

Klasifikasi Tingkat Keparahan Korban Kecelakaan Lalu Lintas Menggunakan *Naïve Bayes*

Said Noor Abdullah^{1*}, Siska Kurnia Gusti^{2*}, Fitri Wulandari³, Fadhilah Syafria⁴
 Program Studi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau,
 Pekanbaru, Indonesia

*e-mail *Corresponding Author*: 11750114964@students.uin-suska.ac.id

Abstract

A traffic accident is an unpredictable and unintentional event between other road users that results in human victims experiencing minor injuries, serious injuries, loss of property, and death. Accidents in the city of Pekanbaru tend to increase every year, based on data obtained from the Polresta in the city of Pekanbaru from 2015 to March 2021. Further analysis is needed regarding the severity of traffic accident victims, so use the data mining method using the naïve Bayes classification technique. The research was carried out through cleaning, transformation, and feature selection processes. Attributes that influence determining the severity of traffic accident victims in Pekanbaru City are time, age, vehicle, type of accident, and crash opponent. Then the testing process was carried out and the results obtained an accuracy of 57%.

Keywords: *Data Mining; Classification; Naïve Bayes; Accident; Traffic.*

Abstrak

Kecelakaan lalu lintas merupakan suatu peristiwa yang tidak dapat diprediksi dan tidak disengaja antara pemakai jalan lainnya yang mengakibatkan korban manusia mengalami luka ringan, luka berat, mengalami kerugian harta benda hingga meninggal dunia. Kecelakaan di kota Pekanbaru setiap tahunnya cenderung semakin bertambah, berdasarkan data yang diperoleh dari Polresta di kota Pekanbaru dari tahun 2015 hingga maret 2021. Diperlukan analisis lebih lanjut mengenai tingkat keparahan korban kecelakaan lalu lintas maka penggunaan metode data mining dengan menggunakan teknik klasifikasi naïve bayes. Penelitian dilakukan melalui proses cleaning, transformasi dan feature selection. Atribut yang berpengaruh dalam menentukan tingkat keparahan korban kecelakaan lalu lintas di Kota Pekanbaru adalah waktu, usia, kendaraan, jenis kecelakaan dan lawan tabrak. Kemudian dilakukan proses pengujian dan didapatkan hasil akurasi sebesar 57%.

Kata kunci: *Data Mining; Klasifikasi; Kecelakaan; Lalu Lintas; Naïve Bayes.*

1. Pendahuluan

Lalu lintas dan angkutan jalan suatu hal yang penting dalam meningkatkan mobilitas sosial dan sangat sering diperlukan oleh masyarakat. Setiap saat masyarakat terus beradaptasi dengan angkutan jalan dengan berbagai macam kepentingan[1]. Pada saat berlalu lintas masyarakat tidak luput dari berbagai peristiwa, salah satunya peristiwa kecelakaan lalu lintas. Kecelakaan lalu lintas merupakan suatu peristiwa yang tidak dapat di prediksi dan tidak disengaja antara pemakai jalan lainnya yang mengakibatkan korban manusia mengalami luka ringan, luka berat, mengalami kerugian harta benda hingga meninggal dunia. Kebanyakan masyarakat lebih memilih menggunakan kendaraan pribadi karena dinilai lebih fleksibel dan juga lebih hemat. Peningkatan jumlah kendaraan mengakibatkan terjadinya permasalahan transportasi seperti pelanggaran lalu lintas yang cukup tinggi sehingga memicu terjadinya kecelakaan lalu lintas [2].

Kecelakaan di kota Pekanbaru setiap tahunnya cenderung semakin bertambah, berdasarkan data yang diperoleh dari Polresta di kota Pekanbaru dari tahun 2015 hingga maret 2021 kasus kecelakaan yang terjadi setiap tahunnya semakin meningkat seperti pada tahun 2019 terdapat 355 kasus kecelakaan dengan jumlah korban meninggal sebanyak 54 jiwa, korban dengan kondisi luka berat sebanyak 105 jiwa dan korban dengan kondisi luka ringan sebanyak 205 jiwa.

Berdasarkan data yang diuraikan diatas, diperlukan analisis lebih lanjut mengenai tingkat keparahan korban kecelakaan lalu lintas maka menggunakan metode data mining dengan menggunakan teknik klasifikasi naïve bayes merupakan solusi yang layak untuk diajukan. Klasifikasi adalah proses pengidentifikasian obyek ke dalam sebuah kategori, kelas atau kelompok berdasarkan prosedur, definisi dan karakteristik yang telah ditentukan sebelumnya. Klasifikasi bertujuan untuk menempatkan objek yang ditugaskan hanya ke salah satu kategori yang disebut kelas [3]. Salah satu metode klasifikasi yang sering digunakan adalah metode naïve bayes. Naïve Bayes memiliki bukti kecepatan dan akurasi yang tinggi saat digunakan ke dalam kumpulan data data yang besar dan melakukan pengklasifikasian dengan menghitung nilai probabilitas untuk setiap kejadian dari atribut target pada setiap kasus. Klasifikasi naïve bayes memiliki fungsi untuk menemukan pengetahuan atau pola-pola kesamaan karakteristik dalam suatu kelompok atau kelas tertentu [4]. Metode Naïve Bayes Classifier juga dinilai berpotensi baik dalam mengklasifikasi dokumen dibandingkan metode pengklasifikasian yang lain dalam hal akurasi dan efisiensi komputasi[5].

Dalam penelitian ini, analisis data diperoleh langsung dari pihak Polresta Kota Pekanbaru. Berdasarkan data yang telah didapat dan variabel yang akan diinputkan, maka hasil klasifikasi diharapkan dapat menentukan atribut yang paling berpengaruh dalam menentukan tingkat keparahan korban kecelakaan lalu lintas.

2. Tinjauan Pustaka

Penelitian mengenai analisis korban kecelakaan lalu lintas telah banyak dilakukan sebelumnya. Diantaranya yaitu Analisis Karakteristik Kecelakaan Lalu Lintas pada Jalan Yos Sudarso Kota Pekanbaru yang dilakukan oleh Tomi Yandi, Fadrizal Lubis, Winayati[6]. Pada penelitian ini, berdasarkan hari kejadian yaitu hari minggu, berdasarkan waktu yaitu pada pagi hari jam 06.00-12.00, berdasarkan tipe tabrakan yaitu pada tipe depan belakang, berdasarkan jenis kendaraan yaitu pada kendaraan bermotor, berdasarkan jenis kelamin yaitu pada laki-laki, berdasarkan usia yaitu pada usia 16-30 tahun, berdasarkan penyebab yaitu pada manusia.

Penelitian yang dilakukan oleh Daniar Mukti Septianingtyas dengan judul Identifikasi Karakteristik Korban Kecelakaan Lalu Lintas di Sidoarjo Tahun 2016[7]. Pada penelitian ini, terlihat bahwa sebagian besar korban kecelakaan lalu lintas di Kabupaten Sidoarjo tahun 2016 cenderung mengalami luka ringan, berjenis kelamin laki-laki dengan usia produktif yaitu lebih besar sama dengan 34 tahun, profesi sebagai pekerja, berkendara, pada waktu sore hari, dengan jenis tabrak lurus, mengendarai sepeda motor dan begitu pula dengan jenis kendaraan lawan berupa sepeda motor. Faktor yang berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas di Kabupaten Sidoarjo tahun 2016 didapatkan pemodelan $3,133X4.1+1,464X5.1$ menunjukkan nilai Y sebesar 4,597.

Penelitian dengan judul K-Means Clustering untuk Data Kecelakaan Lalu Lintas Jalan Raya di Kecamatan Pelaihari oleh Winda Aprianti dan Jaka Permadi[8]. Pada penelitian ini, penyebab kecelakaan pada dataset adalah karena kesalahan pengemudi. Korban kecelakaan yang paling sering rentang usia 9 sampai 28 tahun dan keadaan yang paling sering dialami korban kecelakaan lalu lintas adalah luka ringan. Waktu terjadinya kecelakaan dengan peluang terbesar adalah saat hari kerja dan di pagi hari.

Ada beberapa penelitian tentang metode naïve bayes yang telah dilakukan oleh peneliti sebelumnya, diantaranya penelitian yang berjudul Perancangan Sistem Klasifikasi Menggunakan Metode Naïve Bayes (Studi Kasus Kecamatan Tibawa Kab. Gorontalo) untuk klasifikasi masyarakat miskin menghasilkan kesimpulan bahwa sistem klasifikasi dapat direkayasa dan berdasarkan hasil pengujian confusion matrix dengan teknik *split* validasi, penggunaan metode klasifikasi naïve bayes terhadap dataset yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori *Good*. Sementara nilai *Precision* sebesar 92% dan *Recall* sebesar 86% [9].

Berikutnya penelitian yang dilakukan PDAM Tirta Lihou menggunakan naïve bayes diperoleh klasifikasi dari 19 alternatif yang digunakan, dimana terdapat 8 kelas Layak dan 11 Tidak Layak dengan total *Accuracy* yang diperoleh sebesar 78,95%. Dari hasil yang diperoleh, diharapkan penelitian ini dapat membantu pihak PDAM Tirta Lihou dalam menentukan lokasi yang layak dilakukan pembangunan sumber air sehingga dapat memenuhi kebutuhan masyarakat. Penelitian ini juga diharapkan dapat menjadi referensi bagi peneliti selanjutnya yang berkaitan dengan pengguna algoritma yang digunakan [10].

Selanjutnya penelitian yang dilakukan di STIKOM Tunas Bangsa Pematangsiantar dalam menentukan kepuasan mahasiswa terhadap cara pengajaran dosen menggunakan metode naïve bayes menghasilkan kesimpulan. Dari hasil data yang dapat digunakan sebagai salah satu metode untuk klasifikasi dalam menentukan kepuasan mahasiswa terhadap cara pengajaran dosen. Data testing kelima data menyatakan puas terhadap cara pengajaran dosen di STIKOM Tunas Bangsa. Sedangkan data training yang di proses dengan Rapid Miner memiliki akurasi sebesar 92.00% [11].

Pada prinsipnya semua penelitian [4], [5] dan [6] memiliki kesamaan dengan penelitian yang penulis lakukan, yaitu kesamaan mengklasifikasi tingkat keparahan korban kecelakaan lalu lintas dan penelitian [7], [8] dan [9] ini menggunakan metode Naïve Bayes. Penelitian ini juga menggunakan metode Naïve Bayes yaitu dengan membagi permasalahan ke dalam sebuah kelas-kelas berdasarkan ciri-ciri persamaan dan perbedaan dengan menggunakan statistik yang bisa mengklasifikasi probabilitas sebuah kelas.

3. Metodologi

Adapun tahapan penelitian terlihat pada gambar 1.



Gambar 1. Tahapan Penelitian

- 1) Perumusan Masalah
Perumusan masalah dalam penelitian ini adalah bagaimana penerapan algoritma naïve bayes untuk menentukan tingkat keparahan korban kecelakaan lalu lintas di Kota Pekanbaru dan bagaimana menentukan tingkat akurasi algoritma naïve bayes untuk menentukan tingkat keparahan korban kecelakaan lalu lintas di Kota Pekanbaru.
- 2) Studi Literatur
Studi literatur yang dilakukan penulis adalah dengan membaca sumber-sumber tertulis, seperti jurnal ilmiah dari penelitian terdahulu dan referensi lainnya yang berguna sebagai dasar acuan melakukan penelitian.
- 3) Analisis Kebutuhan Data
Dalam penelitian ini, data yang digunakan adalah kumpulan data Kecelakaan Lalu Lintas di Kota Pekanbaru periode Desember 2015 sampai Maret 2021. Data tersebut diperoleh dari Polresta Kota Pekanbaru. Dari data yang diperoleh, didapatkan data sebanyak 1765 data. Dalam data tersebut terdapat beberapa atribut sebagai berikut:

Tabel 1. Atribut Dalam Data Kecelakaan Lalu Lintas di Kota Pekanbaru

No	Parameter
1	Tanggal Kejadian
2	Hari
3	Waktu
4	Tempat Kejadian
5	Kecamatan
6	Jenis Kelamin
7	Usia
8	Pekerjaan
9	Kondisi Jalan

No	Parameter
10	Kendaraan
11	Jenis Kecelakaan
12	Penyebab
13	Lawan Tabrakan
14	Keadaan Luka

4) Analisa Metode Naïve Bayes

Metode yang digunakan untuk mengklasifikasi tingkat keparahan korban kecelakaan lalu lintas adalah metode klasifikasi dalam data mining yaitu naïve bayes. Metode Naïve Bayes adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk *machine learning* dan data mining. Performa naïve bayes yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi keidependenan atribut (tidak ada kaitan antar atribut). Asumsi keidependenan atribut ini pada data sebenarnya jarang terjadi, namun walaupun asumsi keidependenan atribut tersebut dilanggar, performa naïve pengklasifikasian naïve bayes cukup tinggi [12]. *Naïve bayes* juga mengklasifikasikan data berdasarkan peluang atribut dari setiap label data. Pada model peluang setiap label dan jumlah atribut yang dapat dituliskan seperti persamaan berikut:

$$P(C_i|X_1, X_2, \dots, X_n) \dots\dots\dots(1)$$

Perhitungan *naive bayes* dapat dijelaskan dengan C_i adalah hipotesis data yang merupakan suatu label $P(C_i|X)$ adalah peluang hipotesis label berdasarkan kategori X (*posteriori probability*). $P(C_i)$ adalah peluang hipotesis label (*prior probability*). $P(X|C_i)$ adalah peluang data berdasarkan kategori pada hipotesis label. $P(X)$ adalah jumlah probabilitas data yang nilainya adalah satu. Sehingga didapatkan rumus penghitungan *Naïve Bayes* dituliskan pada persamaan.

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \dots\dots\dots(2)$$

Untuk menentukan peluang dari masing-masing kategori label digunakan persamaan berikut:

$$P(C_i|X) = \prod_{k=1}^n P(X_k|C_i) \dots\dots\dots(3)$$

Untuk menentukan nilai label pada atribut tertentu digunakan persamaan:

$$argmax = \frac{P(X|C_i)P(C_i)}{P(X)} \dots\dots\dots(4)$$

5) Implementasi dan Pengujian

Pengimplementasian metode dalam memproses data dengan menggunakan MATLAB. MATLAB merupakan perangkat lunak yang digunakan untuk pemrograman, analisis, serta komputasi teknis dan matematis berbasis matriks[13]. Pengujian dilakukan untuk melihat akurasi perhitungan yang dilakukan. Pada penelitian ini pengujian dilakukan menggunakan *K-Fold Cross Validation*. *Cross Validation* adalah sebuah metode dari teknik data mining yang bertujuan untuk memperoleh hasil akurasi maksimum ketika data dibagi menjadi dua subset (data latih dan data uji). Salah satu dari jenis pengujian *Cross Validation* adalah *K-Fold Cross Validation* yang berfungsi untuk menilai kinerja proses sebuah metode algoritma dengan membagi sampel data secara acak dan mengelompokkan data tersebut sebanyak nilai K pada *K-Fold*. Pada pendekatan metode *K-Fold Cross Validation*, dataset dibagi menjadi sejumlah buah partisi secara acak. Data partisi tersebut diolah sejumlah K kali eksperimen dengan masing-masing eksperimen menggunakan data partisi ke-K sebagai data *testing* dan menggunakan sisa partisi lainnya sebagai *data training*[14].

6) Penarikan Kesimpulan

Perhitungan dengan menggunakan klasifikasi Naïve Bayes akan menunjukkan hasil akhir berupa kriteria-kriteria apa saja yang mempengaruhi kecelakaan lalu lintas yang terjadi di Kota Pekanbaru dan seberapa akurat metode Naïve Bayes dalam melakukan perhitungan dalam penelitian ini.

4. Hasil dan Pembahasan

1) *Data Selection*

Berdasarkan data kecelakaan lalu lintas di Kota Pekanbaru yang diperoleh dari Polresta Kota Pekanbaru terdapat total keseluruhan data yaitu 1765 data. Data tersebut dapat dilihat pada tabel 2.

Tabel 2. Data Kecelakaan Lalu Lintas di Kota Pekanbaru

No	Tanggal Kejadian	Hari	Waktu	Tempat Kejadian	Kecamatan	...	Keadaan Luka
1	04 Januari 2016	Senin	01.30	Jl. Air Hitam	Tampan	...	Meninggal Dunia
2	05 Januari 2016	Selasa	11.00	Jl. Parit Indah	Tangkerang Tengah	...	Meninggal Dunia
3	05 Januari 2016	Selasa	11.00	Jl. Parit Indah	Tangkerang Tengah	...	Luka Ringan
4	05 Januari 2016	Selasa	05.30	Jl. Arifin Ahmad	Marpoyan Damai	...	Meninggal Dunia
5	01 Januari 2016	Jum'at	20.30	Jl. Yos Sudarso	Rumbai	...	Luka Berat
6	01 Januari 2016	Jum'at	20.30	Jl. Yus Sudarso	Rumbai	...	Luka Ringan
7	08 Januari 2016	Jum'at	08.45	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat
8	09 Januari 2016	Sabtu	20.30	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat
9	12 Januari 2016	Selasa	05.00	Jl. Ahmad Yani	Senapelan	...	Luka Ringan
...
1764	25 Maret 2021	Kamis	21.00	Jl. HR Soebrantas	Tampan	...	Luka Berat
1765	06 Maret 2021	Sabtu	18.30	Jl. Melati	Bina Widya	...	Meninggal Dunia

2) *Pre-Processing / Cleaning*

Dalam proses pengolahan data, tahap pertama yang perlu dilakukan adalah proses *data cleaning* atau membersihkan data. Data mentah yang telah diperoleh perlu diseleksi kembali. Kemudian, hapus atau hilangkan data-data yang tidak lengkap, tidak relevan, dan tidak akurat. Dengan adanya proses *cleaning data*, maka jumlah data akan menjadi sedikit dibandingkan data yang diperoleh sebelumnya. Didapatkan hasil data yang telah *dicleaning* dari 1765 data menjadi 960 data. Data hasil *cleaning* dapat dilihat pada tabel 3.

Tabel 3. Data Hasil *Cleaning Data*

No	Tanggal Kejadian	Hari	Waktu	Tempat Kejadian	Kecamatan	...	Keadaan Luka
1	24 Desember 2015	Kamis	10.30	Jl. Bukit Barisan	Tenayan Raya	...	Meninggal Dunia
2	01 Januari 2016	Jum'at	20.30	Jl. Yos Sudarso	Rumbai	...	Luka Berat
3	01 Januari 2016	Jum'at	20.30	Jl. Yos Sudarso	Rumbai	...	Luka Ringan
4	05 Januari 2016	Selasa	11.00	Jl. Parit Indah	Tangkerang Tengah	...	Meninggal Dunia
5	05 Januari 2016	Selasa	11.00	Jl. Parit Indah	Tangkerang Tengah	...	Luka Ringan
6	06 Januari 2016	Rabu	05.30	Jl. Arifin Ahmad	Marpoyan Damai	...	Meninggal Dunia
7	08 Januari 2016	Jum'at	08.45	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat

No	Tanggal Kejadian	Hari	Waktu	Tempat Kejadian	Kecamatan	...	Keadaan Luka
8	09 Januari 2016	Sabtu	20.30	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat
9	12 Januari 2016	Selasa	05.00	Jl. Ahmad Yani	Senapelan	...	Luka Ringan
...
959	21 Maret 2021	Minggu	19.40	Jl. Garuda Sakti	Tampar	...	Luka Ringan
960	22 Maret 2021	Senin	15.30	Diatas Jembatan	Rumbai	...	Meninggal Dunia

3) Transformation

Selanjutnya tahap transformasi. Tahapan ini akan menggabungkan beberapa data dalam suatu data set, kemudian menyamakan seluruh data yang terkumpul agar dapat mempermudah proses analisis data. Berikut beberapa data yang dikategorikan sebagai berikut:

Tabel 4. Kategori Data	
Usia	(0 –15 tahun) Usia Anak-Anak dan Remaja
	(16 –35 tahun) Usia Muda
	(36 –55 tahun) Usia Dewasa
	(> 55 tahun) Lanjut Usia
Waktu	Dini Hari (22.01 – 06.29)
	Pagi Hari (06.30 – 10.00)
	Siang Hari (10.01 – 14.59)
	Sore Hari (15.00 – 18.00)
	Malam Hari (18.01 – 22.00)

Data hasil transformasi dapat dilihat pada tabel 5.

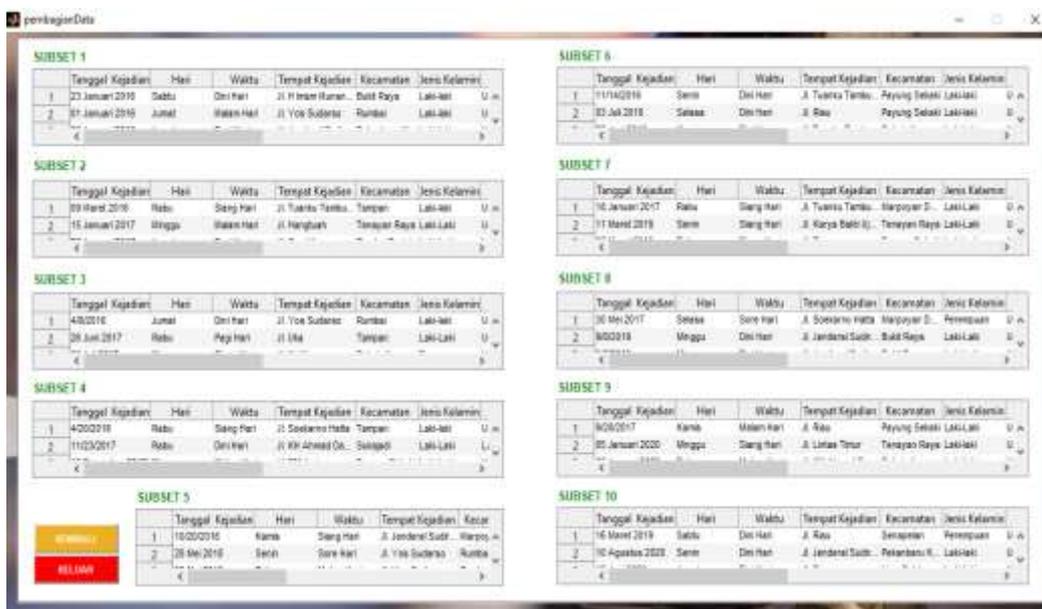
Tabel 5. Data Hasil Transformasi							
No	Tanggal Kejadian	Hari	Waktu	Tempat Kejadian	Kecamatan	...	Keadaan Luka
1	24 Desember 2015	Kamis	Siang Hari	Jl. Bukit Barisan	Tenayan Raya	...	Meninggal Dunia
2	01 Januari 2016	Jum'at	Malam Hari	Jl. Yos Sudarso	Rumbai	...	Luka Berat
3	01 Januari 2016	Jum'at	Malam Hari	Jl. Yos Sudarso	Rumbai	...	Luka Ringan
4	05 Januari 2016	Selasa	Siang Hari	Jl. Parit Indah	Tangkerang Tengah	...	Meninggal Dunia
5	05 Januari 2016	Selasa	Siang Hari	Jl. Parit Indah	Tangkerang Tengah	...	Luka Ringan
6	06 Januari 2016	Rabu	Dini Hari	Jl. Arifin Ahmad	Marpoyan Damai	...	Meninggal Dunia
7	08 Januari 2016	Jum'at	Pagi Hari	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat
8	09 Januari 2016	Sabtu	Malam Hari	Jl. Jendral Sudirman	Pekanbaru Kota	...	Luka Berat
9	12 Januari 2016	Selasa	Dini Hari	Jl. Ahmad Yani	Senapelan	...	Luka Ringan
...

No	Tanggal Kejadian	Hari	Waktu	Tempat Kejadian	Kecamatan	...	Keadaan Luka
959	21 Maret 2021	Minggu	Malam Hari	Jl. Garuda Sakti	Tampan	...	Luka Ringan
960	22 Maret 2021	Senin	Sore Hari	Diatas Jembatan	Rumbai	...	Meninggal Dunia

4.2 Pengujian

Penelitian ini dilakukan menggunakan MATLAB. Pengujian dilakukan dengan *K-Fold Cross Validation*, yaitu kumpulan data yang diberikan dibagi menjadi sejumlah K bagian / *fold* dimana setiap *fold* digunakan sebagai set pengujian di beberapa titik[15]. Pada penelitian ini diterapkan *K-Fold cross validation* sebanyak 10 fold, yaitu dengan membagi data menjadi 10 sampel. Tiap rasio dilakukan 10 kali pengujian dengan output yang diharapkan adalah *Accuracy* rata-rata.

Berikut pembagian data yang dikelompokkan menjadi 10 fold.



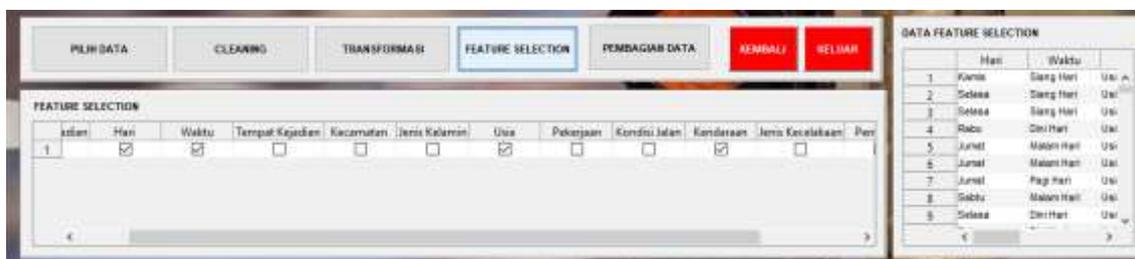
Gambar 2. Tampilan Pembagian Data

Dari pengujian 10 *fold* tersebut didapatkan hasil rata-rata paling tinggi yaitu 57% pada *fold* 2.



Gambar 3. Tampilan Hasil Pengujian

Untuk menentukan atribut yang berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas, dihitung dengan perhitungan *feature selection* dengan menggunakan *Forward selection*. Hasil yang didapat yaitu ada 5 atribut yang berpengaruh dalam menentukan tingkat keparahan korban kecelakaan lalu lintas di kota Pekanbaru. Atribut tersebut adalah hari, usia, kendaraan, jenis kecelakaan dan lawan tabrak. Berikut merupakan tampilan hasil dari *feature selection*:



Gambar 4. Tampilan *Feature Selection*.

Hasil pengujian klasifikasi menggunakan 10 *Fold Cross Validation* pada algoritma *Naïve Bayes Classifier* memperoleh akurasi tertinggi pada *fold* ke-2 dengan akurasi sebesar 57% dengan jumlah data latih sebanyak 1428 data dan jumlah data uji sebanyak 158 data, hal tersebut dapat dilihat pada gambar 3.

5. Simpulan

Pada penelitian ini hasil klasifikasi tingkat keparahan korban kecelakaan lalu lintas menggunakan naïve bayes dapat teridentifikasi dan terklasifikasikan dengan baik pada masing-masing *fold*. Terdapat 5 atribut yang berpengaruh dalam menentukan tingkat keparahan korban kecelakaan lalu lintas di kota Pekanbaru. Atribut tersebut adalah waktu, usia, kendaraan, jenis kecelakaan dan lawan tabrak. Nilai akurasi yang didapat dari hasil pengujian sebesar 57%. Hasil akurasi didapatkan dari hasil pengujian klasifikasi tertinggi dari 10 *Fold Cross Validation*.

Daftar Referensi

- [1] A. Dermawan, "Urgensi Perlindungan Hukum Bagi Korban Kecelakaan Menurut UU No. 22 Tahun 2009 Tentang Lalu Lintas dan Angkutan Jalan," *Doktrina: Journal of Law*, vol. 3, no. 1, pp. 77–86, Apr. 2020.
- [2] I. Fardian Anshori and Y. Nuraini, "Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Tasikmalaya Menggunakan Algoritma K-Means," *JURNAL RESPONSIF*, vol. 2, no. 1, pp. 118–127, 2020.
- [3] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris," *JTIK*, vol. 4, no. 3, pp. 180–184, Sep. 2017.
- [4] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 219, Dec. 2021, doi: 10.30865/jurikom.v8i6.3655.
- [5] D. Alita, I. Sari, A. Rahman Isnain, and Styawati, "Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa," *JDMSI*, vol. 2, no. 1, pp. 17–23, 2021.
- [6] T. Yandi, F. Lubis, and Winayati, "Analisis Karakteristik Kecelakaan Lalu Lintas pada Jalan Yos Sudarso Kota Pekanbaru," *JURNAL TEKNIK*, vol. 14, pp. 17–21, Apr. 2020.
- [7] D. M. Septianingtyas, "Identifikasi Karakteristik Korban Kecelakaan Lalu Lintas Di Sidoarjo Tahun 2016," *The Indonesian Journal Public Health*, vol. 14, pp. 50–60, Jul. 2019, doi: 10.20473/ijph.v14i1.2019.50-60.
- [8] W. Aprianti and J. Permadi, "K-Means Clustering Untuk Data Kecelakaan Lalu Lintas Jalan Raya Di Kecamatan Pelaihari," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 5, no. 5, pp. 613–620, Oct. 2018, doi: 10.25126/jtiik2018551113.
- [9] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes," *ILKOM Jurnal Ilmiah*, vol. 10, no. 2, pp. 160–165, Aug. 2018.

-
- [10] T. Imandasari, E. Irawan, A. Perdana Windarto, and A. Wanto, "Prosiding Seminar Nasional Riset Information Science (SENARIS) Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air," in *SENARIS*, Sep. 2019, pp. 750–761.
- [11] A. Ridwan, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," *Jurnal Sistem Komputer dan Kecerdasan Buatan*, vol. 4, no. 1, pp. 15–21, Sep. 2020.
- [12] A. Ashari Muin and Syarli, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *Jurnal Ilmiah Ilmu Komputer*, vol. 2, no. 1, 2016, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>
- [13] M. Sc. Amir Tjolleng, "Pengantar Pemrograman Matlab," in *KOMPAS GRAMEDIA*, A. Tjolleng, Ed., Jakarta: KOMPAS GRAMEDIA, 2018, pp. 1–217.
- [14] R. Rizqi Robbi Arisandi, B. Warsito, and A. Rachman Hakim, "Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation," *JURNAL GAUSSIAN*, vol. 11, no. 1, pp. 130–139, 2022, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [15] A. Peryanto, A. Yudhana, and R. Umar, "Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation," *Journal of Applied Informatics and Computing (JAIC)*, vol. 4, no. 1, pp. 45–51, Jul. 2020, [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>