KLASIFIKASI KECELAKAAN LALU LINTAS DI KOTA SAMARINDA MENGGUNAKAN ALGORITME *K-NEAREST NEIGHBOR*

Dini Anitasari [1], Fendy Yulianto [2], Taghfirul Azhima Yoga Siswa [3]
[1,2,3] Teknik Informatika, Universitas Muhammadiyah Kalimantan Timur,
Jl.Ir H Juanda, Samarinda 75124, Indonesia
fy415@umkt.ac.id

ABSTRAK

Kecelakaan lalu lintas di Kota Samarinda terus meningkat seiring dengan tingginya volume kendaraan dan berbagai faktor risiko, seperti kondisi jalan, cuaca, serta kelalaian pengemudi. Peningkatan angka kecelakaan ini berdampak pada kerugian material dan korban jiwa, sehingga diperlukan metode prediksi yang efektif untuk mengklasifikasikan tingkat keparahan kecelakaan. Penelitian ini bertujuan untuk mengklasifikasikan tingkat kecelakaan lalu lintas di Kota Samarinda menggunakan algoritme *K-Nearest Neighbor* (KNN), yang dikenal sebagai metode berbasis jarak yang efektif dalam klasifikasi data. Data yang digunakan diperoleh dari Kepolisian Sektor Samarinda Kota dan mencakup faktor-faktor seperti kondisi cahaya, kelas jalan, tipe jalan, dan batas kecepatan. Model KNN diimplementasikan dengan pembagian data latih dan uji menggunakan validasi silang *K-Fold* untuk memastikan keakuratan prediksi. Hasil pengujian menunjukkan bahwa dengan pemilihan parameter yang optimal, model KNN mampu mencapai akurasi sebesar 92,54%. Kesimpulan dari penelitian ini adalah bahwa metode KNN dapat digunakan secara efektif untuk mengklasifikasikan tingkat kecelakaan lalu lintas, sehingga hasilnya dapat menjadi referensi bagi otoritas terkait dalam meningkatkan keselamatan jalan raya di Kota Samarinda.

Keyword : *K-Nearest Neighbor*, klasifikasi, kecelakaan lalu lintas, *K-Fold.*

1. PENDAHULUAN

Kecelakaan lalu lintas banyak terjadi setiap tahunnya di Kota Samarinda yang memiliki volume lalu lintas kendaraan tinggi. Adanya volume kendaraan yang tinggi mengakibatkan kecelakaan lalu lintas seringkali terjadi secara tidak disengaja dengan melibatkan kendaraan di jalan, sehingga dapat mengakibatkan adanya korban jiwa atau kerugian harta benda [1]. Kerugian yang di dapat dari kecelakaan lalu lintas berupa materil dan inmateril yang dimana kecelakaan tersebut biasanya disebabkan oleh berbagai macam faktor antara lain cuaca, kondisi jalan, dan kesalahan pengguna kendaraan [2]. Kesalahan pengguna kendaraan menjadi salah satu faktor yang paling dominan dalam kecelakaan lalu lintas di wilayah kota Samarinda, mengakibatkan banyak korban yang mengalami cedera mulai dari luka ringan hingga luka berat, serta ada juga yang meninggal dunia. Selain itu, kecelakaan-kecelakaan tersebut menyebabkan kerugian materil yang cukup signifikan, baik bagi korban maupun pemerintah setempat [3]. Dari kejadian tersebut maka perlu adanya analisa daerah rawan kecelakaan, yang dimana data kecelakaan ini diambil dari kepolisian sektor samarinda kota melalui wawancara langsung kepada operator aplikasi *Integrated Road Safety Management System* (IRSMS). Dari analisa ini diharapkan langkah penanggulangan kecelakaan lalu lintas akan lebih mudah sehingga risiko kecelakaan di masa mendatang bisa diminimalkan [4]. Dalam melakukan analisa untuk mananggulangi kecelakaan diperlukan sebuah konsep yang dapat memetakan hal tersebut, salah satunya yaitu dengan melakukan klasifikasi terkait jenis kecelakaan lalu lintas yang ada. Klasifikasi merupakan suatu teknik pengolahan data dengan membagi objek menjadi kelas yang berbeda tergantung pada jumlah kelas yang dibutuhkan [5]. Dalam melakukan proses klasifikasi membutuhkan sebuah algoritme untuk memodelkannya antara lain seperti algoritme *K-Nearest Neighbor*, *Decision Tree* dan *Naïve bayes* [6]. Penelitian oleh [7] membandingkan algoritme *K-Nearest Neighbor* dan *Naïve Bayes* untuk klasifikasi data kecelakaan, dengan hasil akurasi masing-masing sebesar 75,86% dan 79,31%. [8] menggunakan integrasi *Decision Tree* dan SMOTE untuk klasifikasi kecelakaan lalu lintas, menghasilkan akurasi 63,56% sebelum SMOTE dan meningkat menjadi 71,12% setelahnya. Penelitian ini menunjukkan perbandingan metode dan pengaruh keseimbangan data terhadap peningkatan akurasi pada analisis kecelakaan lalu lintas. *K-Nearest Neighbor* (KNN) adalah metode yang digunakan untuk mengklasifikasikan objek berdasarkan data latih yang paling dekat dengan objek tersebut [9]. Keunggulan algoritme KNN ini adalah efektif pada kumpulan data dengan jumlah besar dan memberikan hasil klasifikasi yang lebih akurat [10]. Penelitian yang dilakukan oleh [11] juga membuktikan bahwa KNN mampu menghasilkan akurasi yang lebih optimal dalam beberapa skenario dibandingkan metode klasifikasi lainnya, membuatnya cocok untuk digunakan dalam berbagai aplikasi, terutama dalam analisis data yang kompleks dan besar. Dalam penelitian, *algoritme K-Nearest Neighbor* (KNN) telah diterapkan untuk memprediksi berbagai fenomena, seperti bidang

transportasi dan keselamatan lalu lintas. Terdapat perbedaan terkait, yang membedakan penelitian ini dengan studi sebelumnya adalah fokus studi kasus dan indikator yang digunakan untuk memprediksi tingkat kecelakaan lalu lintas. Penelitian ini mengambil studi kasus di Kota Samarinda dengan menggunakan data lalu lintas yang mencakup berbagai faktor seperti kondisi cahaya, cuaca, kelas jalan, tipe jalan, kondisi permukaan jalan, batas kecepatan, status jalan, dan kemiringan jalan [12].
Berdasarkan uraian di atas, penelitian ini akan menggunakan algoritme *K-Nearest Neighbor* (KNN) untuk menghitung akurasi, *Precision*, *Recall* dan *F1-Score* pada proses klasifikasi kecelakaan lalu lintas di Kota Samarinda. Algoritme ini diterapkan dengan tujuan merancang model KNN yang mampu menghasilkan akurasi tinggi dalam memprediksi kecelakaan, dengan mempertimbangkan berbagai faktor yang memengaruhi terjadinya kecelakaan lalu lintas di Samarinda. Uji akurasi akan dilakukan untuk memastikan kehandalan model yang dihasilkan. Diharapkan, hasil penelitian ini dapat memberikan manfaat nyata bagi pihak kepolisian dan otoritas terkait dalam meningkatkan keselamatan pengguna jalan dan transportasi.

2. TINJAUAN PUSTAKA

2.1. Klasifikasi

Klasifikasi adalah proses menemukan model atau fungsi yang dapat menjelaskan atau membedakan antara konsep atau kelas dalam data, dengan tujuan untuk memprediksi kelas dari suatu objek yang belum diketahui [13]. Pada penelitian yang dilakukan oleh [14] mengenai klasifikasi kecelakaan lalu lntas di kota surakarta memperoleh akurasi sebesar 72%. Pada penelitian [15] yang juga membahas tentang klasifikasi tingkat keparahan korban pada kecelakaan lalu lintas dengan menghasilkan akurasi sebesar 86,9%. Selanjutnya juga pada penelitian [16] yang membandingkan metode SVM dan SMOTE dengan akurasi tertinggi pada SVM yaitu 93% sedangkan SMOTE sebesar 87,97%.

2.2. *K-Nearest Neighbor*

Algoritme *K-Nearest Neighbor* (KNN) merupakan metode klasifikasi dataset berdasarkan data pelatihan yang telah diklasifikasikan sebelumnya. Algoritme KNN merupakan metode untuk mengklasifikasikan objek berdasarkan data pelatihan terdekat, dimana prinsip kerjanya adalah menentukan dan mencari jarak terdekat terhadap nilai K tetangga terdekat pada data latih dan data uji [17]. Untuk menghitung nilai tetangga terdekat dapat menggunakan Persamaan 1.

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (xi - yi)^2} \quad (1)$$

Keterangan:
d(x,y) = Jarak
xi = Sampel data
yi = Sampel data *Testing*
n = Dimensi data
I = Variabel data

2.3. Akurasi

Akurasi adalah perbandingan jumlah prediksi dengan nilai positif dari seluruh kelas positif yang diprediksi dengan benar. Akurasi digunakan untuk mendapatkan gambaran umum tentang seberapa baik Model melakukan prediksi secara keseluruhan [18]. Perhitungan nilai akurasi ditunjukkan pada Persamaan 2.

$$Accuracy \frac{\text{TP + TN}}{TP+TN+FP+FN} \quad (2)$$

Keterangan:
TP = *True Positive*
FP = *False Positive*
FN = *False Negative*
TN = *True Negative*

2.4. *Precision*

Precision digunakan untuk menunjukan proporsi prediksi kasus positif yang bernilai positif. *Precision* digunakan untuk memastikan bahwa setiap prediksi positif yang dibuat oleh model memiliki kemungkinan besar benar [19]. Untuk menghitung nilai *Precision* dapat menggunakan Persamaan 3.

$$Precision = \frac{\text{TP}}{TP+FP} \quad (3)$$

Keterangan:
TP = *True Positive*
FP = *False True*

2.5. *Recall*

Recall adalah fraksi dari elemen *True Positive* dibagi dengan keseluruhan unit yang diklasifikasikan sebagai kelas positif yaitu, jumlah *True Positive* [20]. Untuk mengitung nilai *Recall* dapat menggunakan Persamaan 4.

$$Recall = \frac{\text{TP}}{\text{TP+FN}} \quad (4)$$

Keterangan:
TP = *True Positive*
FN = *False Negative*

2.6. *F1-Score*

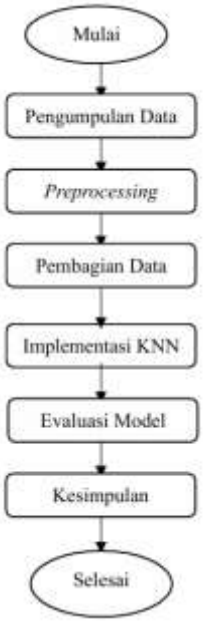
F1-Score sangat berguna untuk menyeimbangkan *Precision* dan *Recall* karena satu metrik saja tidak cukup untuk menggambarkan performa model secara menyeluruh. *F1-Score* membantu memastikan bahwa model tidak hanya berfokus pada peningkatan salah satu dari *Precision* atau *Recall*, tetapi mempertahankan keseimbangan

antara keduanya [21]. Untuk menghitung nilai *F1-Score* dapat menggunakan Persamaan 5.

$$F1\ Score\ = 2 \times \frac{Recall \times Precission}{Recall + Precission} \quad (5)$$

3. METODE PENELITIAN

Penelitian ini dilakukan dalam beberapa tahap utama, mulai dari pengumpulan data hingga analisis hasil model *K-Nearest Neighbor* (KNN). Berikut adalah alur penelitian secara sistematis dapat dilihat pada Gambar 1.



Gambar 1. *Flowchart*

Berikut adalah penjelasan langkah-langkah dalam *Flowchart* pada Gambar 1.

a. Mulai
   Penelitian dimulai dengan merancang metode dan tujuan penelitian.
b. Pengumpulan Data
   Data kecelakaan lalu lintas dikumpulkan dari sumber yang relevan, seperti Kepolisian Sektor Samarinda Kota. Data mencakup faktor-faktor seperti kondisi cahaya, tipe jalan, batas kecepatan, dan lainnya.
c. *Preprocessing*
   Data dibersihkan, diolah, dan disiapkan untuk analisis. Langkah-langkah meliputi menangani nilai yang hilang dan transformasi.
d. Pembagian Data
   Membagi data menjadi data *Training* dan data *Testing.*
e. Implementasi KNN
   Algoritme *K-Nearest Neighbor* (KNN) diterapkan untuk melakukan klasifikasi tingkat kecelakaan. Parameter K optimal ditentukan untuk mendapatkan hasil terbaik.
f. Evaluasi Model
   Model diuji menggunakan metrik evaluasi seperti akurasi, presisi, dan *recall* untuk mengukur performa KNN.
g. Kesimpulan
   Hasil penelitian dianalisis untuk menentukan efektivitas model dalam mengklasifikasikan tingkat kecelakaan.
h. selesai
   Penelitian selesai setelah menyusun laporan hasil dan memberikan rekomendasi bagi otoritas terkait.

3.1. Objek Penelitian

Kecelakaan lalu lintas merupakan penyebab utama kecacatan dan kematian di kalangan generasi muda, terutama laki-laki. Tingginya angka kematian akibat kecelakaan lalu lintas di kalangan generasi muda disebabkan oleh rendahnya kesadaran akan bahaya lalu lintas di jalan raya [22]. Bahaya lalu lintas dapat diklasifikasikan menurut waktu kecelakaan, tingkat keparahan kecelakaan, korban kecelakaan, cuaca pada saat kecelakaan, jenis kecelakaan dan lain lain [23].

Kecelakaan lalu lintas menjadi dampak negatif dari pesatnya peningkatan mobilitas lalu lintas jika tidak didukung oleh infrastruktur yang mengutamakan fitur keselamatan [24]. Adapun data dan keterangan dari kecelakaan lalu lintas yang akan digunakan pada penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Data Kecelakaan Lalu Lintas

Keterangan	**Spesifikasi**
Kondisi permukaan jalan	Berombak Baik Licin Berlubang Basah
Pencahayaan	Terang Gelap
Cuaca	Cerah Hujan Angin kencang Hujan disertai angin kencang Berawan

3.2. Data Penelitian

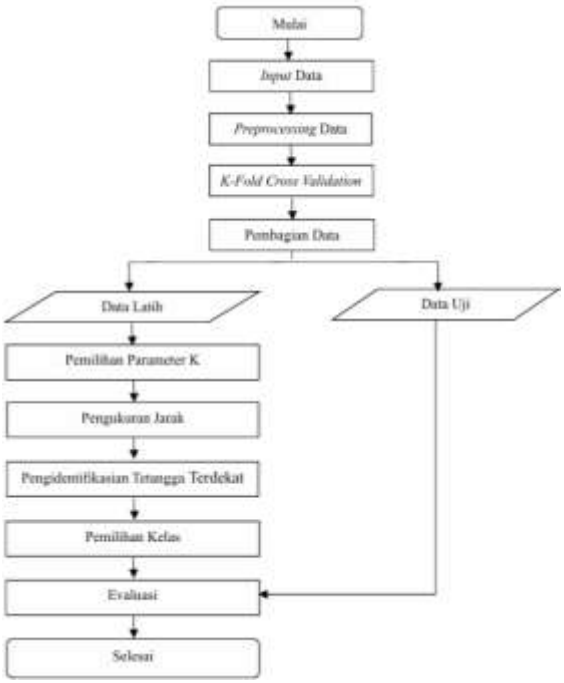
Data yang digunakan pada penelitian ini adalah data sekunder kecelakaan lalu lintas di Kota Samarinda, dimana data tersebut berjumlah 1000 pada tahun 2021-2024 yang didapat melalui Kepolisian sektor Samarinda Kota. Terdapat 3 kelas tingkat kecelakaan yang dapat di klasifikasikan seperti pada Tabel 1, yang Dimana pengumpulan data tersebut melalui wawancara langsung kepada Brigpol Basofi Zoharana Pratama sebagai operator aplikasi *Integrated Road Safety Management System* (IRSMS) di Satlantas Polresta Samarinda yang memiliki pengalaman 12 tahun bertugas di Kepolisian.

Wawancara terkait pengumpulan data dilakukan pada Polsek Samarinda Kota, dimana didapatkan hasil terkait informasi kecelakaan yang akan digunakan pada penelitian ini. Penelitian ini

menggunakan metode *K-Nearest Neighbor* dalam mengklasifikasikan tingkat kecelakaan dengan melakukan analisis terhadap faktor-faktor yang menjadi penyebab kecelakaan lalu lintas.

3.3. Alur Metode

Alur metode adalah proses yang terstruktur dan sistematis untuk memastikan bahwa data dikumpulkan, diproses, dianalisis, serta dimodelkan dengan cara yang efisien. Memahami dan mengikuti alur membantu dalam membangun model penelitian yang berkualitas serta yang bermanfaat dari data yang digunakan. Berikut Alur Metode yang dapat dilihat pada Gambar 2.



Gambar 2. Alur Metode

4. HASIL DAN PEMBAHASAN

4.1. *Preprocessing*

Pada tahap *Preprocessing* ini mencakup beberapa langkah yang dilakukan untuk mempersiapkan data yaitu mengubah kategori menjadi angka numerik di tunjukkan pada Tabel 2.

Tabel 2. *Preprocessing*

Atribut	Spesifikasi	Label
Batas kecepatan	20	0
	30	0
	40	0
	50	1
	60	1
	80	1
Kondisi jalan	Berombak	1
	Baik	0
	Licin	1
	Berlubang	1
	basah	1

Pada Tabel 2 menunjukan data yang telah di ubah menjadi numerik, dimana pada batas kecepatan akan bernilai 0 jika kecepatannya 20km-40km dan akan bernilai 1 jika kecepatannya 50km-80km. Pada kondisi jalan jika jalan tersebut berombak, licin, berlubang, dan basah akan bernilai 1 dan jika kondisi jalan tersebut baik maka bernilai 0.

4.2. *Index K-Fold*

Pengujian *K-Fold* dilakukan sebanyak 5 kali percobaan menggunakan K=25 yang dianggap sebagai K paling optimal berdasarkan hasil penelitian ini. Hasil dari pengujian ini dapat dilihat secara rinci pada Gambar 3.

	Fold	Index	Actual	Predicted
0	1	0	1	1
1	1	1	0	0
2	1	2	1	0
3	1	3	0	0
4	1	4	1	1
...	...	...	...	...
25095	25	999	1	0
25096	25	1000	0	1
25097	25	1001	1	1
25098	25	1002	1	1
25099	25	1003	1	1

Gambar 3. Pengujian *K-Fold* K=25

Pada Gambar 3 menentukan model akurasi dengan membandingkan kolom "*Actual*" dan "*Predicted*". Sebagai contoh, jika dua kolom memiliki nilai yang sama, maka prediksi tersebut akurat. Kesalahan Prediksi: Perbedaan antara nilai "*Actual*" dan "*Predicted*" menunjukkan hasil prediksi. Pada baris pertama, nilai *actual* adalah 1 dan nilai prediksi juga 1, yang menunjukkan bahwa prediksi tersebut akurat. Pada baris kedua, nilai *actual* nya adalah 1, tetapi nilai prediksinya adalah 0, yang mengindikasikan bahwa prediksinya signifikan.

4.3. Pembagian Data

Pada penelitian ini peneliti menggunakan data sebanyak 1004 data kemudian dibagi menjadi 90:10, yaitu 90% data latih dan 10% data uji yang dapat dilihat pada Gambar 4.

```
Jumlah data latih : 903
Jumlah data uji : 101
```

Gambar 4 Pembagian Data

Pada Gambar 4 menunjukan pembagian data dengan rasio 90:10 dengan hasil jumlah data latih sebanyak 903 dan data uji sebanyak 101.

4.4. Data Kelas, *Actual*, Klasifikasi

Pada tahap ini dihasilkan data kelas *actual*, dimana didapatkan hasil klasifikasi seperti yang tertera pada Gambar 5.

100	Sedang	Ringan
101	Ringan	Ringan
102	Ringan	Ringan
103	Ringan	Ringan
104	Ringan	Ringan
105	Berat	Berat
106	Ringan	Ringan
107	Sedang	Ringan
108	Ringan	Ringan
109	Ringan	Ringan
110	Ringan	Ringan

Gambar 5 Data kelas, *actual* dan klasifikasi

Pada Gambar 5 menunjukan kolom pertama berisi data uji, kolom kedua menunjukkan data *actual* atau data sebenarnya dan kolom ketiga menunjukkan hasil dari klasifikasi KNN.

4.5. Parameter Awal *K-Nearest Neighbor*

Parameter awal *K-Nearest Neighbor*, rasio pembagian data, dan *K-Fold* yang digunakan pada model ini didapattkan dari artikel terkait dengan penelitian ini, berikut merupakan penilaian parameter, rasio serta *K-Fold* yang dapat dilihat pada Tabel 3.

Tabel 3. Parameter Awal

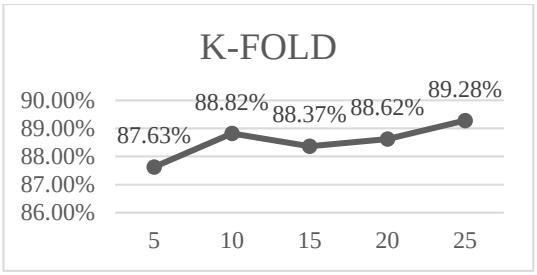
Split data	20:80		
K-fold	10		
Parameter	6		
Kelas	0	1	2
Precision	0,91	0,90	1,00
Recall	0,94	0,84	1,00
F1-score	0,93	0,87	1,00
Akurasi	90,55%		

Pada Tabel 3 menunjukkan nilai-nilai parameter awal yang digunakan dalam penelitian ini, dimana rasio pembagian data adalah 20% untuk data uji dan 80% untuk data latih. Pengujian ini dilakukan dengan menggunakan metode *K-Fold*=10 dan parameter K=6, serta didapatkan hasil akurasi sebesar 90,55%.

4.6. *K-Fold*

Pengujian ini dilakukan menggunakan K=5, 10, 15, 20, 25 dimana, setiap nilai K dilakukan pengujian sebanyak 5 kali untuk memastikan konsistensi hasil. Pengujian *K-Fold* ini dirancang untuk mengevaluasi kinerja model pada setiap nilai K yang dipilih, serta hasil pengujian *K-Fold* dapat dilihat pada Gambar 6.

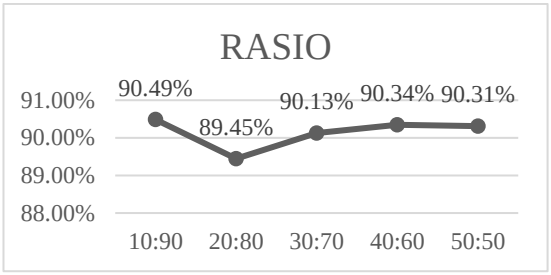
Pada Gambar 6 menunjukkan nilai akurasi *K-Fold* tertinggi yaitu pada K=25 menghasilkan akurasi sebesar 89,28%.



Gambar 6 K-*Fold*

4.7. Pengujian Rasio

Pada tahap pengujian ini dilakukan proses uji rasio pembagian data dengan persentase 10%, 20%, 30%, 40% dan 50% yang dapat dilihat pada Gambar 7.

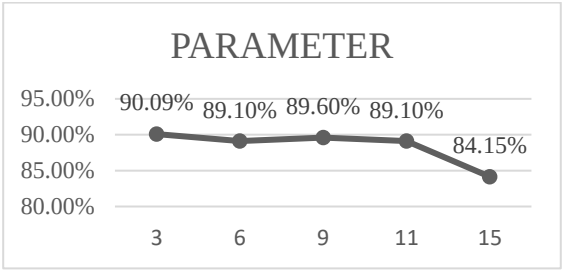


Gambar 7 Grafik Rasio

Pada Gambar 7 menunjukkan bahwa dari 5 kali perulangan menggunakan rasio yang berbeda-beda yaitu 10:90, 20:80, 30:70, 40:60 dan 50:50 menghasilkan akurasi tertinggi pada rasio 10% data uji serta 90% data latih dengan hasil 90,49%.

4.8. Pengujian Parameter K

Pada pengujian parameter K dilakukan eksperimen dengan berbagai nilai K, yaitu K=3, 6, 9, 11, 15 untuk menentukan nilai K yang paling optimal bagi model. Setiap nilai K diuji secara menyeluruh untuk mengevaluasi kinerja model dalam kondisi berbeda, dimana hasil pengujian parameter K dapat dilihat pada Gambar 8.



Gambar 8 Parameter

Pada Gambar 8 menunjukkan nilai dari 5 kali perulangan parameter yaitu K=3, K=6, K=9, K=11 dan K=15, dimana didapatkan parameter tertinggi yaitu pada parameter K=3 menghasilkan akurasi sebesar 90,09%.

4.9. Parameter Akhir *K-Nearest Neighbor*

Setelah pengujian metode *K-Nearest Neighbor, K-Fold*, Rasio dan Parameter didapatkan hasil tertinggi yang dapat dilihat pada Tabel 4.

Tabel 4 Parameter Akhir

Split data	10:90		
K-fold	25		
Parameter	3		
Kelas	0	1	2
Precision	0,97	0,87	1,00
Recall	0,90	0,96	1,00
F1-score	0,93	0,91	1,00
Akurasi	92,54%		

Pada Tabel 4 menunjukkan parameter akhir yang didapatkan dari pengujian menggunakan nilai rasio, *K-Fold* dan parameter yang tertinggi berdasarkan jurnal. Kemudian didapatkan hasil akurasi sebesar 92,54%.

5. KESIMPULAN DAN SARAN

Pada penelitian tingkat kecelakaan di kota Samarinda yang diklasifikasikan menggunakan metode *K-Nearest Neighbor* dapat disimpulkan sebagai berikut. Klasifikasi Tingkat Kecelakaan menggunakan metode *K-Nearest Neighbor* dengan rasio pembagian data 10:90, *K-Fold*=25 dan parameter 3 menghasilkan nilai *Precision* 97% *Recall* 90% *F1-Score* 93% pada kelas 1. Pada kelas 2 menghasilkan *Precision* 87%, *Recall* 96%, *F1-Score* 91%, serta kelas 2 menghasilkan *Precision* 100% *Recall* 100% *F1-Score* 100% dengan akurasi sebesar 92,54%, Dengan pemilihan parameter yang tepat, model KNN berhasil mencapai tingkat akurasi yang memuaskan, diharapkan hasil tersebut bermanfaat untuk mendukung pengambilan keputusan dalam upaya meningkatkan keselamatan jalan wilayah tersebut, Metode *K-Nearest Neighbor* (KNN) dapat diterapkan secara efektif untuk mengklasifikaskan tingkat kecelakaan lalu lintas berdasarkan 1004 data yang diperoleh dari Polsek Samarinda Kota. Pada parameter awal mendapatkan hasil akurasi sebesar 90,55% dan pada parameter akhir sebesar 92,54%.

Diharapkan untuk penelitian selanjutnya dapat memperluas cakupan dengan menggunakan dataset yang lebih besar dan beragam serta membandingkan metode klasifikasi lain seperti *Random Forest* atau *Support Vector Machine*. untuk meningkatkan kualitas generalisasi hasil. Selain itu, optimasi parameter dan penggunaan konsep lain seperti *Ensemble Learning*, dapat membantu meningkatkan akurasi serta performa model. Analisis mendalam tentang faktor penyebab kecelakaan serta eksplorasi faktor sosial dan ekonomi mempengaruhi keparahan kecelakaan juga penting untuk mendapatkan wawasan yang lebih komprehensif, seperti aplikasi pemantauan kecelakaan *Real Time* untuk meningkatkan manfaat praktis penelitian ini.

DAFTAR PUSTAKA

[1] 2022 Asep Nugraha et al., "Hlm 21-39 ANALISIS YURIDIS PENERAPAN RESTORATIF JUSTICE DALAM KECELAKAAN LALU LINTAS GOLONGAN BERAT YANG MENYEBABKAN ORANG LAIN MENINGGAL DUNIA MENURUT UU NOMOR 22 TAHUN 2009 DALAM PERSPEKTIF KEADILAN Kepolisian Resor Kota (Polresta) Cirebon , Indonesia," vol. 8, no. 2, pp. 21–39, 2022.

[2] A. J. Sihombing and H. Widyastuti, "Analisa Kecelakaan Lalu Lintas di Ruas Jalan Tol Cipularang, Purwakarta," *J. Tek. ITS*, vol. 9, no. 2, 2021, doi: 10.12962/j23373539.v9i2.57996.

[3] M. A. F. D. F. F. Alfath S.N. Syaban*1, "Karakteristik Keselamatan Lalu LintasDi Kota Manado," *Keselam. Transp. Jalan (Indonesian J. Road Safety)*, vol. 9, no. 2, pp. 103–109, 2022, doi: 10.46447/ktj.v9i2.421.

[4] L. Ratnawaty, "Upaya Pencegahan Terhadap Kecelakaan Lalu Lintas Di Kabupaten Bogor," *YUSTISI (Jurnal Huk. dan Huk. Islam.*, vol. 9, no. 2, pp. 1–10, 2022.

[5] I. Romli and A. T. Zy, "Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5," *J. Sains Komput. Inform. (J-SAKTI*, vol. 4, no. 2, pp. 694–702, 2020.

[6] R. Rahmadini, Enjel Erika LorencisLubis, Aji Priansyah, Yolanda R.W.N, and Tuti Meutia, "Penerapan Data Mining Untuk Memprediksi Harga Bahan Pangan Di Indonesia Menggunakan Algoritma K-Nearest Neighbor," *J. Mhs. Akunt. Samudra*, vol. 4, no. 4, pp. 223–235, 2023, doi: 10.33059/jmas.v4i4.7074.

[7] 2023 Nabila Abda Salsabila et al., "Classification of The Severity 0f Traffic Accident Victims In The City of Samarinda Uses The K-Nearest Neighbor and Naive Bayes Algorithms," *J. EKSPONENSIAL*, vol. 14, no. 2, pp. 99–106, 2023, [Online]. Available: http://jurnal.fmipa.unmul.ac.id/index.php/exponensial99

[8] A. Franseda, W. Kurniawan, S. Anggraeni, and W. Gata, "Integrasi Metode Decision Tree dan SMOTE untuk Klasifikasi Data Kecelakaan Lalu Lintas," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 3, p. 282, 2020, doi: 10.26418/justin.v8i3.40982.

[9] F. T. Admojo and Ahsanawati, "Klasifikasi Aroma Alkohol Menggunakan Metode KNN," *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 34–38, 2020, doi: 10.33096/ijodas.v1i2.12.

[10] A. Putri *et al.*, "Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi

Kelulusan Mahasiswa Tingkat Akhir," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 20–26, 2023, doi: 10.57152/malcom.v3i1.610.

[11] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.

[12] Y. Priantama and T. A. Yoga Siswa, "Optimasi Correlation-Based Feature Selection Untuk Perbaikan Akurasi Random Forest Classifier Dalam Prediksi Performa Akademik Mahasiswa," *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, p. 251, 2022, doi: 10.26798/jiko.v6i2.651.

[13] S. Febriani and H. Sulistiani, "Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5," *89Jurnal Teknol. dan Sist. Inf.*, vol. 2, no. 4, pp. 89–95, 2021.

[14] B. L. Fauzan, T. Agustin, and A. M. H. Mahmudah, "Prediksi Klasifikasi Kecelakaan Lalu Lintas di Kota Surakarta dengan Menggunakan Metode Regresi Logistik Multinomial," *Sustain. Civ. Build. Manag. Eng.*, vol. 1, no. 4, p. 9, 2024, doi: 10.47134/scbmej.v1i4.3159.

[15] Zulaiha, Mirtawati, and A. Saputra, "Klasifikasi Tingkat Keparahan Korban Pada Kecelakaan Lalu Lintas Dengan Metode Chi-Squared Automatic Interaction Detection (CHAID) Di Kota Bekasi," *Mat. Sains*, vol. 1, no. 2, pp. 16–26, 2023.

[16] B. S. Susanti, R. H. Hirzi, S. A. Wulandhya, and S. H. Hastuti, "Analisis Klasifikasi Kecelakaan Lalu Lintas Lombok Timur Berdasarkan Tingkat Keparahan Korban Kecelakaan Menggunakan Metode Support Vector Machine dan Bootstrap Aggregating," vol. 1, no. 1, pp. 1–8, 2024.

[17] A. Tangkelayuk, "The Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes, dan Decision Tree," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1109–1119, 2022, doi: 10.35957/jatisi.v9i2.2048.

[18] A. D. Adhi Putra, "Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962.

[19] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," pp. 37–63, 2020, [Online]. Available: http://arxiv.org/abs/2010.16061

[20] M. Grandini, E. Bagli, and G. Visani, "Metrics for Multi-Class Classification: an Overview," pp. 1–17, 2020, [Online]. Available: http://arxiv.org/abs/2008.05756

[21] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 273–281, 2024, doi: 10.57152/malcom.v4i1.1085.

[22] Khairul Fahmi, "Faktor Penyebab Kecelakaan Lalu Lintas Dan Perilaku Berkendara Pada Siswa Sekolah Menengah Atas Di Pasir Pengaraian Riau," *J. Ilm. Cano Ekon.*, vol. 10, no. 1, pp. 1–10, 2021, doi: 10.30606/cano.v10i1.1084.

[23] Z. Siregar and I. Dewi, "Analisis Ruas Jalan Lintas Sumatera Kota Tebing Tinggi Dan Kisaran Sebagai Titik Rawan Kecelakaan Lalu Lintas," *J. MESIL (Mesin Elektro Sipil)*, vol. 1, no. 2, pp. 63–73, 2020, doi: 10.53695/jm.v1i2.88.

[24] E. E. S. Putra, S. Y. Ratih, and L. Primantari, "Analisis Daerah Rawan Kecelakaan Lalu Lintas Jalan Raya Ngerong Cemorosewu," *J. Kacapuri J. Keilmuan Tek. Sipil*, vol. 4, no. 2, p. 255, 2022, doi: 10.31602/jk.v4i2.6432.