

Data Mining Untuk Mengestimasi Jumlah Penumpang Pada PT Rajawali Citra Transportasi Medan Dengan Menggunakan Metode Regresi Linear Berganda

Edi Guna Bangun *, YOPI HENDRO SYAHPUTRA**,
RICO IMANTA GINTING **

* Sistem Informasi, STMIK Triguna Dharma

** Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received Jun 12th, 201x

Revised Aug 20th, 201x

Accepted Aug 26th, 201x

Keyword:

Data Mining

Estimasi Jumlah Penumpang

Regresi Linear Berganda

ABSTRACT

Data mining merupakan proses analisa data untuk menemukan suatu pola dari kumpulan data tersebut. Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pengambilan keputusan. Dalam skripsi ini yang akan dibahas adalah mengenai prediksi jumlah penumpang bus. Salah satu teknik yang ada pada data mining adalah estimasi. Teknik estimasi yang akan digunakan adalah Regresi Linear Berganda metode ini dipilih karena mampu membuat suatu prediksi dengan memanfaatkan data-data lama mengenai jumlah penumpang. Metode Regresi Linear Berganda bertujuan untuk merumuskan persamaan regresi dan prediksi terbaik berdasarkan atribut atribut yang ada.

Copyright © 2020 STMIK Triguna Dharma.
All rights reserved.

Corresponding Author:

Nama : Edi Guna Bangun

Program Studi : Sistem Informasi

Kampus : STMIK Triguna Dharma

Email : Edigunabangun227@gmail.com

1. PENDAHULUAN

perusahaan jasa transportasi khususnya bus belom-bela memberikan pelayanan terbaik di mulai dari fasilitas bus yang digunakan perusahaan, serta akses yang cepat dan mudah untuk di berikan kepada konsumen sehingga pengguna jasa sendiri dapat tertarik bahkan sudah menjadi pelanggan tetap. Sistem prasarana transportasi mempunyai ciri utama yaitu membantu pengguna di mana dalam pemilihan dan pengguna jenis moda transportasi perlu di persiapkan tempat moda bergerak seperti jalan raya, rel, kereta api, pelabuhan dan bandar udara. Secara umum moda transportasi untuk angkutan penumpang di bedakan menjadi dua yaitu yaitu kelompok yang memiliki pilihan dalam melakukan Mobilitas akses kendaraan secara pribadi (*choice*) sedangkan kelompok kedua yaitu kelompok yang tergantung pada angkutan umum untuk melakukan pergerakan dengan menggunakan angkutan umum di definisikan sebagai orang berangkat dari rumah tidak mempunyai kendaraan pribadi atau mempunyai pilihan lain kecuali angkutan umum (*captive*).

2. METODOLOGI PENELITIAN

2.1. Transportasi

Transportasi merupakan proses memindahkan manusia atau barang dengan menggunakan sarana yang di gerakkan oleh mesin atau manusia. Transportasi di gunakan untuk mempermudah manusia untuk melakukan aktivitas sehari hari. Transportasi bukan hanya usaha berupa gerakan manusia dan barang dari suatu tempat ke tempat lain dengan gerakan secara statis akan tetapi transportasi akan mengalami perkembangan dan kemajuan dari waktu ke waktu baik sarana dan prasarannya sesuai dengan perkembangan ilmu pengetahuan dan teknologi Transportasi adalah salah satu kemudahan bagi suatu daerah untuk maju dan berkembang serta transportasi dapat meningkatkan perhubungan ke masing masing daerah. Transportasi faktor sangat penting untuk melayani angkutan barang dan manusia dari satu daerah ke daerah lainnya dan menjunjung pengembangan kegiatan-kegiatan sektor lain untuk meningkatkan pembangunan daerah tersebut [1].

2.2. Flowchart Metode Regresi Linear Berganda



2.3. Data Mining

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database[2] yang meliputi pengumpulan, pemakaian data historis untuk menemukan pola atau hubungan dalam data berukuran besar yang mencakup *database* atau informasi lainnya[3]. Data mining yang berisi pencarian pola yang di inginkan pada *database* untuk membantu pengambilan keputusan di waktu yang akan datang [4]. Estimasi adalah suatu metode dimana dapat memperkirakan nilai populasi dengan memakai nilai sampel [5].

2.4. Metode Regresi Linear Berganda

Regresi linear berganda merupakan satu cara prediksi yang menggunakan garis lurus untuk menggambarkan hubungan diantara dua variable atau lebih. Variabel adalah besaran yang berubah-ubah nilainya selanjutnya variable tersebut terbagi atas dua variabel jenis yaitu variabel pemberi pengaruh dan variabel terpengaruh[6].

Ada beberapa tahapan dalam membuat penyelesaian masalah dengan metode regresi linear berganda yaitu

1. Mempersiapkan data *training*. Data *training* biasanya di ambil dari suatu data historis yang pernah terjadi sebelum atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu

2. Bentuk umum regresi linear berganda sebagai berikut :

$$Y = a + b_1X_1 + b_2X_2$$

3. Menentukan nilai konstanta dan koefisien linear berganda

Dalam kasus ini akan diambil dari data yang ada pada PT Rajawali Citra Transportasi selama bulan Januari – Desember. Data tersebut akan digunakan dalam perhitungan metode regresi linear berganda untuk menganalisa mengenai mengestimasi jumlah penumpang. Data penumpang tersebut selanjutnya akan dilakukan proses untuk menghasilkan data yang siap di bentuk menjadi sebuah informasi baru

2.4 Data Penelitian

Berdasarkan hasil kegiatan pengumpulan data melalui observasi dan wawancara, berikut merupakan data yang diperoleh dari perusahaan terkait yaitu :

Tabel 1 Data Bulan Januari sampai Desember

PT Rajawali Citra Transport 2019				
Bulan	Hari Libur (Dalam/Bln)	Jumlah Keberangkatan	Harga Tiket	Jumlah Penumpang
	X1	X2	X3	Y
Januari	1	122	38,000	4020
Februari	1	123	38,000	4025
Maret	1	122	38,000	4011
April	2	124	38,000	4205
Mei	2	124	38,000	4101
Juni	6	131	38,000	4350
Juli	0	120	43,000	3120
Agustus	1	121	38,000	4100
September	0	119	38,000	3090
Oktober	0	121	38,000	3129
November	1	122	38,000	4014
Desember	2	123	38,000	4210
Total	17	1472	461,000	46375

Tabel 2 Jenis Variabel yang Digunakan

Variabel Yang Digunakan		
NO	Variabel	Keterangan
1	X1	Jumlah Hari Libur Dalam Sebulan
2	X2	Jumlah Keberangkatan
3	X3	Harga Tiket
4	Y	Jumlah Penumpang

Tabel 3 Koefisien Regresi

X1*Y	X2*Y	X3*Y	Y ²	X1 ²
4020	490440	152760000	16160400	1
4025	495075	152950000	16200625	1
4011	489342	152418000	16088121	1
8410	521420	159790000	17682025	4
8202	508524	155838000	16818201	4
26100	569850	165300000	18922500	36
0	374400	134160000	9734400	0
4100	496100	155800000	16810000	1
0	367710	117420000	9548100	0
0	378609	118902000	9790641	0
4014	489708	152532000	16112196	1
8420	517830	159980000	17724100	4
71302	5699008	1777850000	181591309	53

Tabel 3 Koefisien Regresi (Lanjutan)

X2 ²	X3 ²	X1.X2	X1.X3	X2.X3
14884	1444000000	122	38000	4636000
15129	1444000000	123	38000	4674000
14884	1444000000	122	38000	4636000
15376	1444000000	248	76000	4712000
15376	1444000000	248	76000	4712000
17161	1444000000	786	228000	4978000
14400	1849000000	0	0	5160000
14641	1444000000	121	38000	4598000
14161	1444000000	0	0	4522000
14641	1444000000	0	0	4598000
14884	1444000000	122	38000	4636000
15129	1444000000	246	76000	4674000
180666	17733000000	2138	646000	56536000

Tabel 4 Koefisien Regresi Hasil

y	A	b1	b2	b3
46375	12	17	1472	461,000
71302	17	53	2138	646,000
5699008	1472	2138	180666	56536000
1777850000	461,000	646000	56536000	17733000000

Untuk memperoleh koefisien regresi a, b_1 , dan b_2 dapat di peroleh dengan cara eliminasi tiga persamaan sebagai berikut

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 \dots \dots \dots (\text{Persamaan 1})$$

$$\sum X_1 Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + b_3 \sum X_1 X_3 \dots \dots \dots (\text{Persamaan 2})$$

$$\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 X_3 \dots \dots \dots (\text{Persamaan 3})$$

Kemudian masukkan angka yang telah di dapat pada Koefisien regresi hasil perhitungan (Tabel 4) dan nilai $\sum X_1, \sum X_2, \sum Y$ (dari tabel 3) sehingga di peroleh hasil persamaan 1,2 dan 3.

$$67249 = -347b_1 + -632b_2 + 85000b_3 \dots\dots\dots(\text{Persamaan 1})$$

$$-124096 = -632b_1 + -1208b_2 + 160000b_3 \dots\dots\dots(\text{Persamaan 2})$$

$$4675000 = 85000b_1 + 160000b_2 + -275000000b_3 \dots\dots\dots(\text{Persamaan 3})$$

Kemudian tiga persamaan diatas diselesaikan hingga memperoleh nilai a, b₁, dan b₂. Langkah-langkah penyelesaiannya adalah sebagai berikut :

1. Melakukan proses eliminasi antara persamaan [1] dengan [2]

$$-67249 \text{ a} = -347b_1 + -632b_2 + 85000b_3$$

$$-124096 \text{ a} = -632b_1 + -1208b_2 + 160000b_3$$

$$42501368 = 21930b_1 + 399424b_2 + -537200000b_3$$

$$43061312 = 219304b_1 + 419176b_2 + -555200000b_3$$

$$-559944 = 0 + -19752b_2 + 1800000b_3$$

Persamaan [4]

2. Kemudian melakukan proses eliminasi antara persamaan [1] dengan [3]

$$-67249 \text{ a} = -347b_1 + -632b_2 + 85000b_3$$

$$44675000 \text{ a} = 85000b_1 + 160000b_2 + -275000000b_3$$

$$-5716165000 = -29495000b_1 + -53720000b_2 + 7225000000b_3$$

$$-15502225000 = -29495000b_1 + -5550000b_2 + 95425000000b_3$$

$$9786060000 = 0 + 1800000b_2 + -88200000000b_3$$

Persamaan [5]

3. Kemudian melakukan proses eliminasi antara persamaan [4] dengan [5]

$$-559944 = -19752b_2 + 1800000b_3$$

$$9786060000 = 1800000b_2 + -88200000000b_3$$

$$-1.0079\text{E}+12 = -35553600000b_2 + 3.24\text{E}+12b_3$$

$$-1,932286\text{E}+14 = -35553600000b_2 + 1.74213\text{E}+15b_3$$

$$1.92286\text{E}+14 = 0 + -1.73889\text{E}+15b_3$$

$$b_3 = 1.92286\text{E}+14 : -1.73889\text{E}+15$$

$$= \mathbf{-0.110580172b_3}$$

4. Kemudian untuk mendapatkan b₂ maka hasil dari b₃ dimasukkan ke persamaan [4]

$$-559944 = -19752b_2 + 1800000b_3$$

$$-559944 = -19752b_2 + 1800000b_3 * (-0.110580172b_3)$$

$$-55994 = -19752b_2 + -199044.3103$$

$$-19752b_2 = (-55994) - (-199044.3103)$$

$$-19752b_2 = -360899.6897$$

$$(-360899.6897) / (-19752) = \mathbf{18.27155172b_2}$$

5. Untuk menghasilkan nilai b_1 maka b_2 dan b_3 dimasukkan ke persamaan [3] dengan mensubstitusikan

$$44675000 = 85000b_1 + 160,000b_2 + -275000000b_3$$

$$44675000 = 85000b_1 + 160,000b_2 * (18.27155172) + -275000000b_3 * (-0.110580172)$$

$$44675000 = 85000b_1 + (2923448.276) + (30409547.41)$$

$$85,000 = 44675000b_1 - (33332995.69)$$

$$85,000 = 11342004.31$$

$$(11342004.31) / (85,000) = \mathbf{133.4353448b_1}$$

6. Untuk menghasilkan nilai a maka b_1 , b_2 dan b_3 dimasukkan ke persamaan [1] dengan mensubstitusikan

$$46375 = 12a + 17b_1 + 1472b_2 + 485000b_3$$

$$46375 = 12a + 17b_1 * (133.4353448) + 1472b_2 * (18.27155172) + 461,000 b_3 * (-0.110580172)$$

$$46375 = 12a + (2268.400862) + (-26895.72414) + (-50977.45948)$$

$$12a = (46375) - (-21813.33448)$$

$$12a = 68188.33448$$

$$(68188.33448) / (12) = \mathbf{5682.361207 a}$$

Sekarang telah di dapat nilai a , b_1 , b_2 dan b_3 dimana nilai-nilainya adalah sebagai berikut :

$$a = 5682.361207$$

$$b_1 = 133.4353448$$

$$b_2 = 18.27155172$$

$$b_3 = -0.110580172$$

Pengujian kasus:

PT. Rajawali Citra Transportasi ingin mengestimasi jumlah penumpang berdasarkan hari libur dalam setahun dan jumlah keberangkatan untuk tahun 2020 dengan data yang ada di tahun-tahun sebelumnya yaitu hari libur setahun 19, jumlah keberangkatan 1480 dan harga tiket 45000

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

$$Y = 5682.361207 + 133.4353448(19) + 18.27155172(1480) + -0.110580172 (45000)$$

$$Y = 5682.361207 + 24509.70259$$

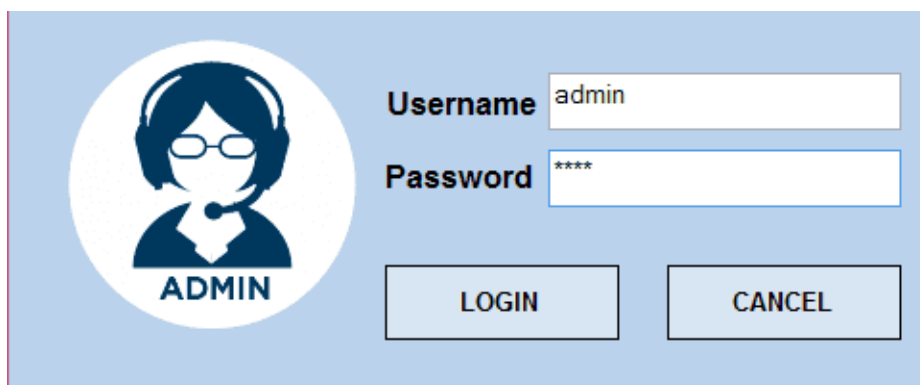
$$Y = 30192.06379$$

Berdasarkan perhitungan diatas dapat diperoleh estimasi jumlah penumpang berdasarkan hari libur dan jumlah keberangkatan pada PT Rajawali Citra Transportasi tahun 2020 adalah 30192.06379

3. ANALISA DAN HASIL

Kegiatan akhir dalam penerapan sistem adalah proses implementasi sistem, yaitu mengoperasikan sistem secara menyeluruh. Dibawah ini merupakan tampilan implementasi sistem dari sistem yang telah diuji sebelumnya yaitu sebagai berikut :

Tampilan Halaman *Login*



Gambar 4.1 Tampilan Halaman *Login*



Gambar 4.2 Tampilan Halaman Utama

Tampilan *Form* Perhitungan Hasil Estimasi

 The image shows a screenshot of a web application form for regression calculation. It includes input fields for "Bulan" (Month) set to "Januari", "Kode Sampel" (Sample Code) set to "P015", "Hari Libur/Bylan (X1)" (Holiday/Bylan) set to "2", "Jumlah Keberangkatan (X2)" (Number of Departures) set to "130", "Harga Tiket (X3)" (Ticket Price) set to "45000", and "Jumlah Penumpang (Y)" (Number of Passengers) set to "3348". A box on the right displays the "Persamaan Regresi" (Regression Equation): $Y = 5682.36120689655 + (133.435344827586)X1 + (18.2715517241379)X2 + (-0.110580172413793)X3$. Below the inputs are buttons for "PROSES", "BERSIH", "CETAK", and "KELUAR". A "Hasil Estimasi" (Estimation Result) field shows the value "3348".
Gambar 4.3 Tampilan Halaman *Regresi*

Tampilan Halaman Laporan

 The image shows a screenshot of a report page. At the top, it says "PT. RAJAWALI CITRA TRANSPORTASI" and "Jl. Ngemban Surbakti No. 30 Medan, Sumatera Utara". Below this is the title "Laporan Estimasi Jumlah Penumpang Bus". The main content is a table with 6 columns: "Bulan", "Hari Libur", "Jlh Keberangkatan", "Harga Tiket", "Jumlah Penumpang", and "Estimasi". The table contains one row of data. Below the table, there is a date "Medan, 26 July 2020" and a signature line "Diketahui Oleh,".

Bulan	Hari Libur	Jlh Keberangkatan	Harga Tiket	Jumlah Penumpang	Estimasi
Januari	2	130	45,000.00	3,348	3,348

Gambar 4.4 Tampilan Halaman Laporan

4. KESIMPULAN

Berdasarkan hasil analisa dan pembahasan data *Mining* dengan teknik estimasi menggunakan metode regresi linear berganda tentang prediksi jumlah penumpang maka di tariklah kesimpulan bahwa. Hasil dari analisa yang di peroleh dari data mining dengan metode regresi linear berganda mengenai prediksi jumlah penumpang dapat PT Rajawali Citra Transportasi dapat membantu pihak perusahaan mengetahui tentang kriteria apa saja yang mempengaruhi laju pertumbuhan penduduk. Dan juga di temukan pola yang saling berkaitan erat antara hari libur dengan harga tiket yang dapat mempengaruhi jumlah penumpang

UCAPAN TERIMA KASIH

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 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 saya ucapkan 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 Yopi Hendro Syahputra S.T.,M.KOM., selaku Dosen Pembimbing I dan Bapak Rico Imanta Ginting S.KOM.,M.KOM., selaku Dosen Pembimbing II yang telah memberikan saran, arahan dan dukungannya serta motivasi, sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Seluruh Dosen, di STMIK Triguna Dharma Medan.

REFERENSI

- [1] S. Gabungan *et al.*, *Riwayat Penulis*. 2016.
- [2] 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.
- [3] N. Variabel, P. Terhadap, A. R. Muhajir, E. Sutoyo, and I. Darmawan, "Forecasting Model Penyakit Demam Berdarah Dengue Di Provinsi DKI Jakarta Menggunakan Algoritma Regresi Linier Untuk Mengetahui Kecenderungan Forecasting Model Penyakit Demam Berdarah Dengue Di Provinsi DKI Jakarta Menggunakan Algoritma Regresi Linier Untuk Mengetahui Kecenderungan Nilai Variabel Prediktor Terhadap Peningkatan Kasus," no. November, 2019, doi: 10.21111/fij.v4i2.3199.
- [4] R. Sistem, "Jurnal resti," vol. 2, no. 1, pp. 282–287, 2018.
- [5] R. Tinggal, P. Pt, and S. Jaya, "Regresi Linear Dalam Mengestimasi Jumlah Pemakaian Bahan Baku Dan Tenaga Pekerja Dalam Pembangunan," pp. 67–78.
- [6] 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.

BILIOGRAFI PENULIS

	Edi Guna Bangun, Laki-laki kelahiran, Namoukur 27 Oktober 1997. Anak pertama dari 3 bersaudara, dan merupakan seorang mahasiswa STMIK Triguna Dharma, Medan. Saat ini sedang dalam proses menyelesaikan skripsi.
	Yopi Hendro Syahputra S.T., M.Kom Beliau merupakan dosen tetap STMIK Triguna Dharma, serta aktif sebagai dosen pengajar khusus pada bidang ilmu Sistem Informasi.
	Rico Imanta Ginting, S.Kom., M.Kom Beliau merupakan dosen tetap di STMIK Triguna Dharma serta aktif sebagai dosen pengajar khusus di bidang ilmu Sistem Informasi.