

DATA MINING UNTUK MENGESTIMASI JUMLAH PENUMPANG PADA PT. PINEM LAU GUNA MEDAN DENGAN MENGGUNAKAN METODE REGRESI LINIER BERGANDA

Yosuah Goktua Siadari*, Darjat Saripuna**, Purwadi**

* Program Studi Sistem Informasi, STMIK Triguna Dharma

** Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Keyword:

Data Mining ,Regresi Linier Berganda

ABSTRACT

Estimasi jumlah data penumpang merupakan salah satu hal yang harus diketahui oleh setiap perusahaan transportasi terutama pada PT. Pinem Lau Guna Medan selama ini belum memiliki aplikasi yang mampu membantu karyawan dalam mengestimasi jumlah penumpang. Hal ini mengakibatkan ketidakseimbangan pengeluaran selama bus beroperasi dengan pendapatan yang didapat dari para penumpang yang hendak berpergian.

Untuk menyelesaikan masalah dalam mngestimasi data penumpang membutuhkan proses data mining dengan menggunakan metode regresi linier berganda(RGB) dan beberapa kriteria yang akan digunakan dalam mengestimasi data penumpang pada PT. Pinem Lau Guna Medan.

Melalui proses data mining akan dapat dilakukan transformasi data untuk proses estimasi jumlah penumpang pada PT. Pinem Lau Guna Medan dalam jangka waktu tertentu. Data mining juga berisi pencarian pola yang diinginkan pada database untuk membantu mengetahui suatu kesimpulan dan pengambilan keputusan di waktu yang akan datang.

Copyright © 2020 STMIK Triguna Dharma.
All rights reserved.

First Author

Nama : Yosuah Goktua Siadari
Program Studi : Sistem Informasi STMIK Triguna Dharma
Email : yosuavtxsiadari@gmail.com

1. PENDAHULUAN

Pesatnya persaingan dalam dunia bisnis jasa bus antar kota antar provinsi (akap) baik bus ukuran kecil maupun yang besar sangat ketat saat ini. Keputusan untuk menggunakan jasa bus antar kota antar provinsi (akap) yang baik dan sesuai dengan kebutuhan adalah alasan yang tepat dari masalah kebutuhan yang banyak dirasakan konsumen dalam hal ini calon penumpang. Oleh karena itulah membangun kepercayaan penumpang dan citra perusahaan yang baik sangatlah penting untuk merahi pangsa pasar transportasi yang diharapkan oleh perusahaan jasa transportasi.

Kondisi tersebut diatas, mengakibatkan perusahaan-perusahaan yang bergerak dibidang jasa transportasi saat ini sedang mempersiapkan dan memperbaiki layanan kepada calon penumpang agar dapat memperediksi jumlah penumpang. Salah satu diantara perusahaan jasa transportasi darat tersebut adalah PT. Pinem Lau Guna Medan. Pada penelitian ini akan menjelaskan bagaimana memanfaatkan data mining dalam mengestimasi jumlah penumpang dengan menggunakan metode *regresi linier berganda*. Dengan memanfaatkan data yang ada maka dapat dilakukan proses mining untuk mengetahui jumlah data yang dicari. Dari proses mining tersebut dapat dilakukan transformasi data untuk proses estimasi jumlah penumpang pada PT. Pinem Lau Guna Medan dalam jangka waktu tertentu. Pada jurnal dijelaskan data *mining* merupakan disiplin ilmu yang bertujuan untuk menemukan, menggali pengetahuan dari data atau informasi yang kita miliki.[1] Data mining berisi pencarian pola yang diinginkan pada database untuk membantu mengetahui suatu kesimpulan dan pengambilan keputusan di waktu yang akan datang.[2] Dalam data *mining* untuk mengestimasi jumlah penumpang dapat menggunakan metode *Regresi Linier Berganda*.

Regresi Linier Berganda merupakan salah satu metode dalam data mining yang memiliki variable bebas lebih dari satu.[3] Selain itu juga pada penelitian lain seperti dijelaskan bahwasannya metode Regresi Linier Berganda dapat diterapkan untuk mengetahui ada tidaknya pengaruh signifikan dua atau lebih variable bebas terhadap variable terikat, untuk mengetahui prediksi produksi barang, dan untuk memprediksi jumlah penumpang.[4]

Dari data yang telah dianalisis dengan Algoritma *Regresi Linier Berganda* akan menghasilkan suatu kesimpulan dalam memprediksi jumlah penumpang pada PT. Pinem Lau Guna Medan. Sehingga kondisi diatas maka di angkatlah judul penelitian yaitu: “**Data Mining Untuk Mengestimasi Jumlah Penumpang Pada PT. PINEM LAU GUNA Medan Dengan Menggunakan Metode Regresi Linear Berganda**”.

2. KAJIAN PUSTAKA

2.1 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat di perkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.[5]. Berdasarkan teknik prediksi dapat dibagi menjadi dua bagian yaitu:

1. Prediksi Kuantitatif : merupakan data kuantitatif yang terjadi pada kejadian yang telah berlalu.
2. Prediksi Kualitatif : adalah data kualitatif yang ada pada masa lalu menggunakan data dari variabel yang akan diprediksi tidak ada, hilang, tidak cukup atau tidak dipercaya.

2.2 Knowledge Discovery Database (KDD)

KDD merupakan proses analisa terstruktur untuk memproleh informasi yang mendekati kebenaran, baru, bermanfaat, dan menemukan pola dari data yang besar dan komplek. *Data mining* (DM) menjadi inti dari proses KDD, yaitu dengan menggunakan algoritma tertentu untuk mengeksplorasi data, membangun model dan menemukan pola yang belum diketahui. Model digunakan untuk memahami fenomena data, analisa maupun prediksi.[7].

2.3 Tahap-Tahap Knowledge Discovery Database (KDD)

Menurut Yuli Mardi dalam proses KDD secara garis besar dapat dijelaskan sebagai berikut:

1. *Data Selection*
Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database(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* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Kesalahan mungkin terjadi karena adanya pembuangan pada duplikasi datanya Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *Knowledge Discovery in Database(KDD)*, seperti data atau informasi eksternal lainnya yang diperlukan.
3. *Transformation*
Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam *Knowledge Discovery in Database (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 menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi.
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 *Knowledge Discovery in Database (KDD)* yang disebut *interpretation*. [8]

2.4 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi didalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan yang berguna dan bermanfaat yang tersimpan didalam database. Perlu diingat bahwa kata *mining* sendiri dapat diartikan sebagai kegiatan atau usaha untuk mendapatkan sedikit barang berharga ataupun seperti informasi dari sejumlah besar material dasar. Oleh Karena ini *Data Mining* sebenarnya memiliki susunan yang panjang bersifat masa lalu dari berbagai bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistika dan penyimpanan

database. Data Mining juga adalah tindakan menganalisa ataupun analisis seseorang secara signifikan dari beberapa data yang umumnya data tersebut berjumlah besar, sehingga databersifat sangat kongkret ataupun kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya.[9]

Secara umum, tugas utama dari data *mining* dibagi menjadi dua yaitu :

1. *Descriptive*

Merupakan kemampuan untuk mengidentifikasi keunikan data, pola, *trend*, hubungan dan *anomaly* pada data. *Descriptive* dibagi menjadi asosiasi, segmentasi, dan *clustering*.

2. *Predictive* merupakan pengembangan model dari beberapa fenomena yang memungkinkan dilakukan estimasi nilai dan prediksi untuk masa depan. *Predictive* dibagi menjadi klasifikasi dan regresi. Regresi termasuk kepada estimasi dan peramalan ataupun prediksi.[11]

Menurut Goldie Gunaidi, Dana Indra Senses dalam data *mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

1. Deskripsi

Terkadang penelitian dan analisa secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

2. Estimasi

Estimasi hamper sama dengan klasifikasi, kecuali variable target estimasi lebih kearah numeric dari pada kea rah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi.

3. Prediksi

Prediksi hamper sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil aka nada di masa mendatang.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklasteran (*Clusterring*)

Pengklasteran merupakan pengelompokan data, pengamatan yang dilakukan atau proses memperhatikan kemudian membentuk kelas objek-objek yang memiliki kemiripan.

6. Asosiasi

Tugas asosiasi dalam data *mining* adalah untuk menemukan atribut yang muncul dalam suatu waktu.

2.5 Regresi Linear Berganda

Metode *regresi linear* berganda merupakan suatu analisis regresi dalam menjelaskan hubungan antara pengubah (variabel dependen) dengan factor-faktor yang mempengaruhi lebih dari satu prediktor (variabel independen).

Persamaan *Regresi Linear* Berganda adalah :

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Keterangan :

Y = Variabel tidak bebas (nilai yang diprediksikan)

X = Variabel Bebas

a = Konstanta (nilai Y apabila $X_1, X_2 \dots X_n = 0$)

b = Koefisien regresi (nilai peningkatan ataupun penurunan).[13]

3. METODOLOGI PENELITIAN

3.1 Metodologi Penelitian

Metode penelitian merupakan cara utama yang di lakukan peneliti yang bertujuan untuk menentukan jawaban atas masalah yang timbul di masyarakat atau yang diajukan dan mencapai tujuan sesuai prosedur. Jadi, metode penelitian adalah prosedur atau cara untuk menyelesaikan masalah yang sedang di hadapi.

Dalam melakukan penelitian ada beberapa cara atau langkah-langkah yang digunakan yaitu sebagai berikut:

1. Pengumpulan data (*Data Collecting*). Dalam pengumplan data terdapat beberapa cara yang dilakukan yaitu:

a. Observasi

Observasi merupakan aktivitas terhadap suatu teknik pengumpulan data yang dilakukan secara langsung dimana peneliti melakukan penelitian di tempat studi kasus tersebut. Dengan dilakukannya penelitian ini, peneliti melakukan observasi pada perusahaan PT. Pinem Lau Guna Medan. Dimana peneliti mengamati data jumlah penumpang yang berangkat dari loket bus PT. Pinem Lau Guna Medan tersebut.

b. Wawancara

Wawancara adalah salah satu metode pengumpulan data yang dilakukan dengan bertanya langsung kepada bagian perusahaan yang berhubungan dengan penelitian tersebut dalam hal ini adalah bagian pemesanan tiketnya. Data yang dapat peneliti peroleh dari perusahaan tersebut adalah data primer dimana data yang diperoleh merupakan data dari perusahaan PT. Pinem Lau Guna Medan secara langsung.

2. Studi Literatur

Dalam studi literatur peneliti banyak menggunakan jurnal-jurnal baik jurnal internasional, jurnal nasional, jurnal local, maupun buku sebagai referensi. Studi literatur ini dilakukan oleh peneliti untuk menentukan topik penelitian dan rumusan masalah sebelum terjun kelapangan dalam pengumpulan data yang diperlukan. Studi Literatur merupakan serangkaian kegiatan yang dilakukan peneliti untuk mengumpulkan data pustaka, membaca dan mencatat yang berkaitan dengan masalah dan tujuan penelitian.

3.2 Metode Perancangan Sistem

Metode adalah "*Methodos* Metode perancangan sistem merupakan salah satu unsur penting dalam penelitian khususnya menggunakan *software* atau perangkat lunak. Dalam metode perancangan sistem sering menggunakan metode algoritma *waterfall* atau algoritma air terjun. Dalam hal ini penelitian menggunakan metode algoritma air terjun. Berikut ini adalah fase yang dilakukan dalam dalam penelitian adalah sebagai berikut:

1. Analisis Kebutuhan Perangkat lunak
2. Desain Sistem
3. Pembuatan Kode Program
4. Pengujian
5. Implementasi atau Pemeliharaan (maintenance)

3.3 Algoritma Sistem

Algoritma tidak terlepas dari sistem komputer. Algoritma diterapkan berbeda-beda dengan kriteria yang sama untuk suatu permasalahan. Algoritma sistem merupakan langkah – langkah logis dan sistematis untuk menyelesaikan pada suatu masalah dalam metode Regresi Liner Berganda dalam memprediksi jumlah penumpang perbulannya pada PT. Pinem Lau Guna Medan.

3.3.1 Penyelesaian

Berikut ini adalah data *penumpang* yang didapat dari PT PINEM LAU GUNA MEDAN, yang akan diamankan. Dalam pengujiannya, sebagai contoh data yang digunakan sebagai sampel dalam penelitian ini yaitu sebagai berikut:

Tabel 3.1 Data Primer Dari Perusahaan

Bulan	Hari Libur (Dalam/Bln)	Jumlah Unit Berangkat	Harga Tiket	Jumlah Penumpang
Januari	1	60	Rp 180.000	165
Februari	1	63	Rp 180.000	172
Maret	1	61	Rp 180.000	176
April	3	68	Rp 185.000	217
Mei	3	65	Rp 185.000	219
Juni	6	72	Rp 200.000	324
Juli	0	60	Rp 180.000	126
Agustus	2	62	Rp 185.000	187

September	1	61	Rp 180.000	176
Oktober	0	62	Rp 180.000	119
November	1	61	Rp 185.000	134
Desember	2	70	Rp 200.000	289

3.3.2 Deskripsi Data Dari Penelitian

Berikut ini merupakan data-data yang digunakan sebagai sampel dalam penelitian yaitu sebagai berikut :

Table 3.1. Data penumpang yang digunakan dalam penelitian

PT PINEM LAU GUNA MEDAN				
Bulan	Hari Libur (Dalam/Bln)	Jumlah Unit Berangkat	Harga Tiket	Jumlah Penumpang
	X1	X2	X3	Y
Januari	1	60	Rp 180.000	165
Februari	1	63	Rp 180.000	172
Maret	1	61	Rp 180.000	176
April	3	68	Rp 185.000	217
Mei	3	65	Rp 185.000	219
Juni	6	72	Rp 200.000	324
Juli	0	60	Rp 180.000	126
Agustus	2	62	Rp 185.000	187
September	1	61	Rp 180.000	176
Oktober	0	62	Rp 180.000	119
November	1	61	Rp 185.000	134
Desember	2	70	Rp 200.000	134
Jumlah	21	765	Rp 2.220.000	2304

3.3.3 Menghitung Koefisien Regresi

Koefisien regresi adalah hal penting dalam menganalisa regresi. Manfaat dari koefisien regresi adalah untuk membentuk model persamaan regresi pada suatu masalah yang di teliti. Menghitung koefisien regresi ada dua cara yaitu dengan cara perhitungan matematika biasa dan menghitung menggunakan *software* statistic (SPSS), namun hasil dari menghitung dengan aplikasi dan manual harus sama atau setidaknya menghasilkan nilai yang mendekati sama.

Tabel 3.2. Koefisien Regresi

X1*Y	X2*Y	X3*Y	Y ²	X1 ²
165	9900	29700000	27225	1
172	10836	30960000	29584	1

176	10736	31680000	30976	1
651	14756	40145000	47089	9
657	14235	40515000	47961	9
1944	23328	64800000	104976	36
0	7560	22680000	15876	0
374	11594	34595000	34969	4
176	10736	31680000	30976	1
0	7378	21420000	14161	0
134	8174	24790000	17956	1
578	20230	57800000	83521	4
5027	149463	430765000	485270	67

Tabel 3.2. Koefisien Regresi (lanjutan)

X2^2	X3^2	X1.X2	X1.X3	X2.X3
3600	32400000000	60	180000	10800000
3969	32400000000	63	180000	11340000
3721	32400000000	61	180000	10980000
4624	34225000000	204	555000	12580000
4225	34225000000	195	555000	12025000
5184	40000000000	432	1200000	14400000
3600	32400000000	0	0	10800000
3844	34225000000	124	370000	11470000
3721	32400000000	61	180000	10980000
3844	32400000000	0	0	11160000
3721	34225000000	61	185000	11285000
4900	40000000000	140	400000	14000000
48953	411300000000	1401	3985000	141820000

3.3.4 Menyederhanakan Persamaan Regresi Linier Berganda

Dari perhitungan koefisien Regresi Linear Berganda diatas maka digunakan persamaan sebagai berikut :

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3$$

Keterangan : Y = Jumlah Penumpang

a = Konstanta

b_1 = Koefisien Regresi X_1

b_2 = Koefisien Regresi X_2

b_3 = Koefisien Regresi X_3

X_1 = Hari Libur Dalam Sebulan

X_2 = Jumlah Keberangkatan

X_3 = Harga Tiket

Selanjutnya melakukan perhitungan Regresi Linear Berganda dengan Persamaan sebagai berikut :

Rumus perhitungan Regresi Linear Berganda antara lain sebagai berikut :

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3$$

$$\sum X_1 Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + b_3 \sum X_1 X_3$$

$$\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 X_3$$

$$\sum X_3 Y = a \sum X_3 + b_1 \sum X_1 X_3 + b_2 \sum X_2 X_3 + b_3 \sum X_3^2$$

Tabel 3.3. Hasil Perhitungan Persamaan Regresi

Y	A	B1	B2	B3
2.304	12	21	765	2.220.000
5.027	21	67	1.401	3.985.000
149.463	765	1.401	48.953	141.820.000
430.765.000	2.220.000	3.985.000	141.820.000	411.300.000.000

1. Untuk persamaan $\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3$ dan koefisien regresi hasilnya sebagai berikut :

$$2.304 = 12a + 21b_1 + 765b_2 + 2.220.000b_3 \dots \dots \dots (1)$$

2. Untuk persamaan $\sum X_1 Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + b_3 \sum X_1 X_3$ dan koefisien regresi, hasilnya sebagai berikut :

$$5.027 = 21a + 67b_1 + 1.401b_2 + 3.985.000b_3 \dots \dots \dots (2)$$

3. Untuk persamaan $\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 X_3$ dan koefisien regresi, hasilnya sebagai berikut :

$$149.463 = 765a + 1.401b_1 + 48.953b_2 + 141.820.000b_3 \dots \dots (3)$$

4. Untuk persamaan $\sum X_3 Y = a \sum X_3 + b_1 \sum X_1 X_3 + b_2 \sum X_2 X_3 + b_3 \sum X_3^2$ dan koefisien regresi, hasilnya sebagai berikut :

$$430.765.000 = 2.220.000a + 3.985.000b_1 + 141.820.000b_2 + 411.300.000.000b_3 \dots \dots (4)$$

Selanjutnya melakukan proses eliminasi antara persamaan [1] dengan persamaan [2] adalah sebagai berikut :

$$\begin{array}{r} 2304 = 12a + 21b_1 + 765b_2 + 2220000b_3 \text{ [1] x 21} \\ 5027 = 21a + 67b_1 + 1401b_2 + 3985000b_3 \text{ [2] x 12} \\ \hline 48384 = 252a + 441b_1 + 16065b_2 + 46620000b_3 \\ 60324 = 252a + 804b_1 + 16812b_2 + 47820000b_3 \\ \hline -11940 = 0 + -363b_1 + -747b_2 + -1200000b_3 \text{ [5]} \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [1] dengan persamaan [3] adalah sebagai berikut :

$$\begin{array}{r} 2304 = 12a + 21b_1 + 765b_2 + 2220000b_3 \text{ [1] x 765} \\ 149463 = 765a + 1401b_1 + 48953b_2 + 141820000b_3 \text{ [3] x 12} \\ \hline 1762560 = 9180a + 16065b_1 + 585225b_2 + 1698300000 \\ 1793556 = 9180a + 16812b_1 + 587436b_2 + 1701840000 \\ \hline -30996 = 0 + -747b_1 + -2211b_2 + -3540000b_3 \text{ [6]} \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [1] dengan persamaan [4] adalah sebagai berikut :

$$\begin{array}{r} 2304 = 12a + 21b_1 + 765b_2 + 2220000b_3 \text{ [1] x 2220000} \\ 430765000 = 2220000a + 3985000b_1 + 141820000b_2 + 4113E+11b_3 \text{ [4] x 12} \\ \hline 511488000 = 26640000a + 46620000b_1 + 1698300000b_2 + 49284E+12b_3 \end{array}$$

$$\begin{array}{r} 5169180000 = 26640000a + 47820000b1 + 1701840000b2 + 49356E+12b3 \\ -54300000 = 0 + -1200000b1 + -3540000b2 + 7200000000b3 \\ \hline \dots\dots\dots[7] \end{array}$$

Setelah melakukan proses eliminasi antara persamaan [1] hingga persamaan [4], maka diperoleh persamaan baru yaitu sebagai berikut :

$$\begin{array}{r} -11940 = -363b1 + -747b2 + -1200000b3\dots\dots\dots[5] \\ -30996 = -747b1 + -2211b2 + -3540000b3\dots\dots[6] \\ -54300000 = -1200000b1 + -3540000b2 + -7200000000b3\dots\dots\dots[7] \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [5] dengan persamaan [6] adalah sebagai berikut :

$$\begin{array}{r} -11940 = -363b1 + -747b2 + -1200000b3 [5] \times -747 \\ -30996 = -747b1 + -2211b2 + -3540000b3 [6] \times -4851 \\ \hline 8919180 = 271161b1 + 558009b2 + 896400000b3 \\ 11251548 = 271161b1 + 802593b2 + 1285020000b3 \\ \hline -2332368 = 0 + -244584b2 + -388620000b3\dots\dots\dots[8] \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [6] dengan persamaan [7] adalah sebagai berikut :

$$\begin{array}{r} -30996 = -747b1 + -2211b2 + -3540000b3 [6] \times -120000 \\ -54300000 = -1200000b1 + -3540000b2 + -7200000000b3 [7] \times -747 \\ \hline 37195200000 = 896400000b1 + 2653200000b2 + 4248E+12b3 \\ 40562100000 = 896400000b1 + 2644380000b2 + 53784E+12b3 \\ \hline -3366900000 = 0 + 8820000b2 + -11304E+12b3\dots\dots[9] \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [8] dengan persamaan [9] adalah sebagai berikut :

$$\begin{array}{r} -2332368 = -244584b2 + -388620000b3 [8] \times 8820000 \\ -3366900000 = 8820000b2 + -11304E+12b3 [9] \times -244584 \\ \hline -205715E+13 = -215723E+12b2 + -342763E+15b3 \\ 82349E+14 = -215723E+12b2 + 276478E+17b3 \\ \hline -844061E+14 = 0 + -279905E+17b3\dots\dots\dots[10] \end{array}$$

$$b3 = -844061E+14 / -279905E+17$$

$$b3 = 0,003015$$

Kemudian untuk menghasilkan nilai b2 maka hasil dari b3 dimasukan ke persamaan [8] atau persamaan [9], dalam hal ini menggunakan persamaan [8] sebagai berikut :

$$\begin{array}{r} -2332368 = -244584b2 + -388620000b3 \\ -2332368 = -244584b2 + (-388620000 * 0,003015) \\ -2332368 = -244584b2 + -1771893 \\ -2332368 - (-1771893) = -244584b2 \\ -1160475 = -244584b2 \\ b2 = -1160475 / -244584b2 \quad b2 = 4,74468 \end{array}$$

Kemudian untuk mendapatkan nilai b1 maka hasil dari b2 dan b3 di masukan ke persamaan [5] atau persamaan [6] ataupun persamaan [7], dalam hal ini menggunakan persamaan [7] sebagai berikut ;

$$\begin{array}{r} -54300000 = -1200000b1 + -3540000b2 + -7200000000b3 \\ -54300000 = -1200000b1 + (-3540000*4,74468) + (-7200000000*0,003015) \\ -54300000 = -1200000b1 + -16796200 + -21711773 \\ -54300000 = -1200000b1 + -38507973 \\ -1200000b1 = (-54300000) - (-38507973) \\ -1200000b1 = -15792027 \\ b1 = -15792027 / -1200000 \end{array}$$

$$b1 = 13,16002$$

Kemudian untuk mendapatkan nilai a maka hasil dari b1, b2, b3 dimasukan ke persamaan [1] atau [2] atau [3] ataupun [4], dalam hal ini menggunakan persamaan [1] sebagai berikut :

$$\begin{array}{r} 2304 = 12a + 21b1 + 765b2 + 2220000b3 \\ 2304 = 12a + 21*13,16002 + 765*4,74468 + 2220000*0,003015 \\ 2304 = 12a + (276,36048) + 3629,68723 + 6694,46328 \\ 2304 = 12a + 10600,511 \\ 12a = 2304 - 10600,511 \\ 12a = -8296,510 \end{array}$$

$$(-8296,510) / 12 = - 691,3759a$$

Dari hasil perhitungan a , b_1 , b_2 , dan b_3 diatas jika hasilnya dimasukan ke dalam persamaan berikut :

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3$$

Maka akan menghasilkan persamaan dibawah ini :

$$Y = - 691,3759 + 13,16002X_1 + 4,74468X_2 + 0,003015X_3$$

Pengujian Kasus :

PT. Pinem Lau Guna Medan ingin mengestimasi jumlah penumpang berdasarkan hari libur dalam sebulan, jumlah keberangkatan dan harga tiket untuk bulan Desember tahun 2020 dengan data yang telah ada dibulan sebelumnya yaitu jumlah hari libur dibulan Desember 2020 adalah 3, jumlah keberangkatan 69 dan harga tiket 185.000 maka hasilnya adalah sebagai berikut :

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3$$

$$Y = - 691,3759 + 13,16002*(3) + 4,74468*(69) + 0,0030155*(185.000)$$

$$Y = - 691,3759 + 39,480 + 327,383 + 557.871$$

$$Y = - 691,3759 + 924,735$$

$$Y = 233,359$$

Berdasarkan perhitungan di atas dapat diperoleh estimasi jumlah penumpang berdasarkan hari libur, jumlah keberangkatan, dan harga tiket pada PT. Pinem Lau Guna Medan pada bulan Agustus tahun 2020 adalah 233,359 Berikut hasil estimasinya :

Tabel 3.4 Hasil Estimasi Jumlah Penumpang Desember Tahun 2020

Bulan	Hari Libur (Dalam/Bln)	Jumlah Keberangkatan	Harga Tiket	Estimasi Jumlah Penumpang
Desember	3	69	Rp 185.000	233,359

4. PEMODELAN SISTEM DAN PERANCANGAN

Model sistem yang akan dibentuk menggunakan dua jenis pemodelan, yang pertama menggunakan UML dan kemudian menggunakan *flowchart* sistem. UML adalah salah satu alat bantu yang sangat handal didunia pengembangan sistem yang berorientasi objek.

1. *Use Case Diagram*
Use case diagram adalah pemodelan yang menggambarkan peranan pengguna pada sebuah sistem.
2. *Activity Diagram*
Activity diagram merupakan gambaran aliran kerja dari menu menu yang terdapat pada sebuah sistem.
3. *Class Diagram*
Class diagram merupakan gambaran aliran kerja pada struktur – struktur dalam membangun sebuah sistem.

5. PENGUJIAN DAN IMPLEMENTASI

Di dalam pengujian dan implementasi aplikasi mengestimasi jumlah penumpang pada PT. Pinem Lau Guna Medan membutuhkan beberapa fasilitas pendukung. Berikut ini merupakan fasilitas pendukung yang dibutuhkan oleh sistem, baik perangkat keras maupun perangkat lunak:

5.1.1 Perangkat Keras (Hardware)

Hardware merupakan komponen yang berbentuk perangkat keras yang akan digunakan di dalam membangun sistem. Dalam membangun system ini di butuhkan perangkat keras – perangkat keras yang dimaksud adalah sebagai berikut:

1. Komputer Prosesor minimal *Dual Core*
2. RAM Minimal (*Random Access Memory*) 2 GB
3. Hardisk Minimal 160 GB
4. Mouse
5. Keyboard
6. Printer

5.1.2 Perangkat Lunak (Software)

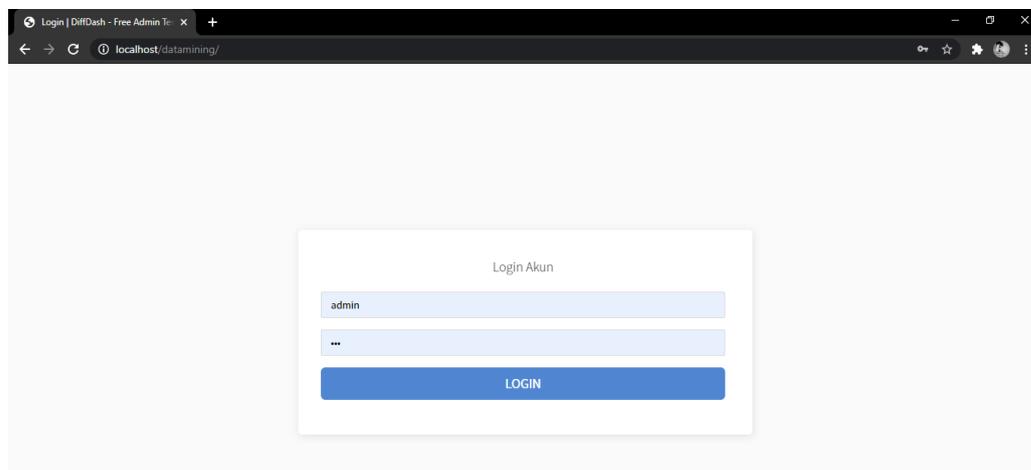
Software merupakan suatu komponen yang berbentuk perangkat lunak yang digunakan dalam membangun sistem. Didalam membangun sistem ini dibutuhkan juga perangkat lunak – perangkat lunak yang dimaksud adalah sebagai berikut :

- a. Sistem Operasi *Microsoft* Minimal Windows 8
- b. *Web Server XAMPP*
- c. *Sublime Text Editor*
- d. *Web Browser*

5.1 Implementasi Sistem

1. *From Login*

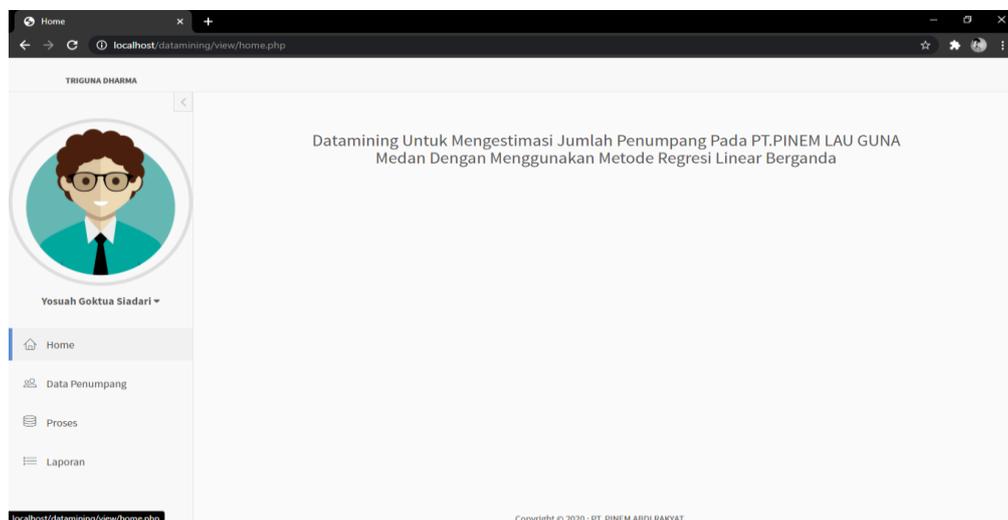
Login adalah suatu proses yang dilakukan oleh user untuk mengakses komputer dengan memasukkan identitas dari akun pengguna dan kata sandi guna untuk mendapatkan hak akses menggunakan suatu aplikasi.



Manfaat dari dibentuknya *form login* adalah untuk membatasi pengguna sistem agar tidak sembarangan orang bisa menjalankan system, oleh sebab itu sistem yang telah dibentuk hanya dapat dipergunakan seseorang yang sudah berwenang dan sekaligus juga sebagai pengaman sebelum memulai untuk menggunakan sistem ini.

2. *Menu Utama*

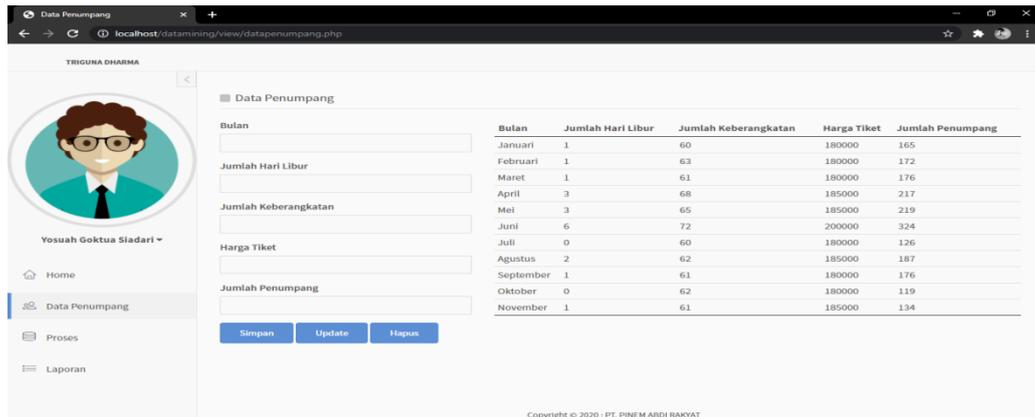
Menu Utama adalah tampilan awal dari setiap aplikasi setelah berhasil login dari form login, menu utama berfungsi untuk memanggil form lainnya. Berikut tampilan menu utama :



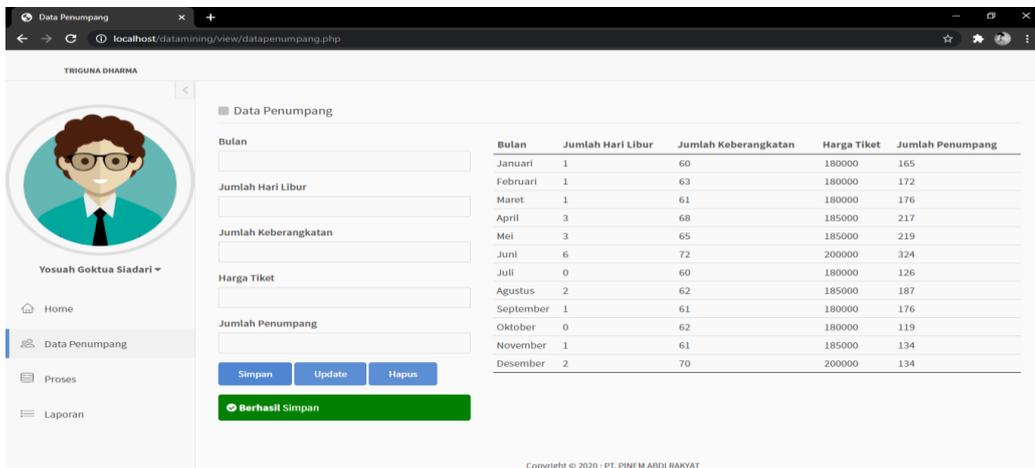
3. *Data Penumpang*

Data penumpang adalah informasi yang diperoleh setiap bulannya berdasarkan jumlah hari libur dalam bulan tersebut, jumlah unit yang diberangkatkan, harga tiket, dan jumlah penumpang sebelum keberangkatan.

Form data penumpang pada aplikasi yang dirancang dapat ditampilkan dengan cara memilih menu data penumpang. Didalam menu data penumpang terdapat form data penumpang, tombol simpan, update, dan hapus yang memiliki fungsi sesuai dengan sebagaimana fungsinya masing-masing. Berikut tampilan dari menu data menu penumpang.



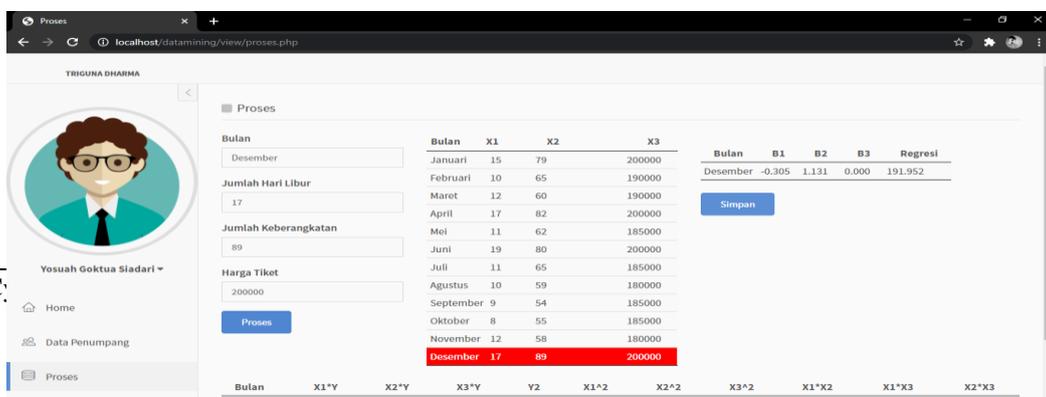
Setelah data penumpang di input kemudian klik tombol simpan maka akan ditampilkan pada tabel Data Penumpang. Berikut tampilannya setelah di input :



4. Proses Perhitungan Regresi Linear Berganda

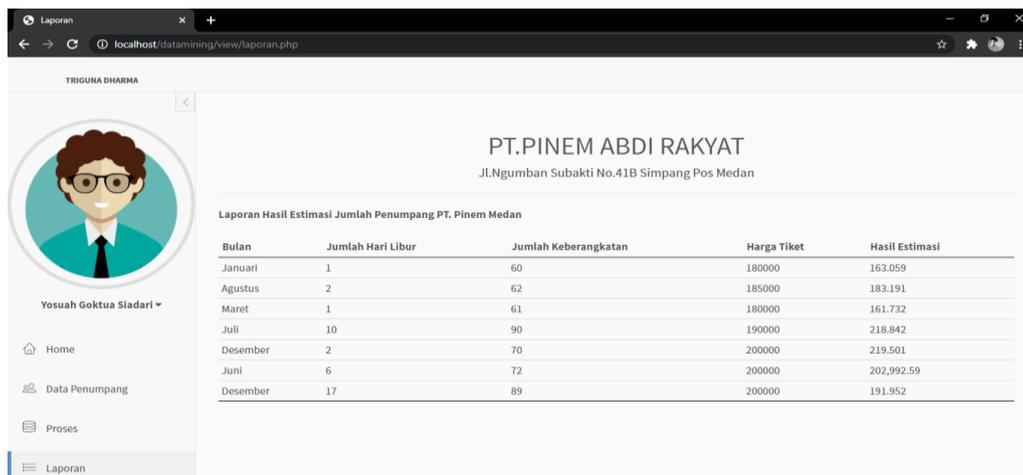
Perhitungan Regresi adalah suatu proses pengolahan data berdasarkan variabel yang telah ditentukan dan informasi masa lalu hingga sekarang yang dimiliki agar dapat menghasilkan informasi dalam mengestimasi jumlah penumpang yang lebih mendekati.

Form proses regresi pada aplikasi yang telah dirancang dapat ditampilkan dengan memilih menu proses. Di dalam *form* proses terdapat fitur tombol proses dan simpan yang berfungsi sebagaimana mestinya. Berikut tampilan menu *form* data hasil mengestimasi jumlah penumpang.



5.3 Pengujian

Setelah melakukan proses implementasi, proses selanjutnya adalah melakukan pengujian terhadap sistem yang telah dirancang. Pengujian ini untuk melihat bahwa hasil perancangan dan perhitungan yang ada di bab III sesuai dengan hasil yang ditampilkan pada sistem. Keluaran yang di hasilkan oleh sistem akan di sesuaikan dengan hasil perhitungan.



Bulan	Jumlah Hari Libur	Jumlah Keberangkatan	Harga Tiket	Hasil Estimasi
Januari	1	60	180000	163.059
Agustus	2	62	185000	183.191
Maret	1	61	180000	161.732
Juli	10	90	190000	218.842
Desember	2	70	200000	219.501
Juni	6	72	200000	202.992.59
Desember	17	89	200000	191.952

Adapun manfaat dari gambar 5.6 di atas ialah menampilkan hasil perhitungan dari Regresi Linear Berganda dan pada *form* laporan ini juga berfungsi untuk *print report* atau mencetak laporan hasil keputusan.

5.4 Kelemahan dan Kelebihan Sistem

Adapun kelemahan dan kelebihan sistem yaitu sebagai berikut :

5.4.1 Kelemahan Sistem

Kelemahan sistem yaitu hal apapun yang menjadi dasar perbaikan dari sistem yang telah dibuat, atau lebih tepatnya kekurangan dari sistem, yaitu :

1. Aplikasi yang dirancang hanya bisa digunakan untuk mengestimasi jumlah penumpang dalam suatu tujuan tertentu tidak dapat mengestimasi jumlah penumpang untuk berbagai tujuan.
2. Sistem ini masih tergantung pada variabel yang sudah ditentukan secara manual. Karena pada sistem ini hanya memiliki 3 variabel.
3. Sistem ini masih terbatas dalam menggunakannya, misalnya pihak terkait yang lain tidak dapat menggunakannya secara massal karena tidak bersifat online.
4. Aplikasi ini belum dilengkapi fitur keamanan data yang baik karena dibutuhkan lagi algoritma pengamanan data.

5.4.2 Kelebihan Sistem

Kelebihan sistem yaitu hal apapun yang menjadi keunggulan dari sistem yang telah dibuat, atau lebih tepatnya nilai tambah dari sistem, yaitu :

1. Sistem ini mampu untuk mengolah data penumpang tiap bulannya dengan cepat dan mudah dalam penggunaannya untuk jumlah data yang cukup banyak.
2. Aplikasi ini juga dapat digunakan ke berbagai perusahaan otobus lainnya dalam membantu perusahaan mengestimasi jumlah data penumpangnya, seperti yang akan digunakan pada PT. Pinem Lau Guna Medan.
3. Sistem ini mampu menampilkan hasil laporan secara otomatis tanpa harus di simpan secara manual.

6. KESIMPULAN DAN SARAN

Berdasarkan analisa pada permasalahan yang terjadi dalam kasus mengestimasi jumlah penumpang pada PT. Pinem Lau Guna Medan dengan menggunakan metode Regresi Linear Berganda, maka dapat ditarik kesimpulan sebagai berikut:

1. Adapun cara menganalisa permasalahan data penumpang pada tahun 2019 dengan cara mengumpulkan setiap data akurat yang terjadi setiap bulannya, kemudian data tersebut dikelompokkan dan selanjutnya dilakukan perhitungan regresi.

2. Untuk merancang sebuah sistem mengestimasi jumlah penumpang dengan teknik data mining menggunakan metode regresi linear berganda dapat menggunakan bahasa pemrograman berbasis website sehingga menjadi sebuah aplikasi estimasi yang dapat digunakan.
3. Untuk mengimplementasikan teknik data mining menggunakan metode regresi linear berganda dalam pemecahan masalah target jumlah penumpang adalah dengan cara menentukan Variabel Y, X1, X2, dan X3 terlebih dahulu, selanjutnya hitung nilai dari setiap persamaan, lalu hitung menggunakan rumus mencari nilai a, b1, b3 dan ditentukan saatnya pengujian sampel untuk menentukan estimasi target jumlah penumpang yang akan diketahui.

6.2 Saran

Untuk meningkatkan khasanah keilmuan berikut ini adalah saran dalam penelitian ini yaitu:

1. Sistem ini ditujukan kepada PT. Pinem Lau Guna Medan, sistem ini diharapkan bisa digunakan di berbagai perusahaan, instansi, lembaga dan pihak lainnya yang membutuhkan.
2. Kedepannya aplikasi tersebut dapat dikembangkan lebih baik lagi daripada saat ini guna mendukung proses estimasi yang lebih efektif, akurat, dan mudah dalam penggunaannya.
3. Peneliti berikutnya dapat menggunakan metode lain di bidang data mining sebagai studi komparasi berikutnya.

UCAPAN TERIMA KASIH

Pada kesempatan ini saya ucapkan terimakasih kepada Bapak, Ibu dan keluarga saya atas segala doa, semangat dan motivasinya. Selain itu, terimakasih sebesar-besarnya kepada semua pihak yang telah membantu untuk menyelesaikan penulisan skripsi ini, yaitu :

1. Bapak Rudi Gunawan, SE, M.Si, Selaku Ketua STMIK Triguna Dharma Medan.
2. Bapak Dr. Zulfian Azmi, ST, M.Kom Selaku Wakil Ketua I Bidang Akademik STMIK Triguna Dharma Medan.
3. Bapak Marsono. S.Kom, M.Kom, Selaku Ketua Program Studi Sistem Informasi STMIK Triguna Dharma Medan.
4. Bapak Darjat Saripurna, S.Kom, M.Kom selaku Dosen Pembimbing I yang membimbing dan menyediakan waktu selama ini.
5. Purwadi, S.Kom., M.Kom selaku Dosen Pembimbing II yang membimbing dan menyediakan waktu selama ini.
6. Seluruh Dosen, Staff dan Pegawai STMIK Triguna Dharma.
7. Terimakasih juga disampaikan kepada PT. Pinem Lau Guna Medan yang telah mengizinkan melakukan penelitian dan memberikan data yang benar sehingga skripsi ini dapat terselesaikan dengan baik.

Akhir kata saya ucapkan rasa terima kasih kepada semua pihak yang terlibat dalam penyelesaian skripsi ini Skripsi ini masih sangat jauh dari sempurna. Oleh karena itu, diharapkan saran dan kritik yang sifatnya membangun dari para pembaca demi kesempurnaan skripsi ini.

REFERENSI

- [1] R. Sistem, "Jurnal resti," vol. 2, no. 1, pp. 361–366, 2018.
- [2] I. Lestari, L. Gaol, S. Sinurat, and E. R. Siagian, "IMPLEMENTASI DATA MINING DENGAN METODE REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI DATA PERSEDIAAN BUKU PADA PT . YUDHISTIRA," vol. 3, pp. 130–133, 2019.
- [3] W. Sulistiyowati, P. Studi, T. Industri, F. Teknik, U. M. Sidoarjo, A. Regresi, L. Jumlah, K. Mesin, and H. B. Baku, "Peramalan Produksi dengan Metode Regresi Linier Berganda," vol. 1, no. 2, pp. 82–89, 2017.
- [4] W. R. Maya, "Penerapan Data Mining Dalam Memprediksi Jumlah Penumpang Pada CV . Surya Mandiri Sukses Dengan Menggunakan Metode Regresi Linier," vol. 2, no. 1, pp. 54–61, 2019.
- [5] D. Sari, O. Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," vol. 7, no. 1, pp. 56–62, 2020.
- [7] A. Zanuardi and H. Suprayitno, "Analisa Karakteristik Kecelakaan Lalu Lintas di Jalan Ahmad Yani Surabaya melalui Pendekatan Knowledge Discovery in Database," vol. 2, no. 1, pp. 45–55, 2018.
- [8] P. Bidang, K. Sains, Y. Mardi, J. Gajah, M. No, and S. Barat, "Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika."
- [9] P. S. Ramadhan and N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang," vol. 18, no. 1, pp. 55–61, 2019.

	<p>Data Diri Nama : Yosuah Goktua Siadari Tempat/Tanggal Lahir : Medan, 10 Januari 1998 Jenis Kelamin : Laki-laki Agama : Katolik Status : Belum Menikah Pendidikan Terakhir : Sekolah Menengah Atas Kewarganegaraan : Indonesia E-mail : yosuavtxsiadari@gmail.com</p> <p>Pendidikan Formal 1. Tahun 2004 - 2010 : SD NEGERI NO. 066656 2. Tahun 2010 - 2013 : SMP SWASTA BAKTI MULIA ONANRUNGU 3. Tahun 2013 - 2016 : SMA SWASTA SANTO MIKHAEL PANGURURAN</p>
	<p>Darjat Saripurna, S.Kom., M.Kom</p> <p>Dosen pengajar tetap STM IK TRIGUNA DHARMA</p>
	<p>Purwadi.,S.Kom.,M.Kom</p> <p>Dosen pengajar tetap STM IK TRIGUNA DHARMA</p>