



**ANALISIS KEPUASAN PENUMPANG LAYANAN KRL COMMUTER
LINE RUTE YOGYAKARTA – SOLO BALAPAN MENGGUNAKAN
ALGORITMA NAÏVE BAYES DENGAN OPTIMALISASI
INFORMATION GAIN**

TUGAS AKHIR

Muhammad Iqbal
41518010066

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



**ANALISIS KEPUASAN PENUMPANG LAYANAN KRL COMMUTER
LINE RUTE YOGYAKARTA – SOLO BALAPAN MENGGUNAKAN
ALGORITMA NAÏVE BAYES DENGAN OPTIMALISASI
INFORMATION GAIN**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Muhammad Iqbal
41518010066

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

LEMBAR PERNYATAAN ORISINALITAS

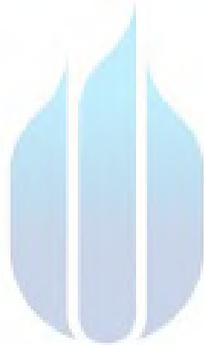
Yang bertanda tangan dibawah ini:

NIM : 41518010066

Nama : Muhammad Iqbal

Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL Commuter
Line Rute Yogyakarta – Solo Balapan Menggunakan
Algoritma Naïve Bayes dengan Optimalisasi Information
Gain

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



UNIVERSITAS
MERCU BUANA

Jakarta, 29 Juni 2022



Muhammad Iqbal

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

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

Nama Mahasiswa : Muhammad Iqbal
NIM : 41518010066
Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL
Commuter Line Rute Yogyakarta – Solo Balapan
Menggunakan Algoritma Naïve Bayes Dengan
Optimalisasi Information Gain

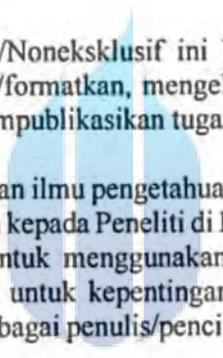
Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

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, 29 Juni 2022



METERAI TEMPEL
04B6AJX910684943
Muhammad Iqbal

SURAT PERNYATAAN LUARAN TUGAS AKHIR

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

Nama Mahasiswa : Muhammad Iqbal
 NIM : 41518010066
 Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL
 Commuter Line Rute Yogyakarta – Solo Balapan
 Menggunakan Algoritma Naïve Bayes Dengan
 Optimalisasi Information Gain

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 RESTI (Rekayasa Sistem dan Teknologi Informasi)			
	ISSN	: 2580-0760 (Online - Elektronik)			
	Link Jurnal	: https://jurnal.iaii.or.id/index.php/RESTI			
	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.

Jakarta, 22 Agustus 2022

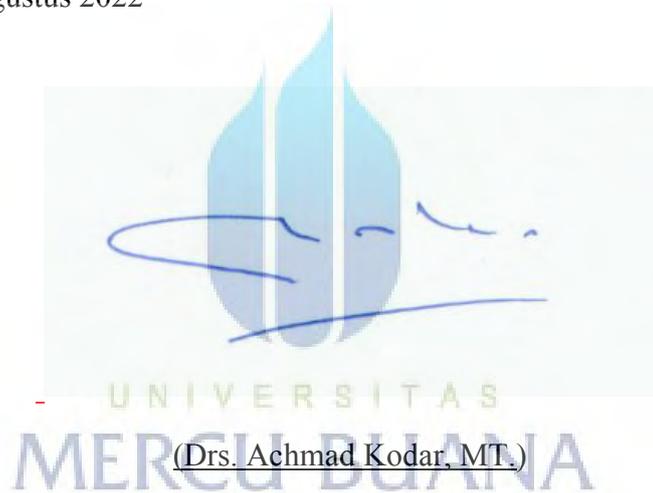

 Muhammad Iqbal

LEMBAR PERSETUJUAN PENGUJI

Nama Mahasiswa : Muhammad Iqbal
NIM : 41518010066
Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL
Commuter Line Rute Yogyakarta – Solo Balapan
Menggunakan Algoritma Naïve Bayes Dengan
Optimalisasi Information Gain

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, 21 Agustus 2022



LEMBAR PERSETUJUAN PENGUJI

Nama Mahasiswa : Muhammad Iqbal
NIM : 41518010066
Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL
Commuter Line Rute Yogyakarta – Solo Balapan
Menggunakan Algoritma Naïve Bayes Dengan
Optimalisasi Information Gain

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, 19 Agustus 2022



(Sabar Rudiarto, M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

Nama Mahasiswa : Muhammad Iqbal
NIM : 41518010066
Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL
Commuter Line Rute Yogyakarta – Solo Balapan
Menggunakan Algoritma Naïve Bayes Dengan
Optimalisasi Information Gain

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, 22 Agustus 2022



(Dr. Puji Rahayu, MT.)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

Nama Mahasiswa : Muhammad Iqbal
NIM : 41518010066
Judul Tugas Akhir : Analisis Kepuasan Penumpang Layanan KRL Commuter
Line Rute Yogyakarta – Solo Balapan Menggunakan
Algoritma Naïve Bayes Dengan Optimalisasi Information
Gain

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, 22 Agustus 2022

Menyetujui,



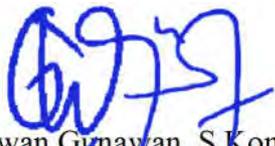
(Harni Kusniyati, ST., M. Kom.)

Dosen Pembimbing

UNIVERSITAS

MERCU BUANA

Mengetahui,



(Wawan Gumayan, S.Kom, MT)

Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)

Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan limpahan keberkahan, karunia, dan rahmatnya sehingga penulis dapat menyelesaikan sebuah kegiatan yang wajib ditempuh oleh mahasiswa semester akhir di Universitas Mercu Buana yaitu Tugas Akhir. Dalam penyelesaian kegiatan laporan Tugas Akhir tersebut, pastinya Penulis menemukan sebuah kendala – kendala dan permasalahan yang dihadapi oleh karena itu, Penulis menyadari bahwa tanpa bantuan dan bimbingan dari beberapa pihak.

Dalam penyusunan laporan Tugas Akhir ini penulis mendapat banyak bantuan dari berbagai pihak. Untuk itu disampaikan rasa terima kasih kepada:

1. PT. Kereta Commuter Indonesia sebagai pihak yang bersedia memberikan perizinan sebagai bahan studi kasus dalam penelitian Tugas Akhir.
2. Bapak Ir. Emil R. Kaburuan, Ph.D., IPM. Selaku Kepala Bidang Program Studi Teknik Informatika di Universitas Mercu Buana.
3. Bapak Wawan Gunawan, S. Kom., MT selaku Koordinator Tugas Akhir di Universitas Mercu Buana.
4. Ibu Harni Kusniyati, ST., M. Kom. selaku Dosen Pembimbing di Mata Kuliah MPTI dan Tugas Akhir.
5. Kedua Orang Tua yang selalu memberikan dukungan secara materil/immaterial.
6. Teman – teman, rekan kerja, dan saudara – saudara yang saya kenal yang telah memberikan doa, dan dukungan.

Akhir kata, penulis berharap laporan Tugas Akhir ini dapat menjadi salah satu jejak sejarah dalam kehidupan penulis serta kedepannya dalam penelitian tugas akhir ini dapat bermanfaat untuk referensi penelitian selanjutnya, mohon maaf jika ada kesalahan dan kekurangan dalam penelitian/penulisan laporan Tugas Akhir ini.

Jakarta, 29 Juni 2022
Penulis

Muhammad Iqbal
xi

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.....	xii
NASKAH JURNAL	1
KERTAS KERJA.....	15
BAB 1. LITERATUR REVIEW	16
BAB 2. ANALISIS DAN PERANCANGAN.....	24
BAB 3. SOURCE CODE	37
sBAB 4. DATASET	44
BAB 5. TAHAPAN EKSPERIMEN	67
BAB 6. HASIL SEMUA EKSPERIMEN.....	72
DAFTAR PUSTAKA	95
LAMPIRAN DOKUMEN HAKI.....	98
LAMPIRAN KORESPONDENSI	101
LAMPIRAN TAMBAHAN	102



Analisis Kepuasan Penumpang Layanan KRL Commuter Line
Rute Yogyakarta – Solo Balapan Menggunakan Algoritma Naïve Bayes
Dengan Optimalisasi Information Gain

Muhammad Iqbal¹, Harni Kusniyati²

^{1,2}Teknik Informatika, Ilmu Komputer, Universitas Mercu Buana

¹41518010066@student.mercubuana.ac.id, ²harni.kusniyati@mercubuana.ac.id

Abstract

Yogyakarta City and Surakarta City are cities that have a very high level of mobility, KRL Commuter Line is an electric rail-based rail transportation which was just launched in 2021 ago. Services for transportation, especially on the KRL Commuter Line, need to be considered so that prospective passengers can experience transportation comfortably and safely. For this reason, customer satisfaction is needed to be a benchmark for a service, the elements of a service aspect, namely (Tangibels, Reliability, Responsiveness, Assurance, Empathy). For this reason, researchers conducted research on passenger satisfaction of Commuter Line KRL services on the Yogyakarta – Solo Balapan route, research using the Naïve Bayes Algorithm which is an algorithm for classification coupled with Information Gain Optimization so that the results of the accuracy of the information are more optimal. After taking the questionnaire survey, there were 132 respondents and the results in this study are the final result of using the Naïve Bayes Algorithm using Information Gain resulting in a higher accuracy of 94.7% compared to the test results with the Naïve Bayes Algorithm without optimizing the Information Gain of 92.42%.

Keywords: Customer Satisfaction, Commuter Line, Naïve Bayes, Information Gain

Abstrak

Kota Yogyakarta dan Kota Surakarta merupakan kota yang memiliki tingkat mobilitas yang sangat tinggi, KRL Commuter Line merupakan transportasi kereta berbasis rel listrik yang baru diluncurkan pada tahun 2021 yang lalu. Pelayanan terhadap transportasi khususnya di KRL Commuter Line perlu diperhatikan agar para calon penumpang dapat merasakan transportasi secara nyaman dan aman. Untuk itu kepuasan pelanggan sangat dibutuhkan untuk menjadi tolak ukur terhadap suatu pelayanan, unsur aspek sebuah layanan yakni (Tangibels, Reliability, Responsiveness, Assurance, Emphaty). Untuk itu peneliti melakukan penelitian terhadap kepuasan penumpang layanan KRL Commuter Line di rute Yogyakarta – Solo Balapan, penelitian menggunakan Algoritma Naïve Bayes yang merupakan algoritma untuk pengklasifikasian ditambah dengan Optimalisasi Information Gain agar hasil akurasi informasi tersebut lebih optimal. Setelah dilakukan pengambilan survei kuesioner, terdapat 132 responden dan hasil dalam penelitian ini adalah hasil akhir dari penggunaan Algoritma Naïve Bayes dengan menggunakan Information Gain menghasilkan akurasi lebih tinggi yaitu sebesar 94,7% dibandingkan dengan hasil uji dengan Algoritma Naïve Bayes tanpa melakukan optimalisasi Information Gain sebesar 92,42%.

Kata kunci: Kepuasan Penumpang, KRL Commuter Line, Naïve Bayes, Information Gain

Diterima Redaksi: xx-xx-2022 | Selesai Revisi: xx-04-2022 | Diterbitkan Online: xx-04-2022

1. Pendahuluan

Kota Yogyakarta dan Kota Surakarta berada di dua provinsi yang berbeda yaitu Provinsi Daerah Istimewa Yogyakarta dan Provinsi Jawa Tengah, kedua kota tersebut sama halnya dengan wilayah Jabodetabek memiliki mobilitas yang tinggi dari suatu tempat ke tempat lain, untuk itu transportasi publik sangat dibutuhkan. Kebutuhan terhadap transportasi publik kian meningkat seiringnya pertumbuhan penduduk, untuk itu pemerintah selaku regulator penyedia jasa transportasi publik terus menggenjot untuk menghadirkan sarana dan prasarana transportasi publik yang lebih baik. Salah satu rencana pemerintah dilakukan sampai saat ini adalah menghadirkan layanan Kereta Rel Listrik (KRL) di jalur Yogyakarta – Solo Balapan untuk tahap awal, program ini sudah disusun dalam Rencana Induk Perkeretapian Nasional 2030 atau bisa disebut RIPNAS 2030[1].

Pada program ini, maka dapat terealisasi dengan dimulainya elektrifikasi tahap pertama yaitu jalur kereta api Yogyakarta sampai Solo Balapan sehingga pada Tanggal 10 Februari 2021, Pemerintah dan PT. Kereta Commuter Indonesia yang merupakan anak perusahaan dari PT. Kereta Api Indonesia (Persero) meresmikan layanan Kereta Komuter berbasis KRL untuk menggantikan kereta api Prambanan Ekspres (Prameks).

Layanan KRL Commuter Line pada rute Yogyakarta – Solo Balapan ini memiliki sepuluh stasiun pemberhentian yaitu Lempuyangan, Maguwo, Brambanan, Srowot, Klaten, Ceper, Delanggu, Gawok, Purwosari, dan tujuan akhir yaitu Solo Balapan. Tarif yang dipatok untuk 1 kali perjalanan yaitu Rp. 8000 secara *flat* dengan menggunakan Kartu Multi Trip (KMT) yang dikeluarkan oleh KAI Commuter, Kartu Uang Elektronik (KUE) yang diterbitkan oleh bank, serta pembayaran melalui dompet digital yaitu Link Aja.

Kepuasan pelanggan dapat dinilai dari baik atau buruknya kualitas suatu pelayanan. Pelanggan akan menilai dari suatu pelayanan yang mereka dapatkan, untuk itu peningkatan kualitas pelayanan sering dilakukan oleh setiap pelaku usaha agar mempertahankan keberlangsungan sebuah bisnis. Peningkatan kualitas pelayanan sebagai salah satu upaya pencapaian kepuasan pelanggan bukanlah tugas yang ringan dapat diselesaikan dengan secepatnya, oleh karena itu perlu diupayakan secara berkesinambungan

serta berorientasi kepada perbaikan untuk mencapai kepuasan pelanggan [2].

Untuk mengetahui tingkat kepuasan pelanggan, biasanya menggunakan perhitungan secara statistik dari data – data yang dikumpulkan, namun pada era teknologi ini perhitungan tingkat kepuasan pelanggan semakin mendekati akurat dengan adanya dukungan dari proses pembelajaran yang dilakukan oleh komputer maka akan semakin cepat untuk mengambil sebuah keputusan dalam membenahi sebuah kualitas pelayanan.

Sebagai beberapa studi kasus penulis temukan, yang pertama dari Efendy Tambunan melakukan analisis kualitas pelayanan Commuter Line rute Parung Panjang – Tanah Abang menggunakan metode deskriptif dan kuantitatif menunjukkan tingkat kehandalan (*reliability*) ketepatan waktu operasional relatif baik yaitu sebesar 4,21% dari batas toleransi PM Perhubungan RI No. 48 Tahun 2015 sebesar 20% [3]. Anggita Safitri Febrianti dan Erna Zuni Astuti melakukan penelitian prediksi kepuasan penumpang bus Trans Semarang dengan mengimplementasikan algoritma C4.5 hasil dari 3 kali percobaan bahwa tiap nilai tertinggi mulai dari akurasi 95%, *precision* 96,67% dan *recall* 97,87% [4]. Fauziah, DKK melakukan penelitian kepuasan pelanggan menggunakan klasifikasi data mining dengan studi kasus Indiiis *Café* Pematangsiantar dalam penelitiannya metode digunakan yaitu melakukan pembagian kuesioner terhadap pengunjung kafe lalu diolah dengan teknik Data Mining dan permodelan Algoritma C4.5 hasilnya mendapatkan akurasi sebesar 98% [5].

Biasanya, untuk melakukan sebuah analisa kepuasan pelanggan, Algoritma C4.5 banyak digunakan karena dapat mampu mengolah dataset secara cepat dan handal, mudah dimengerti, dan mudah dibangun [6]. Dari pemaparan dari trend dan masalah yang ada, penulis ingin melakukan sebuah penelitian yaitu kepuasan penumpang KRL Commuter Line Yogyakarta – Solo Balapan menggunakan pendekatan data mining dan permodelan algoritma Naïve Bayes. Secara singkat algoritma Naïve Bayes merupakan sebuah algoritma pengklasifikasian probalistik secara sederhana dengan menggunakan penjumlahan sebuah frekuensi dan kombinasi nilai dataset yang diberikan [7].

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Agar penelitian mendapatkan hasil yang maksimal, penulis menambahkan optimalisasi pada Algoritma *Naive Bayes* menggunakan *Information Gain*. Secara singkat, *Information Gain* merupakan metode seleksi fitur paling sederhana dengan dengan melakukan pengurutan atribut, biasanya optimaliasi ini digunakan pada katagorisasi teks, analisis data citra, dan analisis data *array* [8]. Dengan ini diharapkan penerapan *Information Gain* dapat meningkatkan hasil dalam analisa kepuasan penumpang KRL Commuter Line rute Yogyakarta – Solo Balapan.

2. Metode Penelitian

Dalam metode penelitian kali ini, peneliti menggunakan metode pendekatan secara kuantitatif untuk pengambilan data, setelah pengambilan data ada beberapa tahap yang harus dipersiapkan sebelum melakukan pembahasan hasil penelitian yang dilakukan yang bisa dilihat pada Gambar 1.



Gambar 0.1. Tahapan Penelitian

2.1. Pengumpulan Data

Pengumpulan data merupakan tahap awal Peneliti dalam penelitian ini. Untuk itu, agar mendapatkan sebuah *dataset* yang diinginkan, maka dilakukan pengambilan data dengan cara melakukan survei kuesioner yang dibagikan secara *online* maupun melakukan terjun langsung ke penumpang yang berada disekitar Stasiun Yogyakarta dan Stasiun Solo Balapan.

Adapun kriteria dalam mengisi kuesioner ini:

1. Minimal usia 15 tahun

2. Setidaknya pernah sekali/sering menggunakan layanan KRL Commuter Line di antara Stasiun Yogyakarta – Stasiun Solo Balapan.

Untuk pertanyaan pun sendiri terdiri dari beberapa pertanyaan seputar identitas responden dan pertanyaan penelitian yang berjumlah 50 bulir pertanyaan yang dibagi 5 bagian.

Dalam membuat pertanyaan untuk mengukur sebuah kepuasan pelanggan tidak bisa dilakukan secara sembarangan, menurut teori Lovelock dan Wright pada tahun 2002, ada 5 aspek dimensi kualitas layanan yaitu [9]:

1. *Tangibels* (Bukti Fisik)

Memberikan bukti secara fisik atau langsung seperti contoh fasilitas dilapangan, peralatan, atau material – material yang dimiliki perusahaan, fungsi dari tersebut untuk menarik minat penumpang.

2. *Reliability* (Kehandalan)

Kehandalan sebuah penyedia transportasi memberikan layanan sesuai dengan dijanjikan, akurat, dapat dipercaya, dan konsisten. Hal ini bertujuan untuk memberikan rasa kepercayaan kepada penumpang terhadap apa yang penyedia transportasi dijanjikan.

3. *Responsiveness* (Tanggap)

Penyedia transportasi harus tanggap dengan kebutuhan penumpang dan dilakukan secara secepat mungkin untuk membantu penumpang dalam memberikan sebuah layanan.

4. *Assurance* (Jaminan)

Penyedia transportasi harus memberikan sebuah jaminan dan memberi rasa aman kepada penumpang dengan pengetahuan yang memadai, kesopanan, sikap professional, dan percaya diri.

5. *Empathy* (Kepedulian)

Penyedia transportasi harus ada rasa kepedulian dan perhatian untuk semua penumpang dari hubungan komunikasi, perhatian pribadi, dan memahami kebutuhan pelanggan.

Berikut merupakan daftar pertanyaan kuesioner yang diajukan:

A. Identitas Responden

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

No.	Pertanyaan	Indikator Jawaban
1.	Nama responden?	Nama Responden
2.	Nomor telepon/Whatsapp?	Nomor telepon responden
3.	Jenis kelamin?	Pria / Wanita
4.	Usia	Rentan usia: 15-20/21-35/35-50/50 tahun keatas
5.	Domisil tempat tinggal?	DIY Yogyakarta / Solo Raya / Lainnya (sebutkan)
6.	Tujuan menggunakan layanan commuter line	Belajar / Bekerja /Berwisata/ Berkunjung
7.	Stasiun asal keberangkatan anda?	Yogyakarta/Lempuyangan/Magwo/Brambanan/Srowot/Klaten/Ceper/Delanggu/Gawok/Purwosari/Solo Balapan
8.	Stasiun akhir tujuan anda?	Yogyakarta/Lempuyangan/Magwo/Brambanan/Srowot/Klaten/Ceper/Delanggu/Gawok/Purwosari/Solo Balapan
9.	Pembayaran layanan KAI Commuter yang anda sering digunakan?	Kartu Multi Trip/Kartu Bank/Link Aja

B. Pertanyaan Penelitian

I. Tangibles		
No.	Pertanyaan	Indikator Jawaban
1.	Tersedianya papan informasi pada stasiun meliputi nama stasiun, nomor jalur, arah jalur, dan tujuan kereta.	1-5
2.	Tersedia fasilitas pembelian tiket lewat loket atau mesin.	1-5
3.	Tersedia fasilitas ruang tunggu untuk KRL	1-5
4.	Tersedia fasilitas toilet pria dan wanita	1-5
5.	Tersedia fasilitas ruang kesehatan dan ruang untuk ibu menyusui.	1-5
6.	Tersedia area komersil (minimarket/toko/ATM)	1-5
7.	Tersedia fasilitas penunjang untuk kaum disabilitas (Lift, akses kursi roda, toilet difabel).	1-5

II. Reliability		
No.	Pertanyaan	Indikator Jawaban
8.	Tersedia fasilitas mushola yang memadai	1-5
9.	Tersedia petugas keamanan di dalam stasiun maupun di dalam kereta	1-5
10.	Tersedia fasilitas di dalam rangkaian kereta (tempat duduk, pendingin ruangan, dan hanger untuk berdiri).	1-5

III. Responsiveness		
No.	Pertanyaan	Indikator Jawaban
1.	Lokasi stasiun yang terjangkau (dekat dengan pusat keramaian atau sambungan antar moda transportasi lain).	1-5
2.	Kesigapan petugas pelayanan dalam stasiun dalam membantu penumpang.	1-5
3.	Kesigapan petugas keamanan dalam mengamankan stasiun.	1-5
4.	Kecepatan petugas loket dalam melayani penumpang.	1-5
5.	Kehandalan petugas kebersihan dalam membersihkan area stasiun	1-5
6.	Ketepatan jadwal keberangkatan dan kedatangan kereta,	1-5
7.	Kehandalan masinis dalam mengendarai kereta	1-5
8.	Kehandalan petugas pelayanan kereta dalam memberikan informasi didalam kereta (rute tujuan, stasiun pemberhentian, peraturan di dalam kereta, keadaan jika ada hal darurat, memberi himbauan dalam memberi hak tempat duduk penumpang prioritas kepada penumpang non prioritas).	1-5
9.	Kehandalan petugas keamanan di dalam kereta untuk mengamankan penumpang kereta.	1-5
10.	Kehandalan petugas kebersihan dalam membersihkan di dalam kereta,	1-5

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

3.	Ketanggapan dalam memberikan informasi kepada penumpang yang baru pertama kali menaiki kereta.	1-5
4.	Ketanggapan petugas dalam memberikan informasi seputar jadwal keberangkatan dan kedatangan kereta api	1-5
5.	Ketanggapan petugas dalam keadaan darurat.	1-5
6.	Ketanggapan petugas keamanan dalam menertibkan penumpang di stasiun atau dalam kereta.	1-5
7.	Ketanggapan petugas dalam mencari kursi untuk penumpang prioritas (ibu hamil, lansia, ibu membawa balita, penyandang disabilitas, memiliki penyakit yang tidak memungkinkan untuk berdiri).	1-5
8.	Ketanggapan petugas keamanan dalam mencari barang penumpang yang tertinggal atau kehilangan.	1-5
9.	Ketanggapan petugas pelayanan kereta dalam membuka atau menutup pintu kereta	1-5
10.	Ketanggapan petugas kebersihan dalam membersihkan stasiun atau didalam kereta jika ada bagian yang kotor.	1-5

IV. Assurance

No.	Pertanyaan	Indikator Jawaban
1.	Tersedianya papan petunjuk informasi yang jelas (tidak membingungkan).	1-5
2.	Tersedianya pengumuman informasi melalui pengeras suara dengan jelas.	1-5
3.	Tersedianya pengumuman informasi melalui sosial media.	1-5
4.	Tersedianya informasi tujuan kereta dan pemberhentian stasiun selanjutnya.	1-5
5.	Tersedianya layanan via telepon untuk keluhan penumpang yang aktif selama 24 jam.	1-5
6.	Tersedianya petunjuk keselamatan dalam keadaan kondisi darurat di dalam kereta	1-5
7.	Pengetahuan petugas tentang informasi seputar kereta	1-5
8.	Jaminan keamanan dalam stasiun dan dalam kereta	1-5
9.	Jaminan ketepatan waktu perjalanan kereta	1-5
10.	Jaminan dalam layanan khusus penumpang prioritas/disabilitas.	1-5

V. Emphaty

No.	Pertanyaan	Indikator Jawaban
1.	Kebersihan dalam stasiun	1-5
2.	Jaminan keselamatan penumpang (rambu peringatan dalam stasiun).	1-5
3.	Keramahan petugas dalam stasiun	1-5
4.	Keamanan dalam stasiun	1-5
5.	Kenyamanan dalam stasiun	1-5
6.	Kebersihan dalam kereta	1-5
7.	Kenyamanan dalam kereta	1-5
8.	Keramahan petugas dalam kereta	1-5
9.	Keamanan di dalam kereta	1-5
10.	Jaminan keselamatan penumpang (rambu peringatan dalam kereta).	1-5

2.2. Pengolahan Data

Pengolahan data atau persiapan data merupakan tahapan persiapan sebuah dataset sebelum dilakukannya ke tahap pengujian dan penggunaan model algoritma. Fungsi persiapan / pengolahan data ini dilakukan untuk mendapatkan gambaran umum dari data proses.

Dalam proses pengolahan data dalam penelitian ini, peneliti melakukan menggunakan prosedur *data mining*.

Data Mining merupakan sebuah penerapan teknologi untuk memproses sejumlah data – data informasi secara kompleks, multi-level dan database yang *terdiversifikasi* dalam proses *data mining* sendiri, data yang berada didalam akan dilakukan penambangan secara teknologi untuk menghasilkan sebuah ekstrak informasi yang berguna dan berharga [10].

Proses data mining dalam penelitian ini, peneliti melakukan beberapa langkah – langkah dilakukan berikut pada di bawah ini:



Gambar 0.2. Tahapan Proses Data Mining

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

1. Selection Data

Selection data dilakukan untuk memilih manakah atribut yang akan digunakan dalam eksperimen Analisis Kepuasan Penumpang Layanan KRL Commuter Line Rute Yogyakarta – Solo Balapan menggunakan Algoritma Naïve Bayes dengan Optimalisasi Information Gain.

2. Preprocessing Data

Preprocessing Data pada proses penelitian ini yaitu proses menyederhanakan data – data agar mudah dilakukan proses transformasi data.

3. Transformation Data

Transformasi Data yaitu proses konversi data atribut menjadi bentuk yang mudah dilakukan perhitungan atau visualisasi data seperti menentukan class kategori label atau mengubah isi atribut menjadi sebuah kode/id agar dapat diseragamkan.

4. Output Dataset

Output dataset yang diharapkan adalah dataset yang sudah dilakukan proses Selection, Proprocessing, dan Transformation. Beberapa atribut dalam dataset sudah dilakukan proses *labeling* dan file format dataset yang ditentukan adalah file berformat .csv (*Comma Separated Value*).

2.3. Pengujian Model Algoritma

Setelah mendapatkan dataset pengujian kepuasan penumpang layanan KRL Commuter Line rute Stasiun Yogyakarta – Solo Balapan dalam berbentuk format .csv . Selanjutnya dilakukan pengujian menggunakan sebuah algoritma Naïve Bayes dengan Optimalisasi Information Gain.

A. Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan sebuah algoritma dalam kategori Machine Learning yang ditemukan oleh seorang pendeta yang juga sekaligus sebagai ilmuwan di Inggris yang bernama Thomas Bayes pada abad Ke – 18 yang dimana variabel X sebagai data tupel, dalam istilah Bayesian, X dianggap sebagai bukti [11].

Metode Naïve Bayes sendiri merupakan sebuah metode pengklasifikasian secara probalistik sederhana yang fungsinya sebagai menghitung sekumpulan probabilitas dengan menunjukkan kombinasi nilai dari dataset yang diberikan. Metode ini ada juga yang menyebut sebagai

metode analisis statistik dimana probabilitas awal diestimasi dari data training [12].

Algoritma Naïve Bayes ini ditentukan probabilitas awal dari setiap parameter dan probabilitas yang memiliki rumus persamaan dibawah secara matematis yaitu:

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

Keterangan:

P (A | B) = Nilai probabilitas A yang terjadi jika B.

P (B | A) = Nilai probabilitas B yang terjadi jika A.

P (A) = Nilai probabilitas A.

P (B) = Nilai probabilitas B.

Klasifikasi Naïve Bayes dapat diasumsikan ada atau tidak adanya sebuah ciri tertentu dari sebuah kelas yang dengan berhubungan dari sebuah ciri dari kelas lainnya [13]. *Output* yang dihasilkan dalam proses algoritma Naïve Bayes ini yaitu kita dapat memprediksikan sesuatu seperti contoh apakah seseorang berasal dari sebuah kelompok tertentu berdasarkan variabel yang melekat padanya.

Untuk pemrosesan algoritma ini ada beberapa tahap – tahap dilakukan agar algoritma ini berjalan yaitu menghitung jumlah kelas/label, menghitung jumlah kasus per kelas, mengkalikan semua variabel kelas dan, membandingkan hasil perkelas [14].

B. Information Gain

Information Gain adalah salah satu metode seleksi fitur yang paling sederhana dengan cara perangkingan atribut dan banyak digunakan seperti aplikasi kategorisasi teks, analisis data secara microarray, dan analisis data citra. Fungsi dari Information Gain pun sendiri yaitu dapat mengurangi sebuah noise yang disebabkan dari fitur – fitur yang tidak relevan dan dapat mendeteksi fitur – fitur paling banyak yang memiliki informasi berdasarkan kelas tertentu [15].

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Untuk mendapatkan hasil dengan maksimal untuk perhitungan permodelan Naïve Bayes melakukan perhitungan entropy dengan persamaan berikut:

$$Entropy(s) = \sum_i^f - P_i \log_2 P_i \dots (1)$$

Setelah mendapatkan nilai entropy, lalu dapat di cari nilai perhitungan *Information Gain* yang dapat dilakukan dengan persamaan berikut:

$$Gain(S, A) = Entropy(s) - \sum_{Values(A)} \frac{|Sv|}{|S|} Entropy(Sv) \dots (2)$$

3. Hasil dan Pembahasan

Pada BAB ini, Peneliti akan menjelaskan hasil eksperimen yang dilakukan selama penelitian mulai dari penggunaan alat perangkat lunak dan perangkat keras, pengolahan data, hingga hasil uji model Algoritma Naïve Bayes dengan Optimalisasi Information Gain.

3.1. Peralatan Pengujian

Pada penelitian kali ini, peneliti menggunakan beberapa peralatan perangkat lunak dan perangkat keras dalam penelitian Analisis Kepuasan Penumpang Layanan KRL Commuter Line Rute Stasiun Yogyakarta – Stasiun Solo Balapan yang bisa dilihat pada tabel dibawah ini.

A. Spesifikasi Perangkat Keras

Tabel 0.3. Spesifikasi Perangkat Keras

No	Spesifikasi	Keterangan
1.	CPU	Intel Core i5-9300H (4 Core, 8 Thread) Frequency 2.40GHz – 4.10GHz
2.	RAM	8GB DDR4 2666MHz SDRAM
3.	STORAGE	1TB SSHD
4.	GPU	NVIDIA GeForce GTX 1050 4GB GDDR5 VRAM
5.	DISPLAY	15,6” (16:9) IPS-level FHD (1920x1080), 60Hz
6.	BATTERY	48WHrs, 3S1P, 3-cell Li-ion
7.	OPERASI SISTEM	Windows 11 Pro Versi. 21H2

B. Spesifikasi Perangkat Lunak

Tabel 0.4. Spesifikasi Perangkat Lunak

No	Nama Software	Versi
1.	Microsoft Excel	Microsoft Office Professional Plus 2019 Versi 1912
2.	Weka	Versi 3.9.6

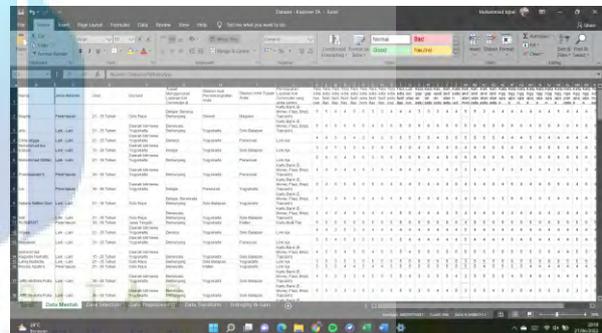
1.	Microsoft Excel	Microsoft Office Professional Plus 2019 Versi 1912
2.	Weka	Versi 3.9.6

3.2. Proses Pengolahan Data

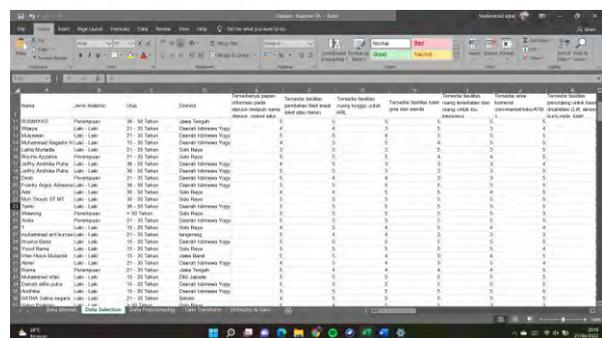
Proses pengolahan data ini menggunakan proses Data Mining, data tersebut bersumber dari data hasil melakukan survei kuesioner kepada penumpang KRL Commuter Line pada rute Yogyakarta – Solo Balapan, dan didapatkan sebesar 132 record data, data sebut akan diolah sebelum dilakukan tahapan permodelan Algoritma Naïve Bayes Dengan Information Gain berikut ini hasil pengolahan data yang dilakukan.

A. Selection Data

Selection Data pada tahapan ini peneliti menyeleksi atribut yang hanya dibutuhkan dalam melakukan eksperimen data, dari semua pertanyaan, peneliti melakukan pemangkasan pada atribut pertanyaan yang terkait dengan identitas responden, hasilnya hanya atribut Nama, Jenis Kelamin, Usia, dan Domisil hanya dipakai untuk tahapan eksperimen, untuk pertanyaan utama seputar penilaian layanan tidak ada perubahan.



Gambar 0.3. Data sebelum dilakukan data selection



Gambar 0.4. Data sesudah dilakukan data selection

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

B. Preprocessing Data

Pada tahapan ini, peneliti melakukan penyederhanaan dataset dengan melakukan perhitungan terhadap tiap aspek pertanyaan seputar layanan KRL Commuter Line yang dilakukan dengan Microsoft Excel menggunakan Formula SUM table. Selain itu, peneliti juga menghitung keseluruhan skor terhadap 50 pertanyaan yang dilakukan penilaian pada responden.

Berikut dibawah ini merupakan beberapa sample dari hasil preprocessing data:

Nama	Jenis Kelamin	Usia	Domisil	HITUNG ASPEK A	HITUNG ASPEK B	HITUNG ASPEK C	HITUNG ASPEK D	HITUNG ASPEK E	Total Skor Keseluruhan Per-responden
Maqda	Perempuan	21 - 35 Tahun	Solo Raya	40	38	41	35	34	188
arta	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	49	50	50	50	50	249
Dito Andas	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	40	49	41	38	40	208
Muhammad Isa Edison	Laki - Laki	15 - 20 Tahun	Daerah Istimewa Yogyakarta	34	40	43	32	42	191
Muhammad Ghafari	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	45	48	50	46	50	239
Pranaboweni K.	Perempuan	36 - 50 Tahun	Daerah Istimewa Yogyakarta	39	39	37	37	40	192
Lia	Perempuan	36 - 50 Tahun	Daerah Istimewa Yogyakarta	47	41	44	41	42	215
Natano Hafino Dion	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	45	46	44	41	49	225
Adi	Laki - Laki	21 - 35 Tahun	Solo Raya	40	40	40	41	40	201
RUSMIYATI	Perempuan	36 - 50 Tahun	Jawa Tengah	50	49	50	50	50	249
Wijaya	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	39	41	40	40	40	200
Mujawan	Laki - Laki	21 - 35 Tahun	Daerah Istimewa Yogyakarta	49	41	43	50	42	225
Muhammad Ragadin	Laki - Laki	15 - 20 Tahun	Daerah Istimewa Yogyakarta	47	49	49	47	50	241
Murhafiz	Laki - Laki	21 - 35 Tahun	Solo Raya	44	48	50	46	46	234
Latha Mustafa	Laki - Laki	21 - 35 Tahun	Solo Raya	46	49	43	42	43	223
Rischa Azahra	Perempuan	21 - 35 Tahun	Daerah Istimewa Yogyakarta	42	42	42	38	43	207
Jeffry Andhika Putra	Laki - Laki	36 - 50 Tahun	Daerah Istimewa Yogyakarta	40	39	31	31	35	178
Jeffry Andhika Putra	Laki - Laki	36 - 50 Tahun	Daerah Istimewa Yogyakarta	40	39	31	31	35	178

Gambar 0.5. Data sesudah dilakukan proses preprocessing data

C. Transformation Data

Pada tahapan ini, peneliti melakukan transformasi data yaitu mengubah value atribut menjadi sebuah kode yang lebih sederhana dan mudah untuk dibaca, transformasi data dilakukan juga untuk kategori labeling terhadap penilaian layanan KRL Commuter Line terhadap para responden.

Pada tahapan pertama transformasi data, dilakukan pengubahan pada field Nama, Jenis Kelamin, dan Domisil yang bisa dilihat pada table dibawah ini:

Tabel 0.5. Tabel Transformasi Data Terhadap Identitas Responden

Field	Isi Record Sebelum	Isi Record Sesudah
Nama Responden	Nama	IDRES#(No. Urutan Responden)
Jenis Kelamin	Laki – Laki / Perempuan	L / P
Asal Wilayah	Wilayah Lain (Sebutkan nama wilayah)	Wilayah Lain

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tahapan selanjutnya, dilakukan kategorisasi label terhadap penjumlahan penilaian responden secara per aspek layanan dan keseluruhan dengan label kategori Sangat Puas, Puas, Cukup Puas, Kurang Puas, dan Tidak Puas. Untuk melakukan sebuah labeling, dilakukanlah pencarian interval kelas sebuah skor yang awalnya mencari nilai terendah terlebih dahulu, nilai terendah didapatkan jika responden menjawab pertanyaan dengan nilai 1 (satu), maka jika ada 10 pertanyaan per-aspek layanan, maka nilai maksimal yang di dapatkan adalah 10 poin dan total keseluruhan dari 50 pertanyaan seputar layanan adalah 50 poin.

Sebaliknya, untuk nilai tertingginya jika responden menjawab pertanyaan dengan nilai 5 (lima), maka nilai maksimal yang didapatkan adalah 50 poin dan total keseluruhan dari 50 pertanyaan seputar layanan adalah 250 poin.

Untuk keterangan jumlah nilai terendah dan tertinggi ada pada tabel dibawah ini:

Tabel 0.6. Tabel Nilai Terendah dan Tertinggi

Field	Min	Max
HITUNG ASPEK A	10	50
HITUNG ASPEK B	10	50
HITUNG ASPEK C	10	50
HITUNG ASPEK D	10	50
HITUNG ASPEK E	10	50
Total Skor Keseluruhan Per-responden	50	250

Jika telah menentukan jumlah nilai terendah dan tertinggi, maka tahapan selanjutnya adalah mencari interval penilaian dengan cara rumus dibawah ini:

$$\frac{\text{Nilai tertinggi} - \text{Nilai terendah}}{\text{jumlah skala kategori label}}$$

Maka perhitungan yang dilakukan terhadap seluruh atribut yang diatas, didapatkan sebuah interval kelasnya:

$$HITUNG ASPEK A = \frac{50 - 10}{5} = \frac{40}{5} = 8$$

$$HITUNG ASPEK B = \frac{50 - 10}{5} = \frac{40}{5} = 8$$

$$HITUNG ASPEK C = \frac{50 - 10}{5} = \frac{40}{5} = 8$$

$$HITUNG ASPEK D = \frac{50 - 10}{5} = \frac{40}{5} = 8$$

$$HITUNG ASPEK E = \frac{50 - 10}{5} = \frac{40}{5} = 8$$

$$Total Skor Keseluruhan Per - responden = \frac{250 - 50}{5} = \frac{200}{5} = 40$$

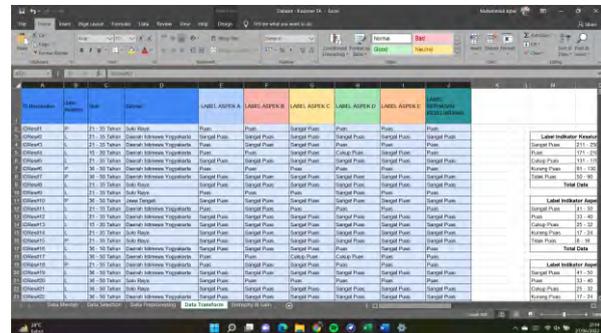
Maka dari hasil pencarian interval kelas tersebut, dapat diketahui kategorisasi label dari terhadap penilaian yang berikut penjabarannya dibawah ini:

Tabel 0.7. Skala Penilaian Label Kepuasan

Label Indikator Aspek A	
Sangat Puas	41 - 50
Puas	33 - 40
Cukup Puas	25 - 32
Kurang Puas	17 - 24
Tidak Puas	8 - 16
Label Indikator Aspek B	
Sangat Puas	41 - 50
Puas	33 - 40
Cukup Puas	25 - 32
Kurang Puas	17 - 24
Tidak Puas	8 - 16
Label Indikator Aspek C	
Sangat Puas	41 - 50
Puas	33 - 40
Cukup Puas	25 - 32
Kurang Puas	17 - 24
Tidak Puas	8 - 16
Label Indikator Aspek D	
Sangat Puas	41 - 50
Puas	33 - 40
Cukup Puas	25 - 32
Kurang Puas	17 - 24
Tidak Puas	8 - 16
Label Indikator Aspek E	
Sangat Puas	41 - 50
Puas	33 - 40
Cukup Puas	25 - 32
Kurang Puas	17 - 24
Tidak Puas	8 - 16
Label Indikator Keseluruhan	
Sangat Puas	211 - 250
Puas	171 - 210
Cukup Puas	131 - 170

Kurang Puas	91 - 130
Tidak Puas	50 - 90

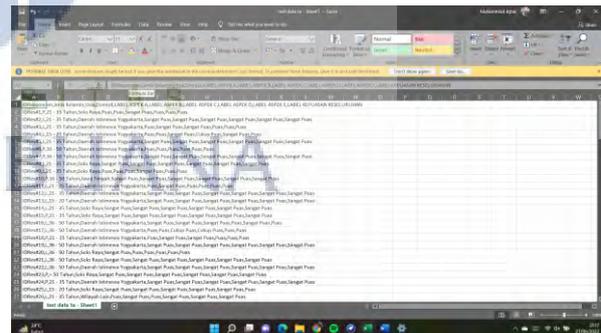
Setelah menentukan perubahan transformasi atribut yang dilakukan, maka berikut dibawah ini merupakan hasil dari transformasi data:



Gambar 0.6. Hasil transformasi data

D. Output Dataset

Setelah melakukan tahapan selection, preprocessing, dan transformasi data, maka peneliti tinggal mengcopy hasil data terakhir yang sudah diolah ke lembar kerja baru Microsoft Excel dan menyimpannya dalam bentuk .csv (Comma Separated Value).



Gambar 0.7. Output dataset berupa file berformat CSV

3.3. Proses Pengujian Algoritma

Pada proses pengujian eksperimen ini, peneliti menggunakan perangkat lunak open source yang

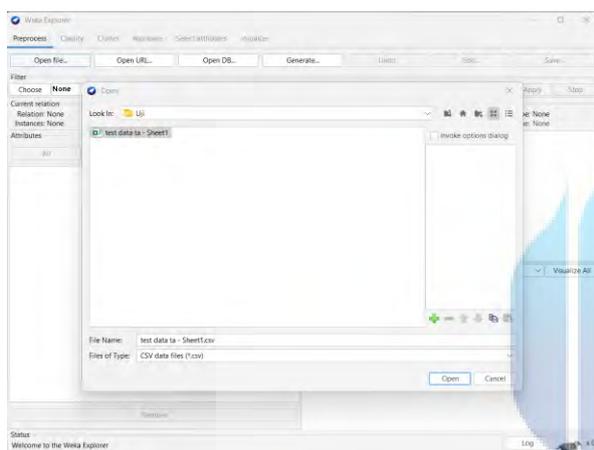
DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

bernama Weka versi 3.9.6 untuk tahapan dan hasil eksperimen dilakukan berikut penjelasannya.

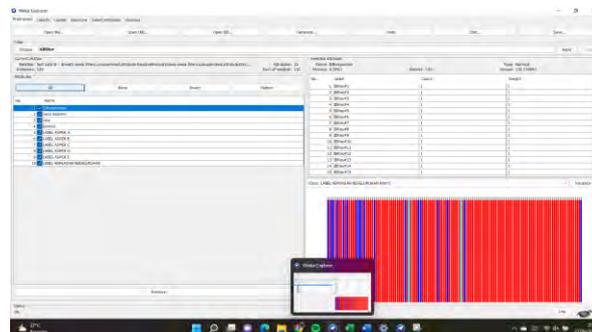
A. Import Dataset

Pada tahapan ini, file dataset csv yang sudah dibuat akan dibaca kedalam software Weka, caranya yaitu buka software Weka, lalu klik tombol Explorer pada bagian Applications, lalu maka akan terbuka tab aplikasi Weka dibagian Preprocess klik tombol “Open File” lalu sesuaikan format file yang ingin dibuka pada tombol Files of Types dan ganti ke format .csv habis itu cari file .csv yang ingin diuji maka file tersebut berhasil terbuka dan langsung simpan dalam bentuk format penyimpanan Software Weka yaitu .arrf (Arrf Data Files).



Gambar 0.8. Proses import file CSV ke Weka

menseleksi atribut yang akan diuji kedalam model Algoritma Naïve Bayes dengan Information Gain.

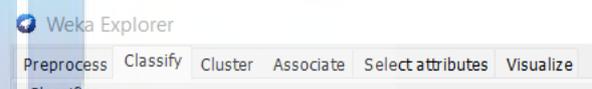


Gambar 0.9. Menu Tab Preprocessing dalam Weka

C. Pengurutan Menggunakan Information Gain

Setelah melakukan proses preprocessing dan seleksi atribut, maka dilakukan pengujian model Algoritma Naïve Bayes dengan Optimalisasi Information Gain. Langkah – langkah yang harus dilakukan yaitu:

1. Tekan tab “Classify”



2. Pada menu Classifier maka tekan tombol “Choose”



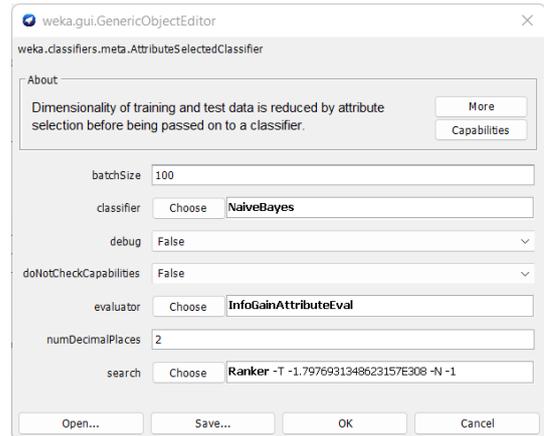
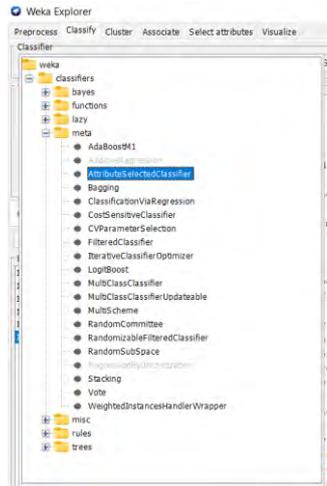
3. Pilih fitur yang ada di Weka > Meta > “AttributeSelectedClassifier” untuk melakukan klasifikasi atribut.

B. Persiapan Pengujian

Setelah melalui tahapan pertama, maka dilakukan persiapan kembali atribut yang ingin dilakukan pengujian seperti melakukan filtering secara supervised atau unsupervised terhadap attribute atau instance seperti, menambahkan, menghilangkan, mengkonversikan, penanganan missing value atau lain – lain. Karena dataset tersebut sudah dilakukan proses preprocessing di excel, maka pada tahapan ini cukup

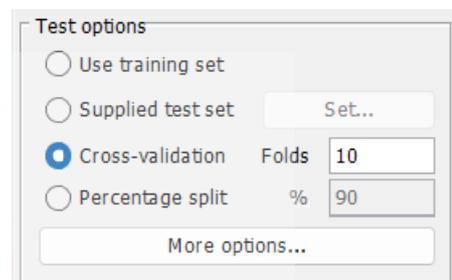
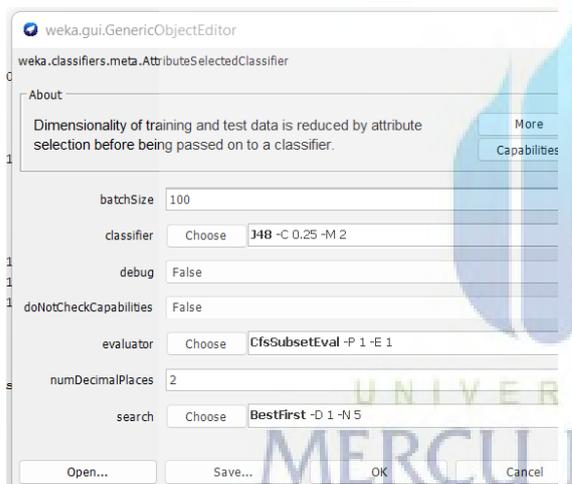
DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



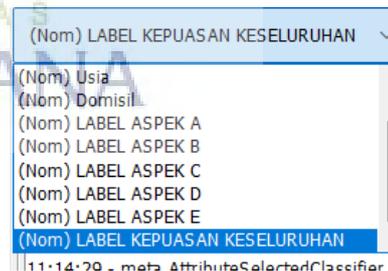
- Pada AttributeSelectedClassifier konfigurasi cukup dengan klik kanan > Show Properties maka terbuka menu weka.gui.GenericObjectEditor.

- Setelah itu pilih metode test options, terdapat Use Training Set, Supplied Test Set, Cross Validation, dan Percentage Split.



- Konfigurasi Classifier dengan “Naïve Bayes”, Evaluator menggunakan “InfoGainAttributEval”, dan jangan lupa untuk mengurutkan manakah atribut terbaik menggunakan search > Ranker.

- Terakhir pilih nominal atribut yang akan diuji dan hasil dari pengujian tersebut akan keluar seperti informasi hasil pengujian.



B. Hasil Eksperimen

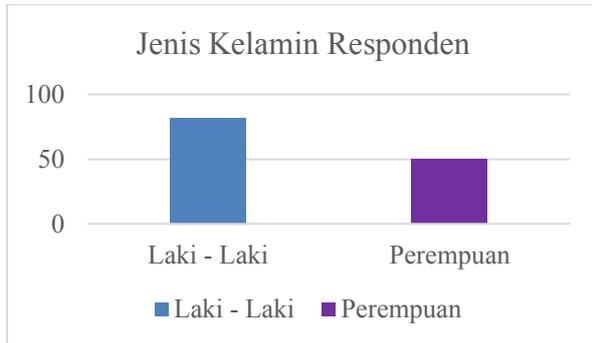
Jika sudah mengikuti langkah diatas, pada selanjutnya peneliti akan menjelaskan hasil eksperimen yang dilakukan.

1. Infografis Responden

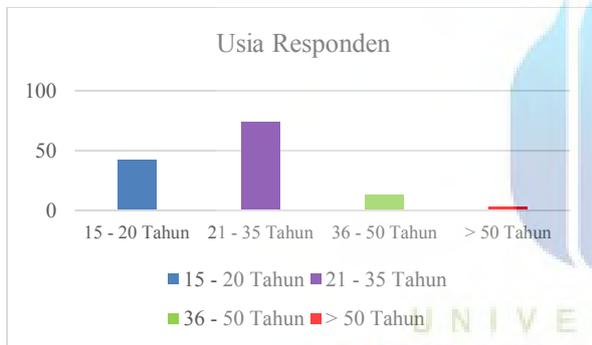
DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

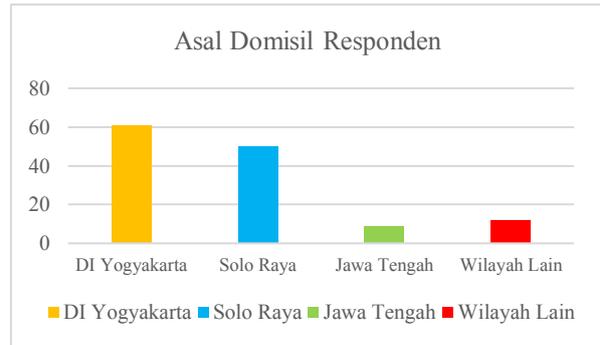
Saat dilakukan survei pembagian kuesioner kepada penumpang KRL Commuter Line rute Stasiun Yogyakarta – Stasiun Solo Balapan, terdapat total 132 responden yang bersedia mengisi formulir kuesioner penelitian ini.



Untuk jenis kelamin para responden, dari 132 responden terdapat sebanyak 82 responden berjenis kelamin laki – laki dan 50 responden berjenis perempuan.



Selanjutnya, untuk usia dari 132 para responden, peneliti membagikan menjadikan 4 kategori usia. Hasilnya untuk kategori usia 15 – 20 tahun terdapat 42 responden, kategori usia 21 – 35 tahun terdapat 74 responden, kategori usia 36 – 50 tahun sebanyak 13 orang dan kategori usia diatas 50 tahun sebanyak 3 orang.



Terakhir merupakan informasi asal domisil para responden yang terbagi menjadi 4 wilayah dari 132 responden yaitu untuk wilayah Daerah Istimewa Yogyakarta terdapat 61 responden, Solo Raya sebanyak 50 responden, Jawa Tengah sebanyak 9 responden, dan Wilayah Lain sebanyak 12 responden.

2. Hasil Kepuasan Penumpang

Setelah mengetahui infografis para responden, selanjutnya, dari sebanyak 132 responden yang didapatkan, setelah peneliti melakukan pengolahan data secara data mining dengan menggunakan Microsoft Excel dan serta melakukan kategorisasi kepuasan pelanggan terhadap lima aspek kepuasan pelayanan (Tangibels, Realiability, Responsiveness, Assurance, Emphaty) serta hasil keseluruhannya. Berikut dibawah ini merupakan hasil tersebut:

Tabel 0.8. Hasil Kepuasan Penumpang

Kategori Kepuasan	A	B	C	D	E	Hasil Keseluruhan Kepuasan
Sangat Puas	118	124	117	114	116	112
Puas	114	6	12	15	6	18
Cukup Puas	0	2	3	3	0	2
Kurang Puas	0	0	0	0	0	0
Tidak Puas	0	0	0	0	0	0

Keterangan:

- A: Aspek Layanan Secara Tangibels
- B: Aspek Layanan Secara Reliability
- C: Aspek Layanan Secara Responsiveness
- D: Aspek Layanan Secara Assurance
- E: Aspek Layanan Secara Emphaty

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

3. Hasil Feature Selection Information Gain pada Weka

Setelah dilakukan pengujian dalam software Weka dengan test options menggunakan K-Fold Cross Validation sebesar 10 Folds dan Random Seed X Validation sebesar 1, berikut ini merupakan hasil pengurutan ranking attribut menggunakan Feature Selection Information Gain terhadap 9 attribut dengan menggunakan Class atribut "LABEL KEPUASAN KESELURUHAN".

Tabel 0.9. Hasil Perangkingan Attribut Dengan Information Gain

Ranking	Nama Attribut	Value
1	IDResponden	0.68468
2	LABEL ASPEK E	0.35415
3	LABEL ASPEK D	0.31985
4	LABEL ASPEK B	0.25379
5	LABEL ASPEK C	0.23354
6	LABEL ASPEK A	0.16417
7	USIA	0.05432
8	DOMISIL	0.04021
9	JENIS KELAMIN	0.00161

4. Hasil Algoritma Naïve Bayes dengan Information Gain

Berdasarkan hasil eksperimen pada Software Weka dengan test options menggunakan K-Fold Cross Validation sebesar 10 Folds dan Random Seed X Validation sebesar 1, didapatkan hasil akurasi, error dan perhitungan confusion matrix berikut dibawah ini.

Tabel 0.10. Hasil Akurasi Eksperimen Naïve Bayes dengan IG

	Jumlah Instance	Value (%)
Akurasi	125	94.70 %
Error	7	5.30%

Tabel 0.11. Hasil Perhitungan Confusion Matrix

Class	Precision	Recall	F-Measure	ROC
Sangat Puas	97,3%	97,3%	97,3%	90,2%
Puas	82,4%	77,8%	80%	88,9%
Cukup Puas	66,7%	100%	80%	100%
Rata - Rata	94,8%	94,7%	94,7%	90,2%

Pada tabel di atas, setelah perhitungan Confusion Matrix dilakukan, nilai precision tertinggi ada pada class Sangat Puas diangka 97,3% lalu untuk nilai recall tertinggi ada pada class Cukup Puas yaitu 100% selanjutnya untuk nilai F-Measure tertinggi ada pada

class Sangat Puas yaitu dan 97,3% dan terakhir nilai ROC tertinggi pada class Cukup Puas yaitu 100%.

5. Hasil Eksperimen menggunakan Algoritma Naïve Bayes tanpa Information Gain.

Pada eksperimen kedua ini, peneliti hanya menggunakan fitur algoritma Naïve Bayes dengan test options K-Fold Cross Validation sebesar 10 Folds dan Random Seed X Validation sebesar 1 sama dengan eksperimen pertama yang telah dilakukan sebelumnya.

Berdasarkan hasil eksperimen, didapatkan hasil akurasi, error dan perhitungan confusion matrix berikut dibawah ini.

Tabel 0.12. Hasil Akurasi Eksperimen Algoritma Naive Bayes

	Jumlah Instance	Value (%)
Akurasi	122	92.42 %
Error	10	7.58%

Tabel 0.13. Hasil Perhitungan Confusion Matrix

Class	Precision	Recall	F-Measure	ROC
Sangat Puas	96,4%	95,5%	96%	90,9%
Puas	72,2%	72,2%	72,2%	89,3%
Cukup Puas	66,7%	100%	80%	100%
Rata - Rata	92,6%	92,4%	92,5%	90,8%

Pada tabel di atas, setelah perhitungan Confusion Matrix dilakukan, nilai precision tertinggi ada pada class Sangat Puas diangka 96,4% lalu untuk nilai recall tertinggi ada pada class Cukup Puas yaitu 100% selanjutnya untuk nilai F-Measure tertinggi ada pada class Sangat Puas yaitu dan 96% dan terakhir nilai ROC tertinggi pada class Cukup Puas yaitu 100%.

4. Kesimpulan

Survei kuesioner yang dilakukan oleh peneliti menunjukkan dari 132 record data menunjukkan bahwa sebanyak 112 responden sangat puas terhadap pelayanan terhadap layanan KRL Commuter Line untuk rute Yogyakarta – Solo Balapan. Disusul dengan 18 responden menjawab secara puas dan 2 orang menjawab secara cukup puas.

Selanjutnya, saat dilakukan pengujian eksperimen dengan model algoritma hasil dari kedua eksperimen diatas menunjukkan bahwa optimalisasi Information Gain dapat membantu menaikkan hasil eksperimen

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

sebesar 2,28% yang dimana hasil eksperimen tersebut menunjukkan bahwa penggunaan algoritma Naïve Bayes menggunakan dataset survei kepuasan layanan KRL Commuter Line Rute Stasiun Yogyakarta – Stasiun Solo Balapan sebesar 92.42% sedangkan untuk hasil eksperimen algoritma Naïve Bayes dengan Information Gain dengan dataset yang sama dapat meningkatkan akurasi sebesar 94,7%.

Selain pengaruh optimalisasi menggunakan fitur seleksi Information Gain, pengaruh dari handling dataset seperti jumlah kecilnya missing value dan pengolahan data dengan baik serta jumlah record isi dataset dapat mempengaruhi tingkat akurasi dalam pengujian model algoritma.

Daftar Rujukan

- [1] Kementerian Perhubungan Ditjen Perkeretaapian, "Rencana Induk Perkeretaapian Nasional (PM No.43 Tahun 2011)," *Rencana Induk Perkeretaapi. Nas.*, no. 8, 2011, [Online]. Available: <https://ppid.dephub.go.id/fileupload/informasi-berkala/RIPNAS-2030.pdf>.
- [2] D. Saidah, "Kualitas Pelayanan Commuter Line Service Quality of Commuter Line," *J. Manaj. Transp. Dan Logistik*, vol. 4, no. 1, p. 51, 2017, doi: 10.25292/j.mtl.v4i1.149.
- [3] E. Tambunan, "Analisis Kualitas Pelayanan Ka Commuter Line Rute Parungpanjang-Tanah Abang," *e-Journal CENTECH 2020*, vol. 1, no. 1, pp. 15–23, 2020.
- [4] A. S. Febriarini and E. Z. Astuti, "Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang," *Eksplora Inform.*, vol. 8, no. 2, pp. 95–103, 2019, doi: 10.30864/eksplora.v8i2.156.
- [5] F. Fauziah, D. Hartama, and I. S. Damanik, "Analisa Kepuasan Pelanggan Menggunakan Klasifikasi Data Mining," *Brahmana J. Penerapan Kecerdasan Buatan*, vol. 2, no. 1, pp. 41–48, 2020, doi: 10.30645/brahmana.v2i1.47.
- [6] J. Rizal Khairul, Muhammad. Suhartini. Sugiantara, "PENERAPAN ALGORITMA C4.5 UNTUK PREDIKSI TINGKAT KEPUASAN MAHASISWA TERHADAP SISTEM INFORMASI AKADEMIK SEKOLAH TINGGI KEGURUAN DAN ILMU PENDIDIKAN HAMZANWADI SELONG," vol. 2, no. 1, pp. 48–62, 2017.
- [7] Wiyanto, "Analisa Tingkat Kepuasan Pelanggan Terhadap Pelayanan Perusahaan Otobus XYZ Menggunakan Metode Naïve Bayes," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019, [Online]. Available: <http://krsti.stmik-amik-riau.ac.id/index.php/satin/article/view/666>.
- [8] R. Khalida and S. Setiawati, "Analisis Sentimen Sistem E-Tilang Menggunakan Algoritma Naive Bayes Dengan Optimalisasi Information Gain," *J. Inform. Inf. Secur.*, vol. 1, no. 1, pp. 19–26, 2020, doi: 10.31599/jiforty.v1i1.137.
- [9] Ezra Laurentia Widjaja, "ANALISA PENGARUH KUALITAS LAYANAN TERHADAP KEPUASAN PENUMPANG MASKAPAI PENERBANGAN BATIK AIR," vol. 3, no. Lcc, p. 2015, 2015, [Online]. Available: <http://weekly.cnbnews.com/news/article.html?no=124000>.
- [10] L. Zhang, K. Liu, I. Ilham, and J. Fan, "Application of Data Mining Technology Based on Data Center," *J. Phys. Conf. Ser.*, vol. 2146, no. 1, 2022, doi: 10.1088/1742-6596/2146/1/012017.
- [11] E. Daniati, "Decision support systems to determining programme for students using DBSCAN and naive bayes: Case study: Engineering faculty of universitas nusantara PGRI kediri," *Proceeding - 2019 Int. Conf. Artif. Intell. Inf. Technol. ICAIT 2019*, pp. 238–243, 2019, doi: 10.1109/ICAIT.2019.8834474.
- [12] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 53, 2017, doi: 10.24843/ijeet.2017.v02.i01.p11.
- [13] F. E. Prabowo and A. Kodar, "Analisis Prediksi Masa Studi Mahasiswa Menggunakan Algoritma Naïve Bayes," *J. Ilmu Tek. dan Komput.*, vol. 3, no. 2, p. 147, 2019, doi: 10.22441/jitkom.2020.v3.i2.008.
- [14] K. A. Padhana and M. Sadikin, "Analisis Sentimen Masyarakat Terhadap Kondisi Perekonomian di Indonesia Pada Masa Pandemi," vol. 5, no. 2, 2021.
- [15] H. S. Utama, D. Rosiyadi, D. Aridarma, and B. S. Prakoso, "Sentimen Analisis Kebijakan Ganjil Genap Di Tol Bekasi Menggunakan Algoritma Naive Bayes Dengan Optimalisasi Information Gain," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 247–254, 2019, doi: 10.33480/pilar.v15i2.705.

DOI: <https://doi.org/10.29207/resti.v6iX.xxx>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan tempat pengumpulan bahan material untuk kelengkapan pembuatan sebuah artikel jurnal dengan judul Analisis Kepuasan Penumpang Layanan KRL Commuter Line Rute Yogyakarta – Solo Balapan Menggunakan Algoritma Naïve Bayes Dengan Optimalisasi Information Gain. Kertas Kerja ini berisikan semua material – material hasil penelitian Tugas Akhir yang tidak dimuat atau disertakan pada artikel jurnal. Pada Kertas Kerja ini terdapat material seperti Literatur Review, Analisis dan Perancangan, Source Code, Dataset, Tahapan Eksperimen, serta Hasil Eksperimen Secara Keseluruhan.

Hasil penelitian ini merupakan hasil akurasi analisis kepuasan penumpang KRL Commuter Line yang dilakukan survei pembagian kuesioner yang dilakukan di Stasiun Yogyakarta – Stasiun Solo Balapan. Hasil akurasi ini menggunakan Algoritma Naïve Bayes dengan Optimalisasi Information Gain, yang diharapkan hasil penelitian ini dapat menjadi bahan evaluasi layanan KRL Commuter Line di rute Stasiun Yogyakarta – Stasiun Solo Balapan kepada PT. Kereta Commuter Indonesia.

UNIVERSITAS
MERCU BUANA