.....	30
Gambar 4. 2 Halaman registrasi.....	36
Gambar 4. 3 Halaman Login.....	37
Gambar 4. 4 Form prediksi pada system.....	37
Gambar 4. 5 Manajemen Dataset.....	38
Gambar 4. 6 Halaman Dataset Pengiriman.....	39
Gambar 4. 7 Skrip Pelabelan Data	40
Gambar 4. 8 Skrip Cleaning Data	41
Gambar 4. 9 Skrip Handling Missing Values	42
Gambar 4. 10 Missing Values.....	42
Gambar 4. 11 Skrip Categorical Encoding	43
Gambar 4. 12 Hasil Preprocessing Data	44
Gambar 4. 13 Form user prediksi.....	45
Gambar 4. 14 Kode Proses Likelihood	46
Gambar 4. 15 Skrip Hitung Posterior	46
Gambar 4. 16 Skrip probabilitas untuk fitur numerik.....	47
Gambar 4. 17 Hasil prediksi	48
Gambar 4. 18 evaluasi.....	50
Gambar 4. 19 Skrip Evaluasi Sistem	52
Gambar 4. 20 Hasil evaluasi akurasi, precision, recall, F1-Score	52
Gambar 4. 21 Panduan Penggunaan Web.....	55

DAFTAR LAMPIRAN

Lampiran 1 Surat Jawaban Izin Penelitian.....	62
Lampiran 2 Surat Penyerahan Hasil Skripsi	63



DAFTAR LAMBANG DAN SINGKATAN

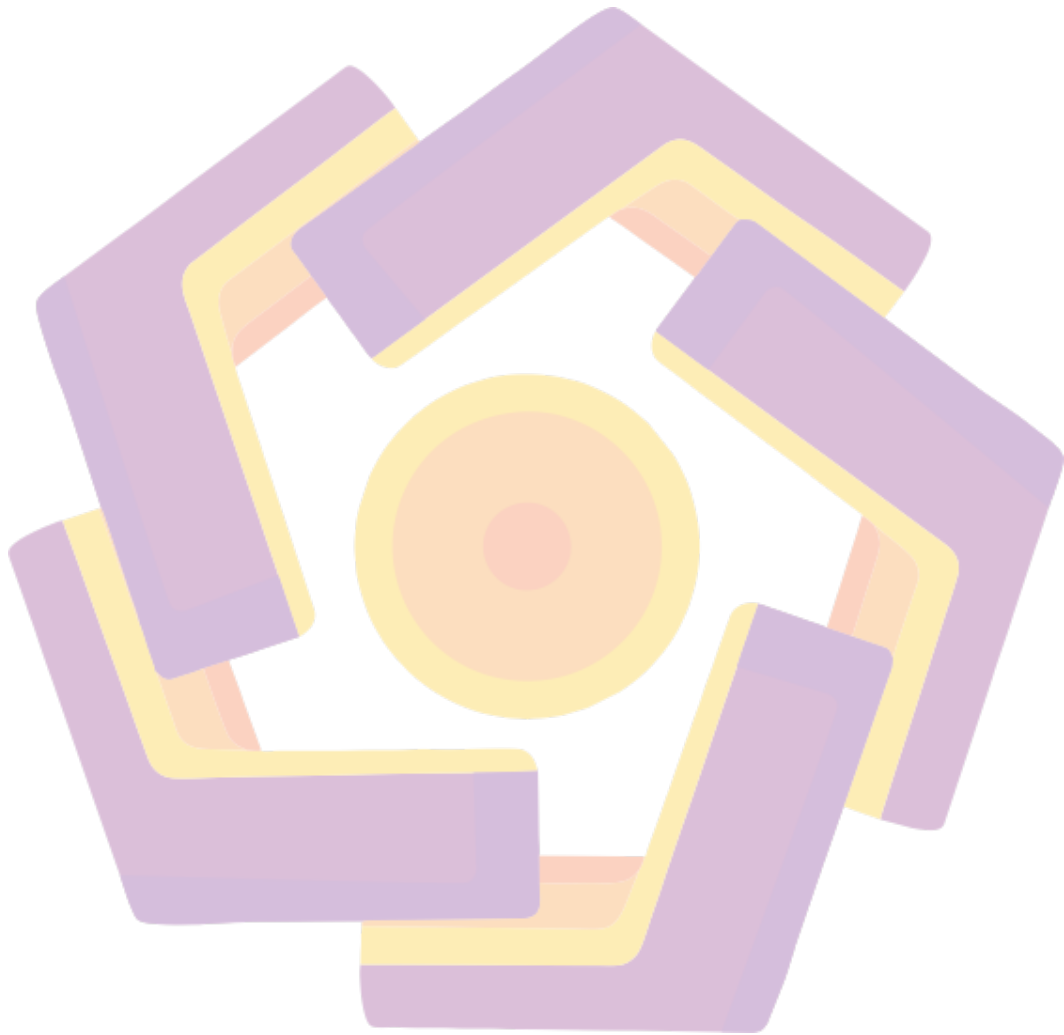
TP	:	True Positive – Prediksi benar untuk kelas positif
FP	:	False Positive – Prediksi salah, model memprediksi positif padahal sebenarnya negatif
TN	:	True Negative – Prediksi benar untuk kelas negatif
FN	:	False Negative – Prediksi salah, model memprediksi negatif padahal sebenarnya positif
PHP	:	Hypertext Preprocessor – Bahasa pemrograman untuk pengembangan web
MySQL	:	Structured Query Language – Sistem manajemen basis data relasional
F1-Score	:	Rata-rata harmonis dari precision dan recall
REG	:	Regular – Jenis layanan pengiriman JNE
YES	:	Yakin Esok Sampai – Layanan cepat dari JNE
OKE	:	Ongkos Kirim Ekonomis – Layanan ekonomis JNE
JTR	:	JNE Trucking – Layanan untuk paket besar atau berat
CSV	:	Comma Separated Values – Format file untuk data tabular
EDA	:	Exploratory Data Analysis – Proses analisis data awal untuk memahami struktur data
KKN	:	K-Nearest Neighbors – Algoritma klasifikasi berbasis kedekatan data
IPK	:	Indeks Prestasi Kumulatif
SKS	:	Satuan Kredit Semester

DAFTAR ISTILAH

Naïve Bayes	:	Algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dan mengasumsikan independensi antar fitur.
Probabilitas Prior	:	Peluang awal dari sebuah kelas sebelum mempertimbangkan fitur lain
Likelihood	:	Kemungkinan suatu fitur muncul dalam kelas tertentu
Posterior	:	Probabilitas suatu kelas berdasarkan bukti fitur yang diamati
Confusion Matrix	:	Matriks yang menampilkan jumlah prediksi benar dan salah dari model klasifikasi
Precision	:	Rasio prediksi positif yang benar dibandingkan seluruh prediksi positif
Recall	:	Rasio prediksi positif yang benar dibandingkan seluruh kejadian aktual yang positif
F1-Score	:	Rata-rata harmonis dari precision dan recall untuk menyeimbangkan keduanya
Preprocessing	:	Proses pembersihan dan normalisasi data sebelum digunakan dalam pemodelan
Encoding Kategorikal	:	Mengubah data kategori menjadi bentuk numerik agar dapat diproses oleh algoritma
Tokenisasi	:	Memecah teks menjadi bagian-bagian kecil seperti kata atau frasa
Stemming	:	Mengubah kata menjadi bentuk dasarnya dengan menghapus imbuhan
Stopword Removal	:	Menghapus kata-kata umum yang tidak memiliki nilai informasi tinggi
Laplace Smoothing	:	Teknik untuk menghindari probabilitas nol dalam perhitungan Naïve Bayes

Supervised Classification : Klasifikasi dengan data pelatihan berlabel

Unsupervised Classification: Klasifikasi tanpa data label, mencari pola alami dalam data



INTISARI

Meningkatnya volume pengiriman dalam industri logistik menjadikan ketepatan waktu sebagai faktor krusial dalam menjaga kepuasan pelanggan. Penelitian ini bertujuan untuk menerapkan algoritma klasifikasi Naïve Bayes guna memprediksi keterlambatan pengiriman pada JNE Sariharjo, Sleman Yogyakarta. Sistem dikembangkan menggunakan bahasa pemrograman PHP dan basis data MySQL, dengan memanfaatkan data historis pengiriman seperti tujuan, jenis layanan, kategori barang, berat, dan jumlah kiriman.

Tahap preprocessing dilakukan untuk memastikan kualitas data, meliputi pembersihan data, normalisasi, penanganan nilai kosong, dan encoding data kategorikal. Model klasifikasi kemudian menghitung probabilitas prior dan likelihood dari setiap fitur, dengan menerapkan Laplace smoothing untuk menghindari probabilitas nol. Fitur numerik seperti berat dan jumlah barang diproses berdasarkan deviasi terhadap rata-rata untuk mendekati nilai probabilitasnya.

Sistem berhasil mengklasifikasikan data pengiriman ke dalam kategori “Tepat Waktu” dan “Terlambat”. Evaluasi menggunakan confusion matrix menghasilkan nilai akurasi sebesar 91,66%, precision 85,81%, recall 73,89%, dan F1-score 79,40%.. Hasil ini menunjukkan bahwa algoritma Naïve Bayes efektif dalam mengidentifikasi potensi keterlambatan pengiriman. Penelitian ini membuktikan bahwa metode pembelajaran mesin dapat mendukung pengambilan keputusan operasional dalam layanan ekspedisi.

Kata kunci: Naïve Bayes, Prediksi Keterlambatan, Klasifikasi, JNE, Logistik, PHP, MySQL

ABSTRACT

The increasing volume of shipments in the logistics industry has made timely delivery a critical factor in customer satisfaction. This study aims to implement the Naïve Bayes classification algorithm to predict delivery delays in JNE Sariharjo, Sleman Yogyakarta. The system was developed using PHP and MySQL, and utilizes historical shipment data including destination, service type, item category, weight, and quantity.

Data preprocessing was conducted to ensure data quality, involving cleaning, normalization, handling of missing values, and categorical encoding. The classification model then calculated prior probabilities and likelihoods for each feature, with Laplace smoothing applied to avoid zero probability. Numerical features such as weight and quantity were processed using deviation from the mean to approximate likelihood values.

The system successfully classifies shipments into “On Time” and “Delayed” categories. Evaluation using a confusion matrix produced an accuracy of 91.66%, a precision of 85.81%, a recall of 73.89%, and an F1-score of 79.40%. These results indicate that the Naïve Bayes algorithm is effective in identifying potential delays in deliveries. This research demonstrates the potential of machine learning methods to support operational decisions in logistics service providers.

Keyword: Naïve Bayes, Delivery Prediction, Classification, JNE, Logistics, PHP, MySQL