

IMPLEMENTASI DATA MINING MENGGUNAKAN MULTI REGRESI UNTUK MEMREDIKSI RERATA KEDATANGAN MASYARAKAT

Dwi marisa Efendi¹, Ferly Ardhy², Riski Oskar Pratama³

STMIK Dian Cipta Cendikia Kotabumi¹²

Jl.Negara No.03 Candimas Kotabumi

E-mail : dwimarisa@dcc.ac.id, Ferly@dcc.ac.id², riski@gmail.com³

ABSTRAK

Dalam penelitian, ini dilakukan analisa dengan menggunakan metode regresi linier, berganda untuk memprediksi, angka kedatangan masyarakat pada kantor disduk capil lampung utara dengan menggunakan Microsoft, Excel 2010 dan aplikasi Rapid miner, dalam penelitian ini terdapat variabel independen, bebas dan variable independen, terikat yaitu variable independen, bebas yaitu : stok blangko, Operator, dan waktu pengerjaan. Sedangkan variable independen, terikat yaitu : jumlah masyarakat yang datang. Prediksi angka kedatangan masyarakat pada kantor disduk capil dengan mengambil sampel $x_1 = 795$ keping orang, $x_2 = 3$ orang, $x_3 = 7$ jam, dan $Y = 296$ orang menunjukkan hasil prediksi $-92,2484452$ menunjukkan nilai eror data $0,548$ untuk angka kedatangan masyarakat

Kata kunci : regresi, Rapid miner, Prediksi, masyarakat, analisa.

ABSTRACTS

This research uses multiple linear regression methods. research is useful to predicate the average number of people's arrival. this research uses Microsoft Excel 2010 and Rapid miner applications. The variables used are independent variables, free: stock blangko, operator, and working time. predicted to use sample $x_1 = 795$, $x_2 = 3$ people, $x_3 = 7$ hours, and $y = 296$ people. Prediction result is -92.2484452 , error value data 0.548 and 0.548 for public arrivals.

Keywords: regression, Rapid miner, Prediction, community, analysis.

1. PENDAHULUAN

Dinas kependudukan, dan pencatatan sipil merupakan instansi pelayanan penerbitan dokumen kependudukan seperti KK, KTP, Akta-Akta. Oleh karena itu instansi ini selalu ramai dikunjungi masyarakat, untuk mengurus dokumen kependudukan. Setiap hari instansi ini dikunjungi Oleh, masyarakat dengan jumlah berbeda, beda. Sering terjadinya, kekurangan, stok blangko sehingga menyebabkan masyarakat menjadi terhamabat di saat pengurusan dokumen, kependudukan, dan harus menunggu, sampai, tersedianya kembali blangko

tersebut.

Membeludaknya antrian masyarakat yang, di akibatkan, stok blangko yang tidak stabil. Belum ditemukan, metode untuk memprediksi, angka kedatangan masyarakat.

Adapun tujuan dari penelitian adalah:

Untuk mengetahui, dan menganalisis angka kedatangan, masyarakat perhari pada kantor Disduk Capil Lampung Utara dengan menggunakan metode regresi. Penulis dapat memberikan referensi, dalam

mengimplementasikan, Rapidminer dalam, menentukan alokasi jumlah blangko pada kantor Disduk Capil Lampung Utara.

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 diperkecil[1] Analisis regresi, merupakan metode sederhana untuk melakukan ,investigasi, tentang hubungan fungsional, di antara beberapa variable[2] penerapan data mining juga di implementasikan dengan metode yang berbeda yaitu Penerapan Algoritma, C5.0 untuk mengetahui permintaan produk kartu, perdana internet pada Vidha Ponsel.[3]

2. METODE PENELITIAN

Pada tahapan ini penulis mengelompokkan data-data yang didapat .Pengelompokkan dimulai dengan menentukan variable independent dan ariabledependen. Adapun variable-variable tersebut terdiri atas sebagai berikut [1].

1. Variable bebas (variable independen) : variable ini menjelaskan dan mempengaruhi variable lain yaitu : stok blangko (X_1) ,operator(X_2),waktu pengerjaan (X_3),
2. Variabel Terikat (variable dependen) : variable dipengaruhi oleh variable independent yaitu jumlah masyarakat yang mendaftar (Y).

Berikut adalah tahapan dalam pengolahan data regresi linear berganda :

1. Menyiapkan data prediksi menggunakan 1 variabel terikat (y) dan 3 variabel bebas.
2. variabel terikat : Y = Jumlah masyarakat yang mendaftar.
3. variabel bebas : x_1 =stok blangko , x_2 =operator , x_3 =waktu pengerjaan,
4. Kemudian mulai melakukan perhitungan untuk memangkatkan dua, X_1 , X_2 , X_3 , X_4 dan Y.
5. Memangkatkan dua variable bebas dan terikat.(X_{12} , X_{22} , X_{32})
6. Kemudian mengalikan antara variable bebas

dengan variable terikat ($X_1.Y$, $X_2.Y$, $X_3.Y$.)

7. Kemudian menghitung dengan mengalikan antara variable bebas dengan
8. variable bebas lainnya ($X_1.X_2$, $X_1.X_3$, $X_2.X_1$, $X_2.X_3$., $X_3.X_1$, $X_3.X_2$.)
9. Rumus persamaan regresi linear berganda 3 variabel yaitu :

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

Keterangan:

Y = Jumlah masyarakat yang daftar

a = Konstanta

X_1 = blangko

b_1 = Koefisien 1

X_2 = operator

b_2 = Koefisien 2

X_3 = waktu pengerjaan

b_3 = Koefisien 3

10. Kemudian tahapan yang terakhir adalah mencari nilai eror dari data yang telah di dapat dari perhitungan sebelumnya.

$$\text{Rumus} = \frac{(Y_i + y_i) - n}{n} \times 100\%$$

Keterangan:

Y_i = Data awal sebenarnya (Data sebenarnya)

y_i = Data akhir (Data hasil estimasi)

n = Jumlah data prediksi

3. HASIL DAN PEMBAHASAN

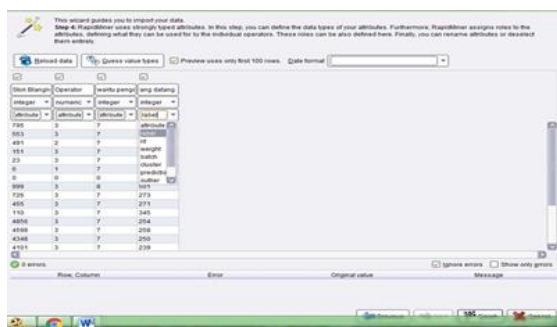
Berikut proses perhitungan angka kedatangan prediksi kedatangan (Studi Kasus Kantor Disdukcapil Lampung Utara) dengan Metode regresi linier berganda dengan menggunakan Microsoft excel 2010. Dengan informasi dari data pembelajaran tahun 2018 s/d 2020 sebagai berikut

Prediksi angka kedatangan

- a) Jumlah masyarakat yang datang tertinggi= 1184 dan kedatangan terendah =0 masyarakat perhari
- b) Jumlah stok blangko paling banyak =7758 dan stok blangko paling sedikit =0
- c) Jumlah operator terbanyak =3 dan operator paling sedikit = 0
- d) Waktu pengerjaan paling lama = 12 jam

dan waktu pengerjaan paling sedikit = 0 jamkuantitatif. Hasil percobaan sebaiknya ditampilkan dalam berupa grafik ataupun tabel.

yaitu prediksi angka kedatangan yang akan dijadikan label



Gamabr 1 tampilan fatribute data uji menjadi tipe data

Selanjutnya memilih metode yaitu metode regresi linier. Kemudian hubungkan antara port-port linear regression, read excel 2 dengan apply model dan hubungkan juga ke result untuk melihat hasil prediksi dari data yang sudah kita buat

Attribute	Coefficient	Std. Error	Std. Coeffi...	Tolerance	t-Stat	p-Value	Code
Operator	108.834	38.825	40.289	0.238	2.803	0.012	**
waktu pengerjaan	-2.806	16.395	-0.917	0.238	-0.171	0.868	
(Intercept)	-29.824	51.908	?	?	-0.575	0.577	

Gambar 2 . Tampilan dari hasil Nilai Error.

Berdasarkan hasil penelitian diperoleh dari hasil prediksi pada variabel terhadap jumlah kedatangan masyarakat mengacu pada tahun 2018 sampai 2020 dengan di lakukan 451 kali prediksi menunjukkan bahwa variabel yang digunakan system berpengaruh terhadap system prediksi angka kedatangan masyarakat. pada perhitungan mengguankan metode regresi linier dalam memprediksi angka kedatangan masyarakat dapat digunakan kantor disdukcapil dalam

mengambil sebuah keputusan dengan nilai prediksi yang mendekati data hasil olahan sebagai data pembelajaran pada prediksi angka kedatangan masyarakat pada kantor disdukcapil dengan mengambil sampel $x_1 = 795$ keping orang , $x_2 = 3$ orang , $x_3 = 7$ jam , dan $Y = 296$ orangmenunjukkan hasil prediksi $-92,2484452$ menunjukan nilai eror data $0,548$ untuk angka kedatangan masyarakat.pada kantor disduk capil lampung utara

4. KESIMPULAN

Dari hasil yang telah dilakukan dapat disimpulkan bahwa :

prediksi angka kedatangan masyarakat pada kantor disdukcapil dengan mengambil sampel $x_1 = 795$ keping belangko , $x_2 = 3$ orang , $x_3 = 7$ jam , dan $Y = 296$ orangmenunjukkan hasil prediksi $-92,2484452$ menunjukan nilai eror data $0,548$ untuk angka kedatangan masyarakat.pada kantor disdukcapil lampung utara.dapat di terima sebagai data pembelajaran untuk prediksi kedepanya

[1] Hutabarat, Carlis. "Penerapan Data Mining, Untuk Memprediksi Permintaan Produk, Kartu Perdana Internet Menggunakan, Algoritma C5. 0 (Studi Kasus: Vidha Ponsel)." Pelita Informatika: Informasi dan Informatika 6.4 (2018): 168-173.

[2] Sembel, Tesyalom., Ronny Gosal, and Sofia Pangemanan. "Partisipasi Masyarakat Dalam Pembangunan Desa (Studi di Desa Singsingon Barat Kecamatan Passi Timur Kabupaten Bolaang Mongondow)." JURNAL EKSEKUTIF 1.1 (2017).

[3] Sulistio, Andreas. "Pengaruh Promotion Mix Terhadap Brand Awareness Pt Hm Sampoerna Tbk Melalui Produk Terbaru Sampoerna A Platinum." Jurnal Strategi Pemasaran 5.2 (2018).

[4] Negara, Dengan Menganalisis Tingkat Dystopia Suatu. "Prediksi Tingkat Kehidupan Masyarakat Di

Mancanegara."

- [5] Syahra, Yohanni, Ismawardi Santoso, And Rini Kustini. "Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi." Seminar Nasional Sains Dan Teknologi Informasi (Sensasi). Vol. 2. No. 1. 2019.
- [6] Gaol, Indah Lestari Lumban, Sinar Sinurat, And Edward Robinson Siagian. "Implementasi Data Mining Dengan Metode Regresi Linear Berganda Untuk Memprediksi Data Persediaan Buku Pada Pt. Yudhistira Ghalia Indonesia Area Sumatera Utara." Komik (Konferensi Nasional Teknologi Informasi Dan Komputer) 3.1 (2019).
- [7] Gusman, Aggy Pramana. "Penerapan Fuzzy Time Series Dalam Menganalisa Peramalan Kedatangan Wisatawan Setelah Diterapkan Konsep Halal Tourism Sumatera Barat." Seminar Nasional Teknologi Informasi Komunikasi Dan Industri. 2018.
- [8] Wulandari, Ni Luh Putu, Et Al. "Prediksi Jumlah Pelanggan Dan Persediaan Barang Menggunakan Metode Regresi Linier Berganda Pada Bali Orchid." Josinfo: Jurnal Online Sistem Informasi 1.1 (2015).
- [9] Budiman, Irwan, And Artesya Nanda Akhlakulkarimah. "Aplikasi Data Mining Menggunakan Multiple Linear Regression Untuk Pengenalan Pola Curah Hujan." Klik-Kumpulan Jurnal Ilmu Komputer 2.1 (2016): 34-33.
- [10] Gunawan, Rudi. "Implementasi Data Mining Untuk Memprediksi Prestasi Siswa Berdasarkan Status Sosial Dan Kedisiplinan Pada Smk Bayu Pertiwi Menggunakan Metode Regresi Linier Berganda." Jurnal Saintikom (Jurnal Sains Manajemen Informatika Dan Komputer) 17.2 (2018): 175-183
- [11] Fikri, Ali. "Penerapan Data Mining Untuk Mengetahui Tingkat Kekuatan Beton Yang Dihasilkan Dengan Metode Estimasi Menggunakan Linear Regression." Universitas Dian Nuswantoro, Semarang (2013).
- [12] Nikmatun, Inna Alvi, And Indra Waspada. "Implementasi Data Mining Untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor." Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer 10.2 (2019): 421-432.
- [13] Yusuf, Ahmad, Hari Ginardi, And Isye Arieshanti. "Pengembangan Perangkat Lunak Prediktor Nilai Mahasiswa Menggunakan Metode Spectral Clustering Dan Bagging Regresi Linier." Jurnal Teknik Its 1.1 (2012): A246-A250.
- [14] Sulthan, Nafis, Nurafni Eltivia, And Nur Indah Riwijanti. "Implementasi Metode Exponential Smoothing Untuk Peramalan Kedatangan Wisatawan Mancanegara Pada Pulau Bali." Media Mahardhika 18.2 (2020): 171-178.
- [15] Supriatna, Agus, Et Al. "Prediksi Wisatawan Mancanegara Ke Jawa Barat Melalui
- [16] Pintu Masuk Bandara Husein Sastranegara Dan Pelabuhan Muarajati Menggunakan Metode Sarima." Prosiding Industrial Research Workshop And National Seminar. Vol. 8. 2017.
- [17] Sovia, Rini, Musli Yanto, And Putri Melati. "Prediksi Jumlah Kunjungan Wisata Mancanegara Dengan Algoritma Backpropagation." Jurnal Media Informatika Budidarma 4.2 (2020): 355-36
- [18] Syafitri, Yuli. "Membangun Sistem Informasi Persediaan Barang Dagang." Jurnal Cendikia 15.2

Oktober (2017): 26-32.

- [19] Efendi, Dwi Marisa, and Ferly Ardhy. "Penerapan Data Mining Untuk Peramalan Penjualan Obat dengan Menggunakan Single Exponential Smoothing di Apotek Hamzah Farma." *Prosiding Seminar Nasional Darmajaya*. Vol. 1. No. 1. 2018.
- [20] Rahmatullah, Sidik, Mukrim Mukrim, And Monica Nindy Pramitha. "Data Mining Untuk Menentukan Produk Terlaris Menggunakan Metode Naive Bayes." *Jurnal Informasi Dan Komputer* 7.2 (2019): 57-6

