

---

## Data Mining Untuk Mengestimasi Tingkat Kecelakaan Lalu Lintas Di Kab. Serdang Bedagai Menggunakan Algoritma Regresi Linier Berganda

M. Harmizi\*,Azanuddin,\*\* Elfitriani.\*\*

\*Program Studi Sistem Informasi, STMIK Triguna Dharma

\*\*Program Studi Sistem Informasi, STMIK Triguna Dharma

---

### Article Info

### ABSTRACT

#### Article history:

Received Jun 12<sup>th</sup>, 201x

Revised Aug 20<sup>th</sup>, 201x

Accepted Aug 26<sup>th</sup>, 201x

---

#### Keyword:

Estimasi, Kecelakaan Lalu Lintas, *Data Mining*, Regresi Linier Berganda

Kepadatan kendaraan yang terus meningkat jumlahnya setiap tahunnya menyebabkan permasalahan lalu lintas diantaranya sering terjadinya kecelakaan. Timbulnya kecelakaan lalu lintas di jalan raya yang semakin tinggi diawali dengan perilaku pengendara yang melanggar aturan perundang-undangan lalu lintas yang ada.

Untuk meninjau nilai keselamatan perlu dilakukan pemahaman terhadap karakteristik kecelakaan seperti dengan melakukan estimasi atau prediksi mengenai data kecelakaan lalu lintas. Dalam perkembangan dunia komputer saat ini, ada beberapa teknik yang dapat digunakan untuk mengestimasi kecelakaan lalu lintas berbasis komputerisasi diantaranya dengan regresi linier berganda.

Dengan adanya sistem ini, diharapkan dapat membantu Polres Serdang Bedagai dalam mengestimasi tingkat kecelakaan lalu lintas di Kabupaten Serdang Bedagai. Sehingga diketahui faktor atau penyebab apa saja yang menjadi penyebab kecelakaan, sehingga. Hal ini tentunya untuk meminimalisir atau mengurangi angka kecelakaan lalu lintas di Kabupaten Serdang Bedagai.

Copyright © 201x STMIK Triguna Dharma.  
All rights reserved.

---

Nama : M.Harmizi  
Kampus : STMIK Triguna Dharma  
Program Studi : Sistem Informasi  
E-Mail : [mharmizi98@gmail.com](mailto:mharmizi98@gmail.com)

## 1. PENDAHULUAN

Di Kabupaten Serdang Bedagai, peningkatan transportasi sangatlah pesat, terutama tersedianya kendaraan bermotor dan jalan raya merupakan tujuan utama untuk memperlancar arus lalu lintas. Kepadatan kendaraan yang terus meningkat setiap tahunnya menyebabkan permasalahan lalu lintas diantaranya sering terjadinya kecelakaan. Timbulnya kecelakaan lalu lintas di jalan raya yang semakin tinggi sebagian besar juga diakibatkan atau diawali dengan oleh perilaku pengendara yang melanggar aturan perundang-undangan lalu lintas yang ada, seperti mengemudikan kendaraan dengan kecepatan tinggi atau tidak dengan hati-hati, mengendarai kendaraan bermotor tidak memiliki surat izin mengemudi, melanggar rambu-rambu lalu lintas dan marka jalan dan berbagai bentuk pelanggaran lainnya[1].

kecelakaan adalah tabrakan, *overtuning* atau slip yang terjadi di jalan terbuka pada lalu lintas umum yang menyebabkan luka, kematian/fatal, kerusakan pada kendaraan atau kerugian material[2]. Kriteria korban kecelakaan lalu lintas menurut pasal 93 pada Peraturan Pemerintah No. 43 tahun 1993 tentang prasarana dan lalu lintas jalan yaitu: korban meninggal, korban luka berat, korban luka ringan. Penyebab kecelakaan antara lain adalah: manusia, sarana dan prasarana (kendaraan dan jalan), alam atau lingkungan[2].

Dalam *Data Mining* terdapat beberapa jenis metode sesuai dengan pemanfaatannya diantaranya: prediksi, asosiasi, klasifikasi, estimasi dan *Clustering*. Pada metode estimasi terdapat beberapa teknik yang dapat digunakan diantaranya adalah algoritma *Regresi Linier Berganda*. Metode regresi adalah metode yang digunakan untuk menyatakan pola hubungan antara satu variabel respon dan satu atau lebih variabel prediktor[3]. *Regresi Linier Berganda* merupakan suatu metode yang digunakan untuk menganalisis hubungan antara variabel. Hubungan tersebut dapat diekspresikan dalam bentuk persamaan yang menghubungkan variabel terkait  $X$  dengan variabel bebas  $Y$ .

## 2. Kajian Pustaka

### 2.1 Lalu Lintas

Lalu lintas adalah kegiatan kendaraan bermotor dengan menggunakan jalan raya sebagai jalur lintas umum sehari-hari. Lalu lintas identik dengan jalur kendaraan bermotor yang ramai yang menjadi jalur kebutuhan masyarakat umum. Oleh karena itu lalu lintas selalu identik pula dengan penerapan tata tertib bermotor dalam menggunakan jalan raya. Tujuan berlalu lintas yang paling utama adalah menumbuhkan sikap dan perilaku untuk menumbuhkan etika dan budaya lalu lintas yang baik. Dalam masyarakat Indonesia beragam budaya dan kebiasaan menjadi hal yang menarik untuk digali[4].

### 2.2 Data Mining

*Data mining* merupakan metode pengolahan data berskala besar oleh karena itu *data mining* ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variabel, dan market basket analisis [5].

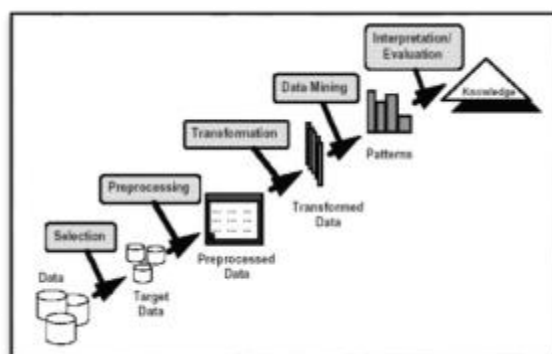
#### 2.2.1 Data Mining Menurut Para Ahli

*Data Mining* adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya [6].

*Data Mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data [6].

### 2.2.2 Knowledge Discovery in Database (KDD)

Pada proses *data mining* yang biasa disebut *Knowledge Discovery in Database* (KDD) terdapat beberapa proses seperti terlihat pada gambar di bawah ini



Gambar 1 Proses KDD

### 2.2.3 Konsep Data Mining

*Data Mining* adalah kegiatan mengekstraksi atau menambang pengetahuan dari data yang berukuran / berjumlah besar, informasi inilah yang nantinya sangat berguna untuk pengembangan. Dimana langkah-langkah untuk melakukan *data mining* adalah sebagai berikut:

1. *Data Selection*  
Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *KDD* dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.
2. *Pre-processing/cleaning*  
Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *KDD*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *KDD*, seperti data atau informasi eksternal.
3. *Transformation*  
*Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *KDD* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.
4. *Data mining*  
*Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode dan algoritma yang tepat sangat bergantung pada tujuan dan proses *KDD* secara keseluruhan.
5. *Interpretation/evaluation*  
Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses *KDD* yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

### 2.2.4 Proses Tahapan Data Mining

*Data mining* merupakan salah satu dari rangkaian *Knowledge Discovery in Database* (KDD). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola

sejumlah data. Serangkaian proses tersebut memiliki tahap sebagai berikut (Tan, 2004) dalam[7]:

1. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di mining)
4. Aplikasi teknik *data mining*, proses ekstraksi pola dari data yang ada
5. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung mengambil keputusan)
6. Prestasi pengetahuan (dengan teknik visualisasi), Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna.

### 2.2.5 Manfaat Data Mining

Pemanfaatan *data mining* dilihat dari dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan.

1. Dari sudut pandang komersial, pemanfaatan *data mining* dapat digunakan untuk menangani meledaknya volume data, dengan menggunakan teknik komputasi dapat digunakan untuk menghasilkan informasi-informasi yang dibutuhkan yang merupakan asset yang dapat meningkatkan daya saing suatu institusi.

Contohnya:

- a. Bagaimana mengetahui hilangnya pelanggan karena pesaing
  - b. Bagaimana mengetahui item produk atau konsumen yang memiliki kesamaan karakteristik
  - c. Bagaimana mengidentifikasi produk-produk yang terjual bersamaan dengan produk lain
  - d. Bagaimana memprediksi tingkat penjualan
  - e. Bagaimana menilai tingkat resiko dalam menentukan jumlah produksi suatu item
  - f. Bagaimana memprediksi perilaku bisnis dimasa yang akan datang
2. Dari sudut pandang keilmuan, *data mining* dapat digunakan untuk *capture*, menganalisis serta menyimpan data yang bersifat *real time* sangat besar, misalnya:
    - a. *Remote sensor* yang ditempatkan pada suatu satelit
    - b. *Telescope* yang digunakan untuk memindai langit
    - c. Simulasi saintifik yang membangkitkan data dalam ukuran *terabytes*

### 2.2.6 Fungsi Data Mining

Fungsi-fungsi yang umum diterapkan dalam *data mining* [7], yaitu:

1. *Association*, adalah proses untuk menemukan aturan asosiasi antara suatu kombinasi item dalam suatu waktu
2. *Sequence*, proses untuk menemukan aturan asosiasi antara suatu kombinasi item dalam suatu waktu dan diterapkan lebih dari satu periode
3. *Clustering*, adalah proses pengelompokan sejumlah data/obyek ke dalam kelompok data sehingga setiap kelompok berisi data yang mirip
4. *Classification*, proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang levelnya tidak diketahui.
5. *Regression*, adalah proses pemetaan data dalam suatu nilai prediksi
6. *Forecasting*, adalah proses pengestimasian nilai prediksi berdasarkan pola-pola di dalam sekumpulan data.
7. *Solution*, adalah proses penemuan akar masalah dan problem solving dari persoalan bisnis yang dihadapi atau paling tidak sebagai informasi dalam pengambilan keputusan.

### 2.2.7 Kelebihan dan Kekurangan Data Mining

*Data Mining* adalah proses menganalisa data yang berjumlah besar untuk menemukan suatu pola yang akan digunakan untuk pengambilan keputusan[8]. Sebagai alat analisis, berikut ini adalah kelebihan

dan kekurangan *data mining*, yaitu:

1. Kelebihan *Data Mining*
  - a. *Data mining* mampu menangani data dalam jumlah besar dan kompleks.
  - b. *Data mining* dapat menangani data dengan berbagai macam tipe atribut.
  - c. *Data mining* mampu mencari dan mengolah data secara semi-otomatis. Disebut semiotomatis karena dalam beberapa teknik *data mining*, diperlukan parameter yang harus di-input oleh user secara manual.
  - d. *Data mining* dapat menggunakan pengalaman ataupun kesalahan terdahulu untuk meningkatkan kualitas dan hasil analisa sehingga mendapat hasil yang terbaik
2. Kekurangan *Data Mining*
  - a. Terkadang ada data yang lebih mudah diselesaikan dengan statistik manual dibandingkan dengan menggunakan teknik *data mining*.
  - b. *Data mining* tidak dapat menemukan pengetahuan secara instant.
  - c. Algoritma *data mining* cukup kompleks.
  - d. Hasil dari *data mining* tidak dapat langsung digunakan, harus dianalisa dan diinterpretasi terlebih dahulu.

### 2.3 Estimasi

Menurut Kamus Besar Bahasa Indonesia, estimasi bisa diartikan sebagai perkiraan, penilaian, atau pendapat. Sementara, pendapat lainnya mengatakan bahwa estimasi adalah keseluruhan proses yang memerlukan serta menggunakan estimator untuk menghasilkan sebuah *estimate* dari suatu parameter. Seorang ahli juga menuturkan bahwa estimasi adalah suatu pengukuran yang didasarkan pada hasil kuantitatif atau dengan kata lain, tingkat akurasi bisa diukur dengan angka.

#### 2.3.1 Pengertian Estimasi

Estimasi adalah suatu metode untuk mengetahui sekitar nilai-nilai suatu populasi dengan menggunakan nilai-nilai sampel. Nilai-nilai populasi sering disebut dengan parameter populasi, sedangkan nilai-nilai sampel sering disebut dengan statistik sampel. Dalam metode simulasi, parameter populasi yang ingin diestimasi itu adalah berupa nilai-nilai rata-rata yang diberi notasi  $\mu$  dan nilai simpangan baku dengan notasi  $\sigma$ . Dengan menggunakan data sampel maka berusaha untuk mengetahui karakteristik populasi[9].

#### 2.3.2 Algoritma Regresi Linier

Regresi linier merupakan suatu persamaan yang menggambarkan hubungan antara variabel terikat dengan variabel bebas, dimana model berhubungan secara linier dengan variabel terikat. Selanjutnya model ini dapat digunakan untuk memprediksi nilai variabel terikat apabila diberikan nilai dari variabel bebas. Oleh karena itu, estimasi model yang didapatkan sebaiknya memenuhi kriteria model yang baik sehingga mampu digunakan prediksi error yang terkecil [9].

#### 2.2.3 Regresi Linier Berganda

*Regresi linier* berganda adalah regresi yang variabel terikat  $Y$  dihubungkan dengan lebih dari satu variabel bebas  $X$ [9]. Pendapat lain dijelaskan dari Amrin[10], yaitu analisis yang memiliki variabel bebas lebih dari satu disebut regresi linier berganda. Teknik regresi linier berganda digunakan untuk mengetahui ada tidaknya pengaruh signifikan dua atau lebih variabel bebas ( $X_1, X_2, X_3, \dots, X_k$ ) terhadap variabel terikat ( $Y$ ).

Pendapat lain mengatakan *regresi linier* berganda merupakan perluasan dari *regresi linier* sederhana. Dengan memperluas *regresi linier* dua atau tiga variabel, maka *regresi* dengan variabel terikat  $Y$  dan  $k$  variabel bebas  $X_1, X_2, X_3, \dots, X_k$  dapat dituliskan sebagai berikut:

1. Menggali data yang akan di proses.
2. Menentukan nilai variabel untuk setiap kriteria dari data yang telah digali.
3. Melakukan perhitungan terhadap data yang telah didapat nilai variabelnya dengan persamaan sebagai berikut:

$$Y = \beta_0 + X_1 \beta_1 + X_2 \beta_2 + X_3 \beta_3$$

Untuk menghitung menggunakan persamaan tersebut, maka terlebih dahulu dilakukan turunan. Yaitu pada langkah ke 4.

4. Dari persamaan di atas, kemudian dilakukan turunan yaitu sebagai berikut:

$$\sum Y = \beta_0 \cdot n + \beta_1 \cdot \sum X_1 + \beta_2 \cdot \sum X_2 + \beta_3 \cdot \sum X_3$$

$$\sum Y = \beta_0 \cdot \sum X_1 + \beta_1 \cdot \sum X_1 \cdot X_1 + \beta_2 \cdot \sum X_1 \cdot X_2 + \beta_3 \cdot \sum X_1 \cdot X_3$$

$$\sum Y = \beta_0 \cdot \sum X_2 + \beta_1 \cdot \sum X_2 \cdot X_1 + \beta_2 \cdot \sum X_2 \cdot X_2 + \beta_3 \cdot \sum X_2 \cdot X_3$$

$$\sum Y = \beta_0 \cdot \sum X_3 + \beta_1 \cdot \sum X_3 \cdot X_1 + \beta_2 \cdot \sum X_3 \cdot X_2 + \beta_3 \cdot \sum X_3 \cdot X_3$$

5. Melakukan eliminasi untuk setiap turunan, yaitu sebagai berikut:

$$\sum Y = \beta_0 \cdot n + \beta_1 \cdot \sum X_1 + \beta_2 \cdot \sum X_2 + \beta_3 \cdot \sum X_3$$

$$\sum X_1 Y = \beta_0 \cdot \sum X_1 + \beta_1 \cdot \sum X_1 \cdot X_1 + \beta_2 \cdot \sum X_1 \cdot X_2 + \beta_3 \cdot \sum X_1 \cdot X_3$$

Proses eliminasi dilakukan secara bertahap sehingga menghasilkan turunan yang baru.

6. Hasil akhir

Hasil akhir merupakan hasil yang didapat dari proses eliminasi, hasil akhir inilah yang merupakan jumlah penjualan mobil secara kredit untuk masa yang akan datang. Untuk mendapatkan hasil akhir, maka digunakan persamaan pada langkah ke 3.

### 3. METODE PENELITIAN

#### 3.1 Metode Perancangan Sistem

Di dalam penelitian ini, diadopsi sebuah metode perancangan sistem yaitu *waterfall algorithm*. Berikut ini adalah fase yang dilakukan dalam penelitian ini yaitu:

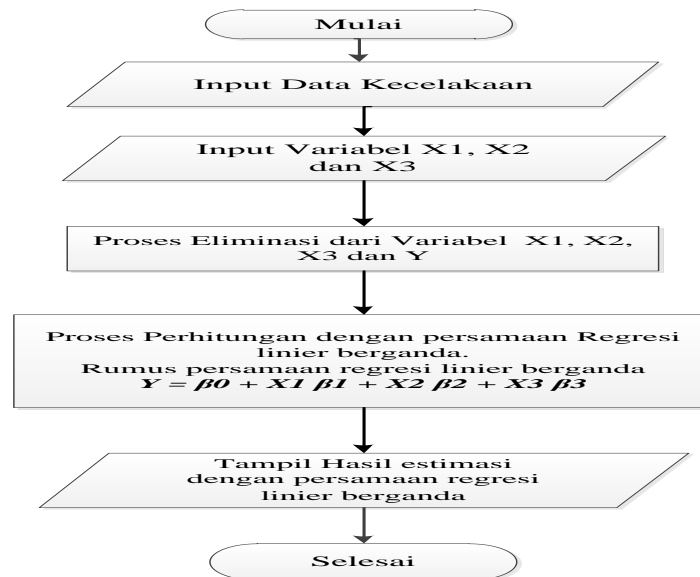
1. Analisis Masalah dan Kebutuhan
2. Desain Sistem
3. Pembangunan Sistem
4. Uji Coba Sistem
5. Implementasi atau Pemeliharaan

#### 3.2 Algoritma Sistem

Algoritma sistem merupakan suatu tahapan yang dilakukan sebelum melakukan proses mengestimasi tingkat kecelakaan lalu lintas.

### 1. Flowchart Metode Penyelesaian

Dibawah ini merupakan *flowchart* metode *regresi linier berganda* adalah sebagai berikut :



Gambar 2 Flowcart Metode

### 2. Menentukan Nilai Variabel untuk setiap Kriteria dari Data yang Ada

Untuk variabel yaitu faktor manusia (X1), faktor kendaraan (X2), faktor alam atau lingkungan (X3) dan jumlah kecelakaan (Y). Dari kesimpulan tersebut terdapat 3 variabel bebas dan 1 variabel terikat. Selanjutnya yaitu menentukan indikator penilaian untuk masing-masing variabel, yaitu sebagai berikut:

#### a. Faktor Manusia

Faktor manusia yang dicatat oleh kepolisian, meliputi jenis kelamin korban, usia korban, profesi korban, dan peran korban dalam berkendara.

#### b. Faktor Kendaraan

Faktor kendaraan yang paling sering terjadi adalah ban pecah, rem tidak berfungsi sebagaimana seharusnya, mesin yang tiba-tiba mati, lampu mati, dan berbagai penyebab lainnya.

#### c. Faktor Alam atau Lingkungan

Faktor lingkungan ini bisa berupa pengaruh cuaca yang tidak menguntungkan, kondisi lingkungan jalan, benda-benda asing yang ada di jalan dan lain sebagainya.

Selanjutnya adalah menghitung nilai variabel untuk kecelakaan lalu lintas, namun untuk mendapatkan hasil yang nyata maka setiap faktor kecelakaan dilakukan penilaian persentase, jadi nantinya dari setiap faktor akan diketahui menyumbang berapa persen (%) untuk setiap kasus kecelakaan. Berikut ini adalah tabel persentase faktor kecelakaan, data yang digunakan adalah kecelakaan tahun 2019, yaitu:

Tabel 1 Tabel Persentase Faktor Kecelakaan Lalu Lintas

Tahun	Bulan	Faktor Manusia %	Faktor Kendaraan %	Faktor Alam atau Lingkungan %	Jumlah Kecelakaan
2019	Januari	18,75	61,46	19,79	96
	Februari	15,91	50	34,09	44
	Maret	12	28	60	25
	April	20	36	44	25
	Mei	23,64	61,82	14,55	55

Tabel 2 Tabel Persentase Faktor Kecelakaan Lalu Lintas(Lanjutan)

Tahun	Bulan	Faktor Manusia %	Faktor Kendaraan %	Faktor Alam atau Lingkungan %	Jumlah Kecelakaan
2019	Juni	23,68	58,55	17,76	152
	Juli	15,49	66,20	18,31	71
	Agustus	10,34	27,59	62,07	29
	September	9,52	38,10	52,38	21
	Oktober	3,23	41,94	54,84	31
	November	2,44	46,34	51,22	41
	Desember	14,17	55,83	30	120

Selanjutnya adalah menghitung nilai variabel untuk kecelakaan lalu lintas, untuk pengujiannya maka dibuat data tabel penilaian variabel kecelakaan lalu lintas sebagai berikut.



Tabel 3 Data Nilai Variabel Kecelakaan Lalu Lintas Tahun 2019

X1	X2	X3	Y	X1.Y	X1^2	X2.Y	X2^2	X3.Y	X3^2	X1.X2	X2.X3	X1.X3
18,75	61,46	19,79	96	1800	351,56	5900	3777,13	1900	391,71	1152,34	1216,36	371,09
15,91	50	34,09	44	700	253,10	2200	2500	1500	1162,19	795,45	1704,55	542,36
12	28	60	25	300	144	700	784	1500	3600	336	1680	720
20	36	44	25	500	400	900	1296	1100	1936	720	1584	880
23,64	61,82	14,55	55	1300	558,68	3400	3821,49	800	211,57	1461,16	899,17	343,80
23,68	58,55	17,76	152	3600	560,94	8900	3428,41	2700	315,53	1386,77	1040,08	420,71
15,49	66,20	18,31	71	1100	240,03	4700	4382,07	1300	335,25	1025,59	1212,06	283,67
10,34	27,59	62,07	29	300	107,02	800	761	1800	3852,56	285,37	1712,25	642,09
9,52	38,10	52,38	21	200	90,70	800	1451,25	1100	2743,76	362,81	1995,46	498,87
3,23	41,94	54,84	31	100	10,41	1300	1758,58	1700	3007,28	135,28	2299,69	176,90
2,44	46,34	51,22	41	100	5,95	1900	2147,53	2100	2623,44	113,03	2373,59	124,93
14,17	55,83	30	120	1700	200,69	6700	3117,36	3600	900	790,97	1675	425
169,17	571,82	459,01	710	11700	2923,08	38200	29224,82	21100	21079,29	8564,78	19392,21	5429,41

Dari tabel 3 di atas, di dapat nilai variabel sebagai berikut:

- n = 12
- $\sum Y = 710$
- $\sum X_1 = 169,17$
- $\sum X_2 = 571,82$
- $\sum X_3 = 459,01$
- $\sum X_1 \cdot Y = 11700$
- $\sum X_2 \cdot Y = 38200$
- $\sum X_3 \cdot Y = 21100$
- $\sum X^2_1 = 2923,08$
- $\sum X^2_2 = 29224,82$
- $\sum X^2_3 = 21079,29$
- $\sum X_1 \cdot X_2 = 8564,78$
- $\sum X_1 \cdot X_3 = 5429,41$
- $\sum X_2 \cdot X_3 = 19392,21$

3. Melakukan perhitungan terhadap data yang telah didapat nilai variabelnya dengan persamaan sebagai berikut:

$$Y = \beta_0 + X_1 \beta_1 + X_2 \beta_2 + X_3 \beta_3$$

Pada proses ini dilanjutkan ke proses 4.

4. Dari persamaan di atas, kemudian dilakukan turunan yaitu sebagai berikut:

$$\sum Y = \beta_0 \cdot n + \beta_1 \cdot \sum X_1 + \beta_2 \cdot \sum X_2 + \beta_3 \cdot \sum X_3$$

$$\sum Y = \beta_0 \cdot \sum X_1 + \beta_1 \cdot \sum X_1 \cdot X_1 + \beta_2 \cdot \sum X_1 \cdot X_2 + \beta_3 \cdot \sum X_1 \cdot X_3$$

$$\sum Y = \beta_0 \cdot \sum X_2 + \beta_1 \cdot \sum X_2 \cdot X_1 + \beta_2 \cdot \sum X_2 \cdot X_2 + \beta_3 \cdot \sum X_2 \cdot X_3$$

$$\sum Y = \beta_0 \cdot \sum X_3 + \beta_1 \cdot \sum X_3 \cdot X_1 + \beta_2 \cdot \sum X_3 \cdot X_2 + \beta_3 \cdot \sum X_3 \cdot X_3$$

Berikut ini adalah nilai variabel dari persamaan di atas:

$$710 = \beta_0 \cdot 12 + \beta_1 \cdot 169,17 + \beta_2 \cdot$$

$$571,82 + \beta_3 \cdot 459,01$$

$$11700 = \beta_0 \cdot 169,17 + \beta_1 \cdot 2923,08 + \beta_2 \cdot 29224,82 + \beta_3 \cdot 5429,41$$

$$38200 = \beta_0 \cdot 571,82 + \beta_1 \cdot 8564,78 + \beta_2 \cdot 29224,82 + \beta_3 \cdot 19392,21$$

$$21100 = \beta_0 \cdot 459,01 + \beta_1 \cdot 5429,41 + \beta_2 \cdot 19392,21 + \beta_3 \cdot 21079,29$$

5. Melakukan eliminasi untuk setiap turunan, yaitu sebagai berikut:

Pertama melakukan eliminasi untuk turunan 1 dan 2. Hasil dari eliminasi ini akan menghasilkan turunan 5. Berikut ini penyelesaiannya:

$$710 = \beta_0 \cdot 12 + \beta_1 \cdot 169,17 + \beta_2 \cdot 571,82 + \beta_3 \cdot 459,01$$

$$11700 = \beta_0 \cdot 169,17 + \beta_1 \cdot 2923,08 + \beta_2 \cdot 29224,82 + \beta_3 \cdot 5429,41$$

$$120110,7 = \beta_0 \cdot 2030,04 + \beta_1 \cdot 28618,489 + \beta_2 \cdot 96736,48 + \beta_3 \cdot 77650,72$$

$$140400 = \beta_0 \cdot 2030,04 + \beta_1 \cdot 35076,96 + \beta_2 \cdot 102777,24 + \beta_3 \cdot 65153,04$$

$$-20289,3 = \beta_1 \cdot -6458,47 + \beta_2 \cdot -6040,76 + \beta_3 \cdot 12497,68 \text{ (turunan 5)}$$

Selanjutnya melakukan eliminasi untuk turunan 1 dan 3. Hasil dari eliminasi ini akan menghasilkan turunan 6. Berikut ini penyelesaiannya:

$$710 = \beta_0 \cdot 12 + \beta_1 \cdot 169,17 + \beta_2 \cdot 571,82 + \beta_3 \cdot 459,01$$

$$38200 = \beta_0 \cdot 571,82 + \beta_1 \cdot 8564,78 + \beta_2 \cdot 29224,82 + \beta_3 \cdot 19392,21$$

$$405999,3 = \beta_0 \cdot 6861,96 + \beta_1 \cdot 96736,48 + \beta_2 \cdot 326989,55 + \beta_3 \cdot 262475,69$$

$$458400 = \beta_0 \cdot 6861,96 + \beta_1 \cdot 102777,24 + \beta_2 \cdot 350697,96 + \beta_3 \cdot 232706,64$$

$$-52400,7 = \beta_1 \cdot -6040,76 + \beta_2 \cdot -23708,41 + \beta_3 \cdot 29769,05 \text{ (turunan 6)}$$

Selanjutnya melakukan eliminasi untuk turunan 1 dan 4. Hasil dari eliminasi ini akan menghasilkan turunan 7. Berikut ini penyelesaiannya:

$$710 = \beta_0 \cdot 12 + \beta_1 \cdot 169,17 + \beta_2 \cdot 571,82 + \beta_3 \cdot 459,01$$

$$21100 = \beta_0 \cdot 459,01 + \beta_1 \cdot 5429,41 + \beta_2 \cdot 19392,21 + \beta_3 \cdot 21079,29$$

$$325897,1 = \beta_0 . 5508,12 + \beta_1 . 77650,72 + \beta_2 . 262475,69 + \beta_3 . 210690,18$$

$$253200 = \beta_0 . 5508,12 + \beta_1 . 65153,04 + \beta_2 . 232706,64 + \beta_3 . 252951,48$$

$$72697,1 = \beta_1 . 12497,68 + \beta_2 . 29769,05 + \beta_3 . -42261,3 \text{ (turunan 7)}$$

Selanjutnya melakukan eliminasi untuk turunan 5 dan 6. Hasil dari eliminasi ini akan menghasilkan turunan 8. Berikut ini hasil turunannya:

$$-215865637,02 = \beta_2 . -116629319,83 + \beta_3 . 116767056,16 \text{ (turunan 8)}$$

Selanjutnya melakukan eliminasi untuk turunan 5 dan 7. Hasil dari eliminasi ini akan menghasilkan turunan 9.

Berikut ini hasil turunannya:

$$215942906,09 = \beta_2 . 116767056,16 + \beta_3 . -116751336,18 \text{ (turunan 9)}$$

Selanjutnya melakukan eliminasi untuk turunan 8 dan 9. Hasil dari eliminasi ini akan menghasilkan turunan 10.

Berikut ini hasil turunannya:

$$-20720702434828 = \beta_3 . 17916477104446 \text{ (turunan 10)}$$

$$215942906,09 - 135024853,5 = \beta_2 . 116767056,16$$

$$80918052,61 = \beta_2 . 116767056,16$$

$$\beta_2 = 80918052,61 / 116767056,16$$

$$\beta_2 = 0,692987006$$

Selanjutnya mencari nilai  $\beta_1$  dengan menggunakan turunan 7

$$72697,1 = \beta_1 . 12497,68 + \beta_2 . 29769,05 + \beta_3 . -42261,3$$

$$72697,1 = \beta_1 . 12497,68 + (0,692987 . 29769,05) + (-1,15652 . -42261,3)$$

$$72697,1 = \beta_1 . 12497,68 + 20629,56 + 48876,04$$

$$72697,1 - 20629,56 - 48876,04 = \beta_1 . 12497,68$$

$$3191,497 = \beta_1 . 12497,68$$

$$\beta_1 = 3191,497 / 12497,68$$

$$\beta_1 = 0,255378824$$

Terakhir mencari nilai  $\beta_0$  dengan menggunakan turunan 1  $710 = \beta_0 . 12 + \beta_1 . 169,17 + \beta_2 . 571,82 + \beta_3 . 459,01$

$$710 = \beta_0 . 12 + (0,2554 . 169,17) + (0,693 . 571,82) + (-1,15652 . 459,01)$$

$$710 = \beta_0 . 12 + 42,18442 + 396,2713 + (-530,8543)$$

$$714 - 42,18442 - 396,2713 + 530,8543 = \beta_0 . 12$$

$$801,37947 = \beta_0 . 12$$

$$\beta_0 = 801,37947 / 12$$

$$\beta_0 = 66,7816225$$

#### 6. Hasil akhir

Proses terakhir dari algoritma sistem metode *regresi linier berganda* untuk mengestimasi tingkat kecelakaan lalu lintas di Kab. Serdang Bedagai menggunakan persamaan yaitu sebagai berikut:

$$Y = \beta_0 + X_1 \beta_1 + X_2 \beta_2 + X_3 \beta_3$$

Penyelesaian dengan persamaan di atas yaitu:

$$Y = 66,7816225 + X_1 \cdot 0,255378824 + X_2 \cdot 0,692987006 + X_3 \cdot -1,156516558$$

Untuk menentukan nilai  $X_1$ ,  $X_2$  dan  $X_3$  adalah dengan melihat nilai rata rata dari total data untuk setiap variabel, sehingga di dapat hasil sebagai berikut:

$$Y = 66,7816 + (5,83 \times 0,2554) + (41,75 \times 0,693) + (52,42 \times -1,15652)$$

$$Y = 66,7816 + 1,48886 + 28,932 + (-60,6246)$$

$$Y = 36,5781$$

Dari hasil diatas, di dapat kesimpulan yaitu jumlah estimasi tingkat kecelakaan lalu lintas di Serdang Bedagai untuk bulan Mei tahun 2020 adalah 36,5781 kasus dibulatkan menjadi **37** kasus. Dari hasil tersebut dapat di lihat bahwa kasus kecelakaan di perkirakan mengalami penurunan.

## 4 PENGUJIAN DAN IMPLEMENTASI

### 4.1 Tampilan Form Login

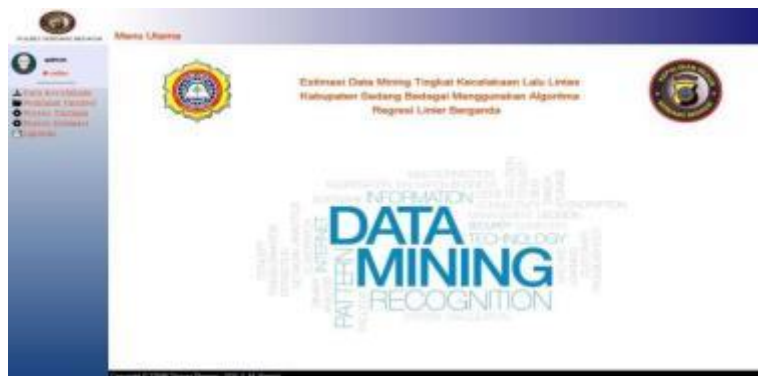
Berikut ini merupakan tampilan dari *form* login yang berfungsi untuk melakukan proses validasi *username* dan *password* pengguna.



Gambar 3 Tampilan Form Login

**4.2 Tampilan Form Menu Utama**

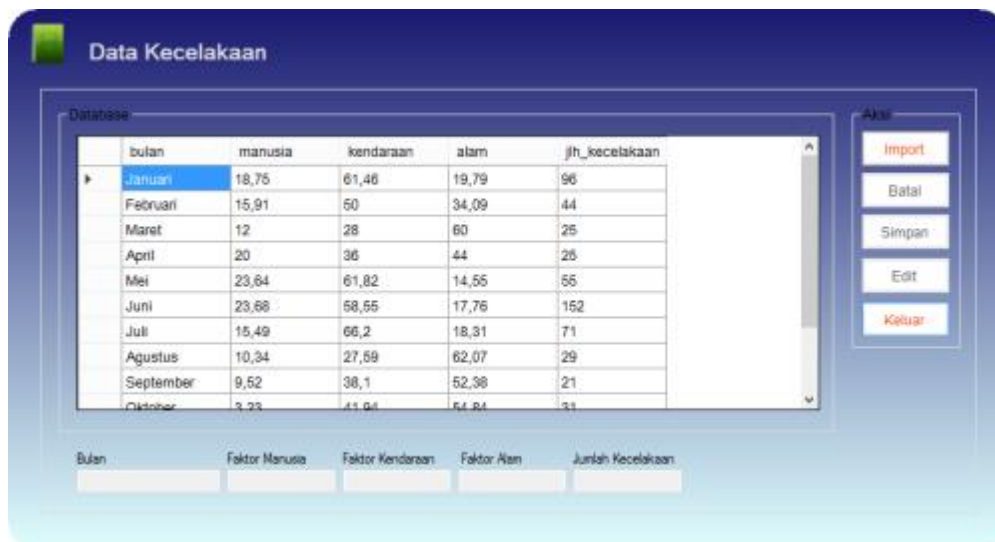
Berikut ini merupakan tampilan menu utama dari aplikasi untuk estimasi *data mining* tingkat kecelakaan lalu lintas Kabupaten Serdang Bedagai menggunakan algoritma *regresi linier* berganda:



Gambar 4 Tampilan Form Menu Utama

**4.3 Tampilan Form Masukan Data**

Berikut ini merupakan tampilan dari *form* masukan data yang berfungsi untuk menginput data-data kecelakaan lalu lintas:



Gambar 5 Tampilan Form Masukan Data

**4.4 Tampilan Form Proses Penilaian**

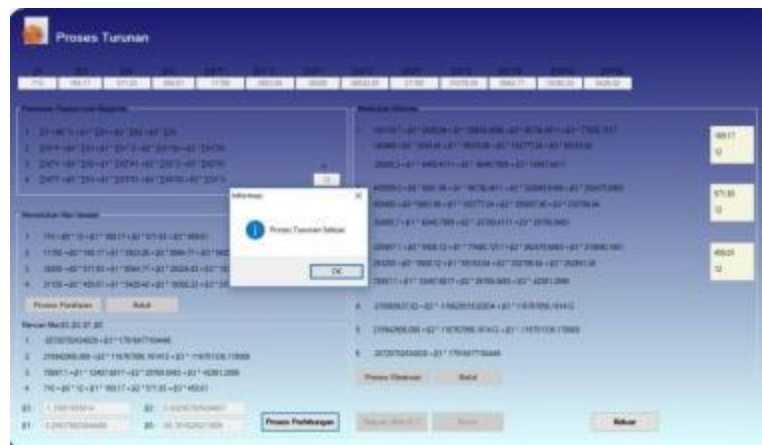
Berikut ini merupakan tampilan dari *form* proses penilaian yang berfungsi untuk melakukan proses penilaian terhadap data kecelakaan lalu lintas:



Gambar 6 Tampilan Form Proses Penilaian

**4.5 Tampilan Form Proses Turunan**

Berikut ini merupakan tampilan dari *form* proses turunan yang berfungsi untuk melakukan proses turunan terhadap data kecelakaan lalu lintas:



Gambar 7 Tampilan Form Proses Turunan

**4.6 Tampilan Form Proses perhitungan Estimasi**

Berikut ini merupakan tampilan dari *form* proses perhitungan estimasi kecelakaan lalu lintas yang berfungsi untuk melakukan proses estimasi terhadap data kecelakaan lalu lintas:



Gambar 8 Tampilan *Form* Masukan Data Basis Aturan

**4.7 Tampilan Form Laporan Hasil Estimasi**

Berikut ini merupakan tampilan dari *form* laporan hasil estimasi yang berfungsi untuk menampilkan hasil dari estimasi jumlah kecelakaan lalu lintas:



Gambar 9 Tampilan *Form* Laporan Hasil Estimasi

Pada laporan hasil estimasi diatas diketahui bahwa kasus kecelakaan di perkirakan mengalami penurunan, hal tersebut dikarenakan penerapan algoritma *regresi linier berganda*.

**5 KESIMPULAN**

Berdasarkan perumusan dan pembahasan bab-bab sebelumnya dapat diambil kesimpulan sebagai berikut :

1. Dalam menerapkan algoritma *regresi linier berganda* untuk estimasi tingkat kecelakaan lalu lintas di Kabupaten Serdang Bedagai yaitu dengan menentukan terlebih dahulu variabel yang menjadi sebab terjadinya kecelakaan, kemudian melakukan perhitungan dengan algoritma *regresi linier berganda* terhadap data kecelakaan, selanjutnya memasukkan perhitungan algoritma *regresi linier berganda* ke dalam *coding* program untuk melihat apakah proses perhitungan berjalan dengan baik atau tidak.
2. Dalam merancang aplikasi untuk estimasi tingkat kecelakaan lalu lintas di Kabupaten Serdang Bedagai yang menerapkan algoritma *regresi linier berganda* yaitu dengan membangun

aplikasi berbasis *desktop programming* kemudian membuat *form-form* yang berkaitan dan mendukung untuk proses estimasi seperti membuat *form* data kecelakaan, *form* penilaian variabel, *form* proses turunan dan membuat *form* proses estimasi. Setelah semua *form* dibuat selanjutnya melakukan proses estimasi.

3. Sistem yang telah dirancang selanjutnya diuji dan diimplementasikan dengan memasukkan data-data sesuai dengan yang ada pada bab-bab sebelumnya, kemudian jika hasil *output* nya sesuai dengan data manual maka dalam pengujian ini sistem berjalan dengan baik, menambahkan data ke *database*, perintah *update* untuk merubah data di *database*, perintah *delete* untuk menghapus data di *database*.

## 5.1 Saran

Untuk lebih mengembangkan dan meningkatkan sistem dalam mengestimasi tingkat kecelakaan lalu lintas di Kabupaten Serdang Bedagai ada beberapa saran yang dapat dijadikan pertimbangan, yaitu :

1. Perangkat lunak sebaiknya menggunakan bahasa pemrograman berbasis web sehingga dapat dengan mudah diakses oleh pihak Polres Serdang Bedagai dengan media apa saja.
2. Sistem dirasa belum sempurna, untuk itu perlu dilakukan pengembangan bagi penelitian selanjutnya seperti penambahan variabel baru secara otomatis.

## UCAPAN TERIMA KASIH


Puji syukur kehadiran Allah SWT atas izin-Nya yang telah melimpahkan rahmat dan karunia-Nya sehingga dapat menyelesaikan jurnal ilmiah ini. Pada kesempatan ini diucapkan terima kasih yang sebesar-besarnya kepada kedua Orang Tua tercinta yang selama ini memberikan do'a dan dorongan baik secara moril maupun materi sehingga dapat terselesaikan pendidikan dari tingkat dasar sampai bangku perkuliahan dan terselesaikannya jurnal ini. Di dalam penyusunan jurnal ini, banyak sekali bimbingan yang didapatkan serta arahan dan bantuan dari pihak yang sangat mendukung. Oleh karena itu dengan segala kerendahan hati, diucapkan terima kasih yang sebesar-besarnya kepada Bapak Rudi Gunawan, SE., M.Si., selaku Ketua Sekolah Tinggi Manajemen Informatika Dan Komputer (STMIK) Triguna Dharma Medan. Bapak Dr. Zulfian Azmi, ST., M.Kom., selaku Wakil Ketua I Bidang Akademik STMIK Triguna Dharma Medan. Bapak Marsono, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi STMIK Triguna Dharma Medan. Bapak Azanuddin, S.Kom., M.Kom., selaku Dosen Pembimbing I yang telah memberikan saran, arahan dan dukungannya serta motivasi, sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Ibu Elfutriani, S.Pd., M.Si., selaku Dosen Pembimbing II yang telah memberikan bimbingan tata cara penulisan, saran dan motivasi sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Seluruh Dosen, Staff dan Pegawai di STMIK Triguna Dharma Medan.



**REFERENSI**

- [1] Aztria Dharma and Bambang MT Edison, "Identifikasi Kecelakaan Lalu Lintas (Study Kasus Jalan Dalu-Dalu Sampai Pasir Pengaraian),".
- [2] "15912-ID-faktor-faktor-yang-mempengaruhi-keparahan-korban-kecelakaan-lalu-lintas-di-kota".
- [3] Jurnal Sains and Dan Seni Its, "928X Print," 2016.
- [4] Maritim Raja Ali Haji, Marisa Elsera, and Universitas Maritim Raja Ali Haji, "Makna Keteraturan Berlalu Lintas," vol. 4, no. 1, 2016.
- [5] Rini Astuti, "Rini Astuti *Data Mining* untuk Lasifikasi dengan Algoritma CART,".
- [6] Gunadi Widi Nurcahyo Dicky Nofriansyah, *Algoritma Data Mining dan Pengujiannya*. Yogyakarta, Indoensia: Budi Utama, 2015
- [7] Retno Tri Wulandari, *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta, Indonesia: Gava Media, 2017.
- [8] Eka Miranda, Hans Yonathan, Ivander G Lokanatha, ; Josep, And Benedictus Suryawan, "*Data Mining* Untuk Menganalisis Gagal Serah Dana Pada Transaksi Jual Beli Saham,".
- [9] "Estimasi Parameter Model Regresi Linier Pada Data Yang Mengandung Outlier Dengan Metode Maximum Likelihood Estimation,". Amrin, "*Data Mining* Dengan Regresi Linier Berganda Untuk Peramalan Tingkat Inflasi," 2016.

**BIOGRAFI PENULIS**

	<p><b>M. Harmizi.</b> Laki - laki kelahiran Tualang, 28 Mei 1998, Anak kelima dari lima bersaudara ini merupakan seorang mahasiswa STMIK Triguna Dharma yang sedang dalam proses menyelesaikan skripsi.</p>
	<p><b>Azanuddin, S.Kom., M. Kom.</b> Beliau merupakan dosen tetap di STMIK Triguna Dharma Medan dan aktif sebagai pengajar pada bidang ilmu Sistem Informasi.</p>
	<p><b>Elfitriani, S.Pd., M.SI</b> Beliau merupakan dosen tetap di STMIK Triguna Dharma Medan dan aktif sebagai pengajar pada bidang ilmu Bahasa Inggris.</p>