

KLASIFIKASI PEMBAGIAN ARUS LALU LINTAS MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN MODEL LINEAR

Muhammad Firdaus Abdi¹⁾, Sri Yanto Qodarbaskoro²⁾, Aisha Alfani³⁾, Kusri⁴⁾, Dina Maulina⁵⁾

¹Magister Teknik Informatika, Universitas AMIKOM Yogyakarta
mfirdaus.abdi29@gmail.com

²Magister Teknik Informatika, Universitas AMIKOM Yogyakarta
sri.1269@amikom.ac.id

³ Magister Teknik Informatika, Universitas AMIKOM Yogyakarta
aisha.1285@amikom.ac.id

⁴Universitas AMIKOM Yogyakarta
kusri@amikom.ac.id

⁵Universitas AMIKOM Yogyakarta
dina.m@amikom.ac.id

Abstract

The density of traffic flow is a problem for every big city, especially as it is easy to have a private vehicle, causing the flow to increase every year. So to overcome traffic flow, a system that can make optimal traffic performance is needed. The purpose of this study is to determine whether the road conditions are empty, smooth, dense and very congested so as to produce a prediction of road options whether to continue passing the road or find another way, as well as to test the accuracy of traffic flow using the naive bayes method and the liner model. The classification stages carried out are data input, data preprocessing, classification, and the results of accuracy, precision, and recall. And the results of this study the naive bayes method obtained higher accuracy than the linear model, namely for naive bayes accuracy 95.70%, precision 95.67%, and recall 100%, while for naive bayes accuracy 92.10%, precision 95.68%, and recall 96.20%. then the result is the naive bayes method is superior in the traffic flow data classification process. And the results of decision making obtained results from traffic flow data obtained that the road is empty so that the road can be passed without having to find another way.

Keywords - Classification, Naive Bayes, Traffic, Linear Model, Flow Density

Abstrak

Kepadatan arus lalu lintas menjadi masalah setiap kota-kota besar, apalagi seiring mudahnya dalam memiliki kendaraan pribadi sehingga menimbulkan arus yang meningkat pada setiap tahunnya. Maka untuk penanggulangan arus lalu lintas dibutuhkan sistem yang bisa membuat kinerja lalu lintas yang optimal. Tujuan penelitian ini adalah mengetahui kondisi jalan apakah lengang, lancar, padat dan sangat padat sehingga menghasilkan prediksi opsi jalan apakah tetap melewati jalan tersebut atau mencari jalan lain, serta menguji tingkat akurasi arus lalu lintas menggunakan metode naive bayes dan model liner. Dengan tahapan klasifikasi yang dilakukan yaitu input data, preprocessing data, klasifikasi, dan hasil accuracy, precision, dan recall. Dan hasil penelitian ini metode naive bayes mendapatkan accuracy lebih tinggi dari model linier yaitu untuk naive bayes accuracy 95.70%, precision 95.67%, dan recall 100%, sedangkan untuk naive bayes accuracy 92.10%, precision 95.68%, dan recall 96.20%. maka hasilnya metode naive bayes lebih unggul dalam proses klasifikasi data arus lalu lintas. Dan hasil dari pengambilan keputusan didapat hasil dari data arus lalu lintas didapatkan jalan tersebut lengang sehingga jalan tersebut dapat dilalui tanpa harus mencari jalan lain.

Kata Kunci - Klasifikasi, Naive Bayes, Lalu Lintas, Model Linier, Kepadatan Arus

1. PENDAHULUAN

Kepadatan arus lalu lintas pada daerah kota-kota besar meningkat pada setiap tahunnya

(Yuniar Farida, dkk 2020), bertambahnya kepemilikan kendaraan bermotor baik roda dua maupun roda empat menjadi faktor membuat padatnya arus lalu lintas itu sendiri, apalagi oleh

pihak pemerintah khususnya di Indonesia dipermudah dalam proses memiliki kendaraan bermotor baik roda dua maupun 4 (Teuku Didi Rosadi, dkk 2019), tidak ada batasan untuk memiliki kendaraan pribadi.

Penelitian tentang arus lalu lintas yang dilakukan oleh (Stathopoulos, 2018) tentang analisis risiko kinerja sistem transportasi dan pengembangan metode pengendalian lalu lintas di daerah perkotaan. Adapun yang dilakukan oleh (Yao and Ye, 2020) memprediksi keselamatan lalu lintas jalan bebas hambatan dan mewujudkan arus lalu lintas dalam data besar nonlinier menggunakan algoritma *deeplearning* dan *naive bayes*. Dari penelitian tersebut bahwa lalu lintas yang padat dapat menimbulkan resiko kecelakaan dan dapat merugikan penggunaannya.

Seperti yang dilakukan oleh (Zhang *et al.*, 2019) arus lalu lintas dapat dilihat dan digambarkan kecepatan arus laju kendaraan dan keadaan arus jalan tersebut menggunakan *Model Based on Cellular Automata and Naive Bayes*. Prediksi kondisi lalu lintas berdasarkan jenis kemacetan yang dilakukan (Chung *et al.*, 2020) berdasarkan kemacetan arus dikategorikan ekstrim, berat, dan sedang dengan strategi kontrol *METANET*.

Penggunaan metode *naive bayes* dalam klasifikasi aktifitas pengendara oleh (Haniah Mahmudah, dkk 2020) akurat tertinggi dalam penerapan algoritma *naive bayes* yaitu 99.27%. penelitian tersebut dilakukan untuk mendapatkan parameter indentifikasi aktifitas berkendara yang dapat diterapkan sebagai deteksi kecelakaan.

Parameter dalam aktifitas jalan raya seperti yang dilakukan oleh (Muhammad Yusuf Fadhlán, dkk 2021) yaitu padat, lengang, ramai dan sangat ramai menjadi acuan terhadap kelancaran suatu jalan yang dilintasi. Yang dimana suatu kendaraan akan di deteksi jumlahnya untuk mencari acuan penentuan parameter tersebut.

Arus lalu lintas yang lancar mempermudah setiap individu untuk bepergian ke tempat tujuan yang diinginkan, sehingga jika kondisi arus yang tidak lancar dapat menyebabkan arus macet (Marwan Lubis, dkk 2020) dan bisa menyebabkan kepadatan antrian kendaraan bermotor baik roda dua maupun empat (AR Muslikh, dkk 2018).

Dari penelitian penelitian terdahulu *naive bayes* digunakan sebagai pilihan untuk mengklasifikasi arus lalu lintas (Dhina Setyo

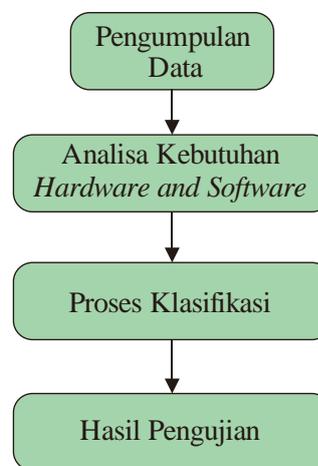
Oktaria, 2020), untuk itu dalam penelitian ini dilakukan *comparasi* atau menggabungkan metode *naive bayes* dan *model linier* (Iis Nurhasanah, dkk 2020) untuk mencari akurasi tertinggi dari dua algoritma yang dilakukan.

Selanjutnya penelitian ini akan menggunakan parameter seperti yang dilakukan oleh (Muhammad Yusuf Fadhlán, dkk 2021) kondisi jalan raya untuk menentukan jalan tersebut lengang, lancar, padat dan sangat padat. Maka penelitian ini menggunakan 4 kelas, dan akan menghasilkan usulan terhadap pengendara agar dapat melewati jalan tersebut atau melewati jalan lain.

2. METODE PENELITIAN

Tahapan penelitian ini dilakukan 4 tahapan yang dimana setiap tahapan dari mendapatkan data sampai hasil dari pengujian metode *naive bayes* dan *model linier*.

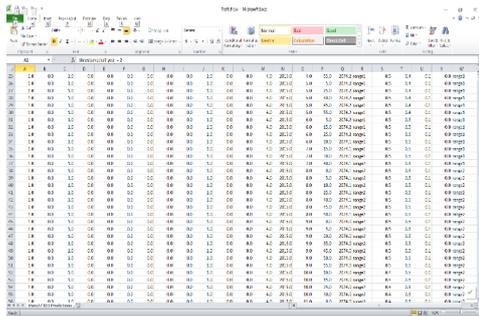
Tahapan yang dilakukan diantaranya : pengumpulan data, analisa kebutuhan hardware dan software, proses klasifikasi, hasil dari perbandingan dari metode *naive bayes* dan *model linier*.



Gambar 1. Alur Penelitian

A. Pengumpulan Data

Data yang digunakan pada penelitian ini yaitu data *Traffic flow* yang diperoleh dari <https://www.kaggle.com/coplin/traffic/data> sebagai data untuk pengujian menentukan apakah kondisi jalan lengang, lancar, padat maupun sangat padat. Penelitian ini juga membandingkan akurasi metode terbaik dari perbandingan *naive bayes* dan *model linier*.



Gambar 2. Dataset *Traffic Flow*

B. Analisa Kebutuhan *Hardware and Software*

Tahapan ini analisa tentang kebutuhan perangkat lunak dan perangkat keras yang dibutuhkan untuk menjalankan sebuah replikasi data.

Tabel 1. Kebutuhan Perangkat Lunak

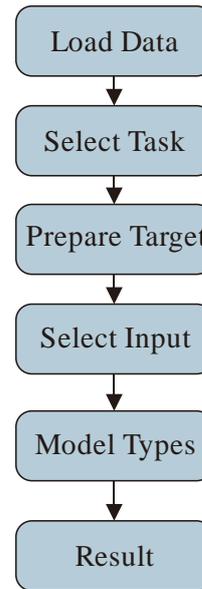
No.	Perangkat Lunak	Versi
1	Sistem operasi	Windows 10
2	Aplikasi	Rapidminer

Tabel 2. Kebutuhan Perangkat Keras

No.	Perangkat Lunak	Versi
1	Prosesor	Intel(R) Core(TM) i5-6200U CPU @2.30GHz 2.40GHz
2	Penyimpanan (RAM)	8 GB DDR4
3	Hardisk	500 SSD
4	Grafis	NVIDIA GeForce 2GB

C. Proses Klasifikasi

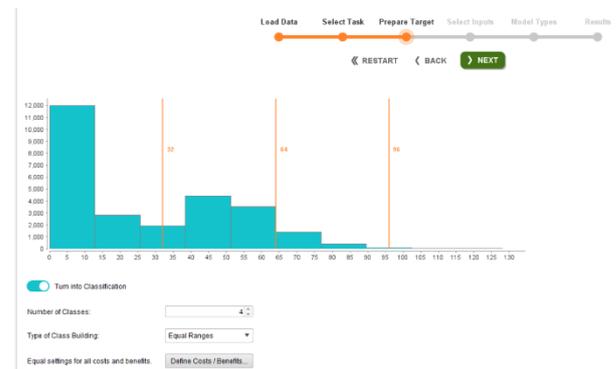
Proses klasifikasi ini dilakukan menggunakan aplikasi rapidminer, yang dimana tahapan-tahapan dilakukan dari *load data*, *select task*, *prepare target*, *select input*, *model types*, *result*.



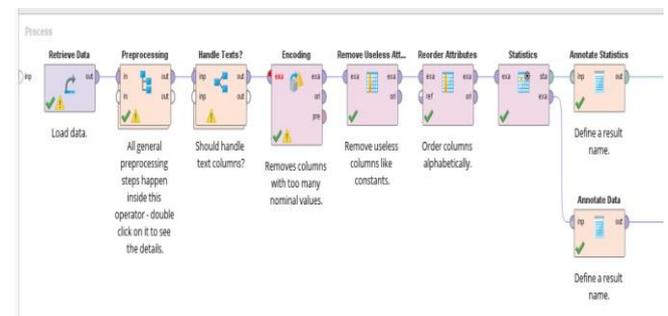
Gambar 3. Tahapan Klasifikasi

D. Tahapan Pengujian

Berikut gambar dari tahapan pengujian menggunakan rapidminer yang dimana kelas ditentukan yaitu 4 kelas.



Gambar 4. Pembagian Kelas Klasifikasi



Gambar 5. Proses pada rapidminer

Gambar diatas yaitu bagaimana langkah proses klasifikasi pada rapidminer tahapan-tahapan yang dilakukan pada penelitian ini.

E. Tahapan Pengambilan Keputusan

Tahapan ini untuk menentukan keputusan jalan yang akan diambil untuk terus jalan atau mencari alternatif jalan lain.

Tabel 3. Tahapan Pengambilan Keputusan

Kondisi Jalan	Opsi jalan
Lengang	jalan terus
Lancar	jalan terus
Padat	jalan terus/ lewat jalan lain
Sangat Padat	Lewat jalan lain

Jika kondisi jalan lengang dan lancar maka akan menghasilkan opsi jalan terus atau bisa dilewati, untuk padat bisa dilewati atau mencari jalan lain, dan sangat padat bisa mencari jalan lain untuk terhindar dari kemacetan.

Tabel 4. Tahapan Pengambilan Data

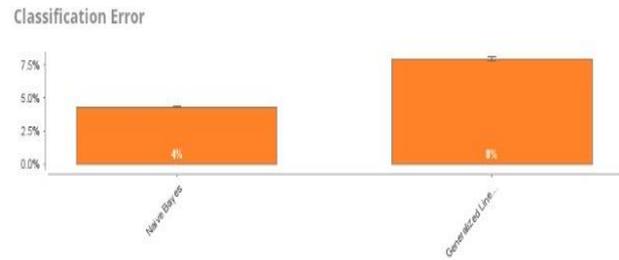
Data dibagi menjadi 4 kelas	Jumlah (range/kondisi) terjadi
range1	354853
range2	11973
range3	3248
range4	870

Dari data diatas kondisi jalan dibagi dalam 4 kelas yang dimana range 1 itu lengang,range 2 itu lancar, range 3 itu padat dan range 4 itu sangat padat. Dari data itu nantinya akan diklasifikasikan untuk mendapatkan hasil apakah jalan tersebut lengang, lancar, padat maupun sangat padat.

3. HASIL DAN PEMBAHASAN

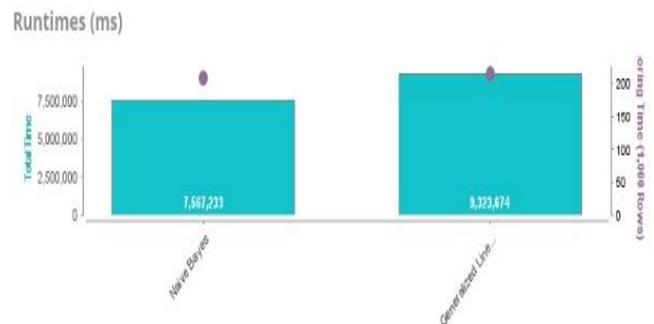
3.1. Hasil pengujian

Hasil dari penelitian yang telah dilakukan didapatkan hasil yaitu :



Gambar 6. Classification error

Dalam proses yang dilakukan didapat kesalahan atau eror pada klasifikasi yaitu 4,3% pada Naive Bayes dan 7,9% pada Model Linier.



Gambar 7. Proses waktu klasifikasi

Proses waktu total untuk klasifikasi dengan Naive Bayes yaitu 2 jam 6 menit, dan untuk dengan Model Linier yaitu 2 jam 35 menit, dibutuhkan waktu sedikit lebih lama untuk proses klasifikasi menggunakan Model Linier.

Tabel 5. Hasil Pengujian

	Accuracy	Precision	Recall
Naive Bayes	95,70%	95,67%	100%
Model Linier	92,10%	95,68%	96,20%

Pengujian yang telah dilakukan didapatkan Metode Naive Bayes mendapat accuracy 95.70%, precision 95.67%, dan recall 100%, untuk Metode Model Linier mendapat accuracy 92.10%, precision 95.68, dan recall 96.20%.

Keputusan yang didapat dari tahapan pengambilan keputusan pada data traffic flow yang digunakan didapat parameter kondisi jalan yaitu lengang, lancar, padat dan sangat padat

untuk menentukan opsi jalan apakah jalan terus atau mencari jalan lain. Hasil klasifikasi menggunakan Naive Bayes dan Model Linier menghasilkan prediksi bahwa kondisi jalan lengang.

Tabel 5. Hasil Prediksi Menggunakan Naive Bayes dan Model Linier

Prediksi	Kondisi Jalan	Opsi Jalan
Naive Bayes	Lengang	Jalan terus
Model Linier	Lengang	Jalan terus

Dari hasil prediksi diatas menghasilkan opsi jalan yang bisa diambil yaitu tetap jalan terus karena prediksi jalan dapat dilewati atau kondisi lengang.

4. KESIMPULAN

Dari penelitian ini dapat disimpulkan bahwa proses klasifikasi *traffic flow* dengan 2 metode yang diterapkan yaitu Naive Bayes dan Model Linier, didapatkan Metode Naive Bayes Lebih unggul dibanding Model Linier dengan *accuracy* yang didapat dengan Naive Bayes yaitu 95.70% sedangkan Model Linier hanya 92.10%. Dan untuk pengambilan keputusan yang dilakukan dari kondisi jalan lengang, lancar, padat, sangat padat, menghasilkan usulan opsi jalan lengang dari prediksi yang dilakukan menggunakan Naive Bayes dan Model Linier.

Untuk usulan penelitian selanjutnya bisa dengan metode yang berbeda, berbeda hardware yang digunakan juga bisa mempengaruhi hasil akurasi.

5. REFERENSI

A. Artikel Jurnal

Chung, W. *et al.* (2020) 'Development of Decision Support System for Integrated Active Traffic Management Systems Considering Travel Time Reliability', *Transportation Research Record*. doi: 10.1177/0361198120905591.

Stathopoulos, A. (2018) 'Decision Support System for Network Traffic Control Risk Management', *International Journal of Scientific*

& Engineering Research. doi: 10.14299/ijser.2018.10.17.

Yao, J. and Ye, Y. (2020) 'The effect of image recognition traffic prediction method under deep learning and naive Bayes algorithm on freeway traffic safety', *Image and Vision Computing*. doi: 10.1016/j.imavis.2020.103971.

Zhang, H. *et al.* (2019) 'The study of traffic flow model based on cellular automata and Naive Bayes', *International Journal of Modern Physics C*. doi: 10.1142/S0129183119500347.

Yuniar Farida, dkk. (2020) 'Pemodelan Arus Lalu Lintas Dan Waktu Tunggu Total Optimal Di Persimpangan Jl. Jemur Andayani Ahmad Yani Sebagai Upaya Mengurai Kemacetan', *Jurnal Ilmu Matematika dan Terapan*.

Marwan Lubis, dkk (2020) 'Evaluasi Kinerja Persimpangan Akibat Adanya Fly Over Jamin Ginting Terhadap Pergerakan Arus Lalu Lintas', *Buletin Utama Teknik Vol. 15, No. 2*.

Haniah Mahmudah, dkk (2020) 'Metode Naive Bayes Classifier – Smoothing pada Sensor Smartphone untuk Klasifikasi Aktivitas Pengendara', *Jurnal Nasional Teknik Elektro dan Teknologi Informasi / Vol. 9, No. 3*.

Dhina Setyo Oktaria (2020) 'Pengaruh Lalu-lintas Kereta Api dan Shift Jaga di Era Pandemi Covid-19 Terhadap Persepsi dari Penjaga Perlintasan Kereta Api (Pjl) di Wilayah Kerja Daerah Operasi (DAOP) 1 Jakarta', *Jurnal Dinamika Volume 1 No. 2*.

Iis Nurhasanah, dkk (2020) 'Hubungan Antara Volume, Kecepatan Dan Kepadatan Lalu Lintas Dengan Menggunakan Metode Linier', *Jurnal Rekayasa Sipil*.

Teuku Didi Rosadi, dkk (2019) ‘Penggunaan Okupansi dan Komposisi Kendaraan Untuk Menentukan Ekivalensi Mobil Penumpang (Emp) Pada Lalu Lintas Campuran Di Bundaran Empat Lengan’, *Teras Jurnal vol.9 no.2*.

Teuku Didi Rosadi, dkk (2019) ‘Klasifikasi Data Time Series Arus Lalu Lintas Jangka Pendek Menggunakan Algoritma Adaboost Dengan Random Forest’, *Jurnal Teknologi Informasi CyberKU Vol.14 no 1*.

Muhammad Yusuf Fadhlan, dkk (2021) ‘Implementasi algoritma pendeteksi tingkat kepadatan lalu lintas menggunakan metode background subtraction’, *JITEL (Jurnal Ilmiah Telekomunikasi, Elektronika, dan Listrik Tenaga) Vol. 1, No. 1, Maret 2021, pp. 59-68*.