

The Influence Of Various Toll Accident Factors For The Wound Suffered By The Victim Using C4.5 Algorithm

Pengaruh Berbagai Faktor Kecelakaan Di Jalan Tol Terhadap Luka Yang Diderita Korban Menggunakan Algoritma C4.5

Andi Azhar Mustary Husein
{andiazhar55@gmail.com}

Program Studi Informatika Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo

Abstract. *The toll road is the main road that connects one city to another, used to speed up the journey of public and corporate vehicles. However, toll roads are no longer safe because of human negligence who does not obey existing regulations, resulting in victims. The C4.5 algorithm is one of the algorithms used to form a decision tree based on training data. This algorithm is a very powerful and well known classification and prediction method. The process of processing accident data in toll road management companies has not used data mining methods. So it is necessary to create a data mining method using the C4.5 algorithm. With this method, it is expected to be able to assist toll road managers in dealing with existing accidents based on the results of data processing using this method. The results of this study are the results of the prediction presentation, where using manual calculations the truth rate is 86% and the error rate is 14%, for calculations using the Weka application the truth is 79% and the error rate is 21%.*

Keywords – Classification; C4.5 algorithm; Data Mining; TollRoads; Victim's wound

Abstrak. *Di Jalan tol merupakan jalan utama yang menghubungkan satu kota dengan kota yang lain, dimana kegunaannya untuk mempercepat perjalanan kendaraan masyarakat maupun perusahaan. Tetapi jalan tol sudah tidak aman karena kelalaian manusia yang tidak mentaati peraturan yang ada sehingga menimbulkan adanya korban. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk decision tree berdasarkan training data. Algoritmaini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Proses pengolahan data kecelakaan diperusahaan pengelola jalan tol belum menggunakan metode data mining. Maka perlu dibuat suatu metode data mining menggunakan algoritma C4.5. Dengan metode tersebut diharapkan dapat mampu membantu pihak pengelola jalan tol dalam menangani kecelakaan yang ada berdasarkan hasil dari pengolahan data menggunakan metode tersebut. Hasil dari penelitian ini berupa hasil presentasi prediksi, dimana menggunakan perhitungan manual mendapatkan tingkat kebenaran sebesar 86% dan tingkat kesalahan diperoleh sebesar 14%, untuk perhitungan menggunakan aplikasi weka mendapatkan kebenaran sebesar 79% dan tingkat kesalahan diperoleh sebesar 21%.*

Kata Kunci – Klasifikasi; Algoritma C4.5; Data Mining; Jalan Tol; Luka Korban

I. PENDAHULUAN

Kecelakaan lalu lintas merupakan salah satu hal yang sangat sering terjadi di jalan tol, karena merupakan jalur bagi kendaraan yang berjalan dengan kecepatan tinggi. Jalan tol yang pertamanya berfungsi memperlancar perjalanan dan pengiriman barang dari satu tempat ke tempat lain, akhir-akhir ini menjadi sumber kecelakaan.

Jalan tol sudah tidak aman karena kelalaian manusia yang tidak mentaati peraturan yang ada sehingga menimbulkan adanya korban, yang terdiri dari korban luka ringan, luka berat, dan meninggal. Ada berbagai macam kemungkinan yang dapat menyebabkan kecelakaan, antara lain adalah kurang antisipasi, lengah, mengantuk, mabuk, tidak tertib (jarak rapat), ban pecah, rem blong dan masih banyak lagi.

Namun masih banyak orang yang menganggap ringan masalah ini, mereka menganggap hal kecil seperti mengecek ban yang dimana dapat menyebabkan pecah ban dan terjadinya kecelakaan itu tidak penting. Tidak hanya itu, ada beberapa yang perlu diperhatikan juga seperti kondisi mesin, kondisi kesehatan saat berkendara dan aturan lalu lintas yang ada. Ada berbagai macam kemungkinan yang dapat menyebabkan kecelakaan, antara lain adalah kurang antisipasi, lengah, mengantuk, mabuk, tidak tertib (jarak rapat), ban pecah, rem blong dan masih banyak lagi.

Menanggapi permasalahan diatas, peneliti akan melakukan penelitian menggunakan algoritma C4.5 yang dapat mengklasifikasi pengaruh berbagai faktor kecelakaan terhadap dampak luka yang akan diderita korban saat kecelakaan. Dari penelitian ini akan muncul dua kelas klasifikasi, yaitu luka ringan dan luka berat.

II. METODE

A. Kecelakaan lalu lintas

Kecelakaan lalu lintas merupakan berbagai bentuk dari tabrakan, slip maupun hilangnya kendali kendaraan yang terjadi pada ruas jalan, dimana mengakibatkan adanya korban, baik itu cedera atau kematian pada manusiamaupun kerusakan pada benda - benda yang ada di lokasi kecelakaan [1].

B. Data mining

Menurut Gartner Group, data mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika [2].

Data mining merupakan gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar [3].

Data mining menurut David Hand, Heikki Mannila, dan Padhraic Smyth dari MIT adalah analisis terhadap data (biasanya data yang berukuran besar) untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut [4].

Definisi sederhana dari data mining adalah ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di database yang besar. Dalam jurnal ilmiah, data mining juga dikenal dengan nama Knowledge Discovery in Databases (KDD) [5].

Knowledge Discovery In Databases (KDD) adalah sebuah proses non trivial yang dimana untuk mencari dan mengidentifikasi pola (pattern) dalam data. Knowledge Discovery In Database (KDD) merupakan proses yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau sebuah hubungan dalam data berukuran besar [6].

C. Algoritma C4.5

Algoritma C4.5 merupakan pengembangan dari Algoritma ID3 yang ditemukan oleh Ross Quinlan, Algoritma C4.5 digunakan untuk mengklasifikasikan data yang memiliki atribut berupa numerik ataupun kategorial, Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskrit dari record yang baru [7]

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk decision tree berdasarkan training data. Algoritma ini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal [8]. Algoritma C4.5 dan *decision tree* merupakan dua model yang tidak terpisahkan. Karena untuk membangun sebuah decision tree, dibutuhkan algoritma C4.5 [9].

D. Weka

WEKA adalah sebuah paket tools machine learning praktis. WEKA merupakan singkatan dari “Waikato Environment for Knowledge Analysis”, yang dibuat di Universitas Waikato, New Zealand untuk penelitian, pendidikan dan berbagai aplikasi [10]

E. Metode penelitian

Metode yang dipakai dalam penelitian ini adalah metode eksperimental atau penelitian terapan. Dimana menerapkan algoritma C4.5 ke dalam sistem serta menganalisis apa yang menyebabkan adanya korban sesuai kriteria yang ada di PT. Jasa Marga Cabang Surabaya, dimana data kecelakaan lalu lintas bulan Januari sampai Agustus Tahun 2018 dijadikan sebagai data *Training* dan Agustus sampai dengan Desember Tahun 2018 dijadikan sebagai *Sampel* data.

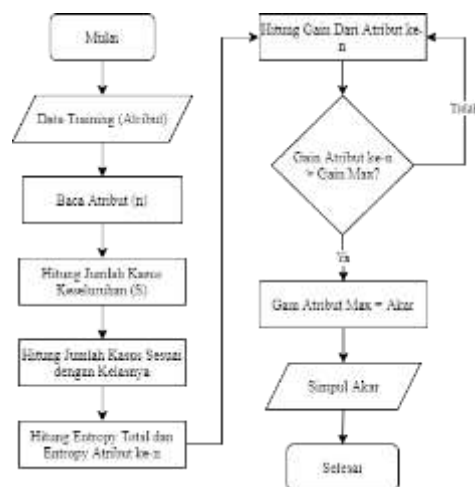
Kemudian dari data yang digunakan akan dilakukan proses *Preprocessing* yang selanjutnya masuk ke proses perhitungan algoritma C4.5, kemudian melakukan uji akurasi dari perhitungan yang dilakukan..



Gambar 1. Flowchart Penentuan Simpul Akar

Langkah Pertama yang dilakukan adalah memasukkan data training yang telah melalui *preprocessing*. Lalu sistem akan membacanya dan melakukan perhitungan untuk menentukan atribut yang akan digunakan sebagai simpul akar.

Flowchart untuk menentukan simpul akar dijabarkan pada gambar Flowchart penentuan simpul akar dibawah ini :



Gambar 2. Flowchart Penentuan Simpul Akar

Hal pertama yang dilakukan yaitu menghitung jumlah kasus secara keseluruhan, lalu dilanjutkan dengan menghitung jumlah kasus sesuai dengan kelasnya. Berikut langkah – langkahnya :

1. Mempersiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan dipilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus:

$$Entropy(S) = \sum_{i=1}^n - P_i \times \log_2 \times P_i$$

Di mana:

S = himpunan kasus

N = jumlah partisi S

P_i = proporsi S_i terhadap S

3. Menghitung nilai Gain menggunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} \times Entropy(S_i)$$

Di mana:

S = himpunan kasus

A = fitur

N = jumlah partisi atribut A

|S_i| = proporsi S_i terhadap S

|S| = jumlah kasus dalam S

4. Ulangi langkah ke-2 dan langkah ke-3 hingga semua record terpartisi.

5. Proses partisi decision tree akan berhenti saat:

6. Semua record dalam simpul N mendapat kelas yang sama.

7. Tidak ada atribut di dalam record yang dipartisi lagi.

8. Tidak ada record di dalam cabang yang kosong

III. HASIL DAN PEMBAHASAN

A. Pengujian

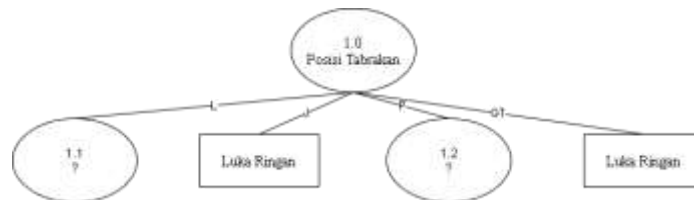
Hasil dari penelitian yang dilakukan di ruas tol Surabaya – Gempol PT. Jasa Marga. Setelah dilakukan pengambilan data, peneliti langsung melakukan uji coba. Dalam uji penelitian ini, menggunakan metode klasifikasi dengan algoritma C4.5

Tabel 1. Perhitungan Awal Untuk Pencarian Simpul Akar Menggunakan Excel

Node Akar		Jumlah	Luka Ringan	Luka Berat	Entropy	Gain
Total		174	104	70	0,972	
Jumlah Kendaraan Terlibat	Satu Kendaraan	87	49	38	0,988	0,004
	Dua Kendaraan	65	41	24	0,950	
	Tiga Kendaraan atau Lebih	22	14	8	0,946	
Kondisi Kendaraan	Ringan	60	31	29	0,999	0,010
	Berat	114	73	41	0,942	
Sebab Kecelakaan	Human Error	146	81	65	0,991	0,032
	Kondisi Kendaraan	28	23	5	0,677	
Hari Kejadian	Hari Kerja	119	76	43	0,944	0,015
	Akhir Minggu	26	15	11	0,983	
	Libur/Minggu	29	13	16	0,992	
Jam Kejadian	Waktu 1	58	34	24	0,978	0,006
	Waktu 2	45	30	15	0,918	
	Waktu 3	41	24	17	0,979	
	Waktu 4	30	16	14	0,997	

						0,000
Kondisi Cuaca	Cuaca Panas	148	89	59	0,970	
	Polusi	5	3	2	0,971	
	Cuaca Lembab	21	12	9	0,9852	
						0,033
Posisi Tabrakan	Lajur	149		61	0,976	
	Jembatan	2	2	0	0,000	
	Pembatas	18	9	9	1,000	
	Gerbang Tol	5	5	0	0,000	
Jenis Kecelakaan	Tabrakan Tunggal	74		28	0,957	
	Menabrak Objek	10	5	5	1,000	
	Tabrakan Dua Kendaraan	71	40	31	0,988	
		19	13	6	0,900	
Jenis Kendaraan	Mobil	88	51	37	0,982	
	Bus	13	6	7	0,996	
	Truk 1	23	16	7	0,887	
	Truk 2	50	31	19	0,958	
						0,013
Jenis Kelamin	Laki - Laki	171	101	70	0,976	
	Perempuan	3	3	0	0,000	

Dapat dilihat gain terbesar dari seluruh atribut merupakan atribut posisi tabrakan dengan nilai 0,033,selanjutnya atribut posisi tabrakan dijadikan sebagai simpul akar. Berikut gambar pohon keputusan:



Gambar 3. Pohon Simpul Akar yang didapat

Setelah dilakukannya perhitungan dan didapatkan simpul akar, selanjutnya dilakukan perhitungan lagi sampai nilai atribut berada ditempat yang sama.

B. Hasil

Pohon keputusan

Setelah dilakukan perhitungan berulang kali, didapatkan hasil sebagai berikut:



Gambar 4. Hasil Akhir Pohon

Hasil perbandingan

Dari hasil pohon keputusan yang didapat, diketahui basis aturan yang terbuat dari pohon keputusan. Kemudian dilanjutkan dengan proses perbandingan dengan data testing, berikut hasilnya:

Tabel 2. Hasil Perbandingan

Jenis Kendaraan Terlibat	Kondisi Kendaraan	Jumlah Kendaraan Terlibat	Sebab Kecelakaan	Hari Kejadian	Jam Kejadian
Mobil	Berat	Satu Kendaraan	Human Error	Hari Kerja	Waktu 3
Bus	Berat	Satu Kendaraan	Human Error	Hari Kerja	Waktu 2
Truk 1	Ringan	Satu Kendaraan	Human Error	Hari Kerja	Waktu 2
Truk 1	Berat	Dua Kendaraan	Human Error	Hari Kerja	Waktu 3
Mobil	Berat	Satu Kendaraan	Kondisi Kendaraan	Hari Kerja	Waktu 4
Bus	Berat	Satu Kendaraan	Kondisi Kendaraan	Libur/Minggu	Waktu 2
Mobil	Ringan	Satu Kendaraan	Human Error	Libur/Minggu	Waktu 2
Bus	Berat	Tiga Kendaraan atau Lebih	Human Error	Hari Kerja	Waktu 2
Truk 2	Berat	Dua Kendaraan	Human Error	Hari Kerja	Waktu 1
Mobil	Berat	Tiga Kendaraan atau Lebih	Human Error	Hari Kerja	Waktu 2
Truk 2	Ringan	Satu Kendaraan	Human Error	Hari Kerja	Waktu 1
Truk 2	Berat	Satu Kendaraan	Kondisi Kendaraan	Hari Kerja	Waktu 4
Mobil	Berat	Tiga Kendaraan atau Lebih	Human Error	Hari Kerja	Waktu 2
Truk 2	Berat	Dua Kendaraan	Human Error	Hari Kerja	Waktu 4
Mobil	Berat	Tiga Kendaraan atau Lebih	Human Error	Hari Kerja	Waktu 3
Mobil	Berat	Dua Kendaraan	Human Error	Hari Kerja	Waktu 2

Lanjutan Tabel 1.

Kondisi Cuaca	Posisi Tabrakan	Jenis Kecelakaan	Jenis kelamin	Luka Korban	Prediksi	Hasil
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Dua Kendaraan	Laki - Laki	Luka Ringan	Luka Berat	Tidak Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Berat	Luka Ringan	Tidak Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Perempuan	Luka Berat	Luka Ringan	Tidak Benar
Polusi	Lajur	Tabrakan Beruntun	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Dua Kendaraan	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan	Laki -	Luka	Luka Ringan	Benar

		Beruntun	Laki	Ringan		
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Berat	Luka Berat	Benar
Cuaca Panas	Lajur	Tabrakan Tunggal	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Beruntun	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Dua Kendaraan	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Beruntun	Laki - Laki	Luka Ringan	Luka Ringan	Benar
Cuaca Panas	Lajur	Tabrakan Dua Kendaraan	Laki - Laki	Luka Ringan	Luka Ringan	Benar

Tabel 3. Hasil Persentasi

Hasil Prediksi	Jumlah	Persentase %
Benar	37	86%
Tidak Benar	6	14%

Hasil perhitungan menggunakan weka

Selanjutnya peneliti melakukan proses algoritma c4.5 menggunakan aplikasi weka, berikut hasilnya:

```

Classifier output

--- Evaluation on test set ---

Time taken to test model on supplied test set: 0 seconds

--- Summary ---

Correctly Classified Instances      34           79.0698 %
Incorrectly Classified Instances    9            20.9302 %
Kappa statistic                     0.5666
Mean absolute error                  0.3066
Root mean squared error              0.366
Relative absolute error              63.7501 %
Root relative squared error         74.1594 %
Total Number of Instances           43

--- Detailed Accuracy By Class ---

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	ROC Area	PRC Area	Class
	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
Weighted Avg.	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat

```

--- Confusion Matrix ---

 a  b  <-- classified as
21  4  | a = Luka Ringan
 5 13 | b = Luka Berat

```

Gambar 5. Hasil Perbandingan Menggunakan Weka

Sample Data Testing_perbandingan

Model	Hasil	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	ROC Area	PRC Area	Class
S1	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S2	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S3	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S4	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S5	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S6	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S7	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S8	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S9	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S10	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S11	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S12	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S13	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S14	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S15	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S16	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S17	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S18	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S19	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S20	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S21	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S22	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S23	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S24	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S25	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S26	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S27	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S28	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat
S29	Benar	0,840	0,278	0,808	0,840	0,824	0,567	0,878	0,875	Luka Ringan
S30	Salah	0,723	0,169	0,765	0,723	0,743	0,567	0,878	0,856	Luka Berat

Gambar 6. Tampilan Perbandingan Pada Aplikasi Weka

VII. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, persentase dari hasil prediksi data kecelakaan menggunakan 80% dari data training sebanyak 174 data dan 20% dari data testing sebanyak 43 data kecelakaan. Didapatkan hasil prediksi menggunakan perhitungan manual tingkat kebenaran sebesar 86% dan tingkat kesalahan diperoleh sebesar 14%, untuk perhitungan menggunakan aplikasi weka mendapatkan kebenaran sebesar 79% dan tingkat kesalahan diperoleh sebesar 21%. Hasil perhitungan manual lebih tepat dan tingkat kebenaran tinggi dari pada aplikasi weka, karena persentase dari hasil prediksi lebih besar menggunakan perhitungan manual dari pada aplikasi weka.

REFERENSI

- [1] A. Dharma, B. M. Edison, and Rismalinda, "IDENTIFIKASI KECELAKAAN LALU LINTAS (Study Kasus Jalan Dalu-Dalu sampai Pasir Pengaraian)."
- [2] D. Wira and T. Putra, "ALGORITMA C4.5 UNTUK MENENTUKAN TINGKAT KELAYAKAN MOTOR BEKAS YANG AKAN DIJUAL," vol. 4, no. 1, 2016.
- [3] D. T. Larose, "Discovering Knowledge in Data: An Introduction to Data Mining," 2005.
- [4] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5".
- [5] H. Susanto, "DATA MINING UNTUK MEMREDIKSI PRESTASI SISWA BERDASARKAN SOSIAL EKONOMI, MOTIVASI, KEDISIPLINAN DAN PRESTASI MASA LALU DATA MINING TO PREDICT STUDENT'S ACHIEVEMENT BASED ON SOCIO-ECONOMIC, MOTIVATION, DISCIPLINE AND ACHIEVEMENT OF THE PAST," 2014.
- [6] E. Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti," vol. 2, no. 1, 2017.
- [7] M. Muhamad, A. P. Windarto, and S. Suhada, "PENERAPAN ALGORITMA C4.5 PADA KLASIFIKASI POTENSI SISWA DROP OUT," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 3, no. 1, Dec. 2019, doi: 10.30865/komik.v3i1.1688.
- [8] Rismayanti, "IMPLEMENTASI ALGORITMA C4.5 UNTUK MENENTUKAN PENERIMA BEASISWA DI STT HARAPAN MEDAN," pp. 1–5, 2016.
- [9] A. Andriani, "SISTEM PREDIKSI PENYAKIT DIABETES BERBASIS DECISION TREE," 2013.
- [10] S. Purnama, J. Henharta, Y. Perdana S, F. Ihsani, and I. W. Simri W, "Get Easy Using WEKA," Mar. 2013.