

PREDIKSI ANGKA KELAHIRAN BAYI PADA DATA MINING DENGAN METODE *MULTI REGRESSION*

Zulham¹, Buyung Solihin Hasugian²

¹⁾²⁾Fakultas Teknik dan Ilmu Komputer
Universitas Dharmawangsa

Corresponding Email: zulham@dharmawangsa.ac.id¹, buyung@dharmawangsa.ac.id²

RINGKASAN- Angka Kelahiran Bayi di suatu daerah merupakan indikator untuk mengetahui jumlah bayi dan jumlah bayi yang lahir setiap tahunnya. Hal ini akan menunjukkan perkembangan baru dan mulai meniru perilaku orang-orang terdekatnya namun tetap mengandalkan pengasuh untuk menyediakan segala kebutuhan dasarnya. Gizi buruk pada balita di Desa Sibolangit masih belum membaik, meskipun ada sedikit peningkatan. Dengan adanya Data mining maka kita dapat memprediksi jumlah kelahiran sehingga dapat mengetahui jumlah bayi yang lahir setiap tahun. Dengan menerapkan metode regresi linier berganda diharapkan dapat membantu menyelesaikan permasalahan yang ada selama ini di Desa Sibolangit. Dengan menggunakan Regresi Linier Berganda kita dapat memprediksi kelahiran bayi yang dapat dijadikan acuan untuk menentukan jumlah kelahiran. Proses pengolahan data menjadi lebih efektif dan efisien serta hasil prediksi jumlah bayi yang lahir lebih akurat dan tepat sasaran.

Kata Kunci: *Data Mining, Multi Regression, Algoritma, Prediksi.*

PENDAHULUAN

Angka kelahiran di Indonesia setiap tahun mengalami peningkatan seperti halnya di desa Sibolangit. Bayi lima tahun yang sering disebut balita merupakan anak yang mempunyai usia 12-59 bulan. Usia balita merupakan masa penting dalam perkembangan seorang anak, yang sering disebut dengan “masa emas” yang artinya dimana perkembangan fisik, psikis dan sosial seorang anak berkembang dengan sangat cepat serta tidak terulang kembali. Pada masa inilah balita akan menunjukkan perkembangan seperti senang melakukan hal-hal yang baru dan mulai meniru perilaku orang-orang terdekatnya. Oleh sebab itu perlu dibangun sebuah sistem yang mengadopsi dalam mengambil keputusan yang akan membantu pihak sekolah dalam merekrut tenaga pengajar.

Metode yang digunakan dalam Sistem Pendukung Keputusan ini adalah Weighted Sum Model yang banyak digunakan untuk membantu pengambil

keputusan dalam mengambil keputusan karena caregiver menyediakan semua kebutuhan dasarnya. Hingga saat ini masalah status gizi buruk pada balita di Desa Sibolangit belum juga membaik meskipun ada sedikit peningkatan. Pemantauan status gizi balita berdasarkan Indeks Berat Badan per umur yaitu 79,7% balita gizi baik, 14,9% balita gizi buruk, 3,8% balita gizi buruk dan 1,5% balita gizi lebih.

Data mining adalah proses yang menggunakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis. Penggunaan data mining juga bertujuan untuk mempermudah proses penggalian data yang akan dikelompokkan. Metode yang akan digunakan dalam mengekstrak data adalah Multi Regression. Dimana pengertian regresi berganda adalah memprediksi nilai variabel kontinu yang diberikan berdasarkan nilai variabel lain, dengan asumsi model ketergantungan linier atau non-linier.

Hal inilah yang melatarbelakangi penulisan tugas akhir ini dengan judul “Prediksi Angka Kelahiran Bayi Pada Data Mining Dengan Metode Multi Regresi” untuk memprediksi angka kelahiran bayi, karena seperti yang terlihat setiap tahun jumlah bayi di desa Sibolangit semakin meningkat. Desa Sibolangit belum mampu menghadapi tingginya angka kelahiran bayi karena berbagai sebab. Sehingga bayi di desa sibolangit semakin bertambah dan melewati batas yang telah diberitahukan oleh Pemerintah. Oleh karena itu perlu dibangun Sistem Pendukung Keputusan yang akan membantu pihak sekolah dalam merekrut tenaga pengajar.

Metode yang digunakan dalam Sistem Pendukung Keputusan ini adalah Weighted Sum Model yang banyak digunakan untuk membantu pengambil keputusan dalam mengambil keputusan karena dalam konsepnya metode ini hanya melakukan perkalian antara bobot kriteria dan alternatif untuk mengolah data penilaian tertentu dan memberikan hasil yang akurat dan cepat.

METODE PENELITIAN

1. Konsep Data Mining

Tan pada tahun 2006 mengartikan bahwa data mining merupakan suatu proses untuk untuk menghasilkan informasi yang berguna dalam jumlah data

yang besar. Data mining juga dapat diartikan sebagai penggalian informasi baru dari potongan besar data yang membantu dalam pengambilan keputusan. Istilah penambangan data kadang-kadang disebut sebagai penemuan pengetahuan.

2. Data Mining

Fajar Astuti Hermawati (2009: 3) mengartikan bahwa data mining adalah suatu proses iteratif dan interaktif dalam menentukan suatu pola maupun model baru yang berguna dan dapat dipahami.

- 1 Sahih: Bisa digeneralisir untuk masa depan.
2. Baru: memberikan pengetahuan awal.
3. Berguna: bisa dipergunakan untuk mengerjakan sesuatu tindakan.
4. Iteratif: membutuhkan proses yang berulang.
5. Interaktif: membutuhkan sebuah interaksi dari manusia.

3. Tahap Knowledge Discovery Database stages (KDD)

Pada Discovery Database terdapat beberapa tahapan yaitu sebagai berikut:

1. Seleksi

Mensortir data dari kumpulan data operasional sebelum masuk pada tahap penambangan di Knowledge Discovery Database (KDD) dimulai. Data kemudian dikumpulkan untuk digunakan pada proses data mining.

2. Prapemrosesan/Pembersihan

Proses preprocessing dimana dilakukan tindakan menghapus beberapa data yang double, pengecekan data yang tidak valid atau konsisten, kemudian mengoreksi kesalahan pada data, seperti kesalahan pada proses pencetakannya.

3. Transformasi

Transformasi dilakukan bagi entitas yang belum jelas agar menjadi entitas yang valid sehingga siap dilakukan penambangan data.

4. Penambangan Data

Pada proses ini yang adalah dengan mengimplementasikan algoritma atau metode pencarian pengetahuan yang dipakai.

5. Interpretasi/Evaluasi

Yang terakhir yang dilakukan adalah proses pembentukan informasi dari apa yang dihasilkan untuk menjadi suatu pengetahuan.

4. Linier Regression

Dalam metode linier regresi terdapat dua macam, yaitu Sederhana dan Berganda. Secara umum dapat kita katakan bahwasannya regresi linier sederhana hanya terfokus kepada satu variabel yang mempengaruhinya, sedangkan regresi linier berganda berfokus kepada lebih dari satu variabel.

4.1 Multiple Linier Regression

Melakukan wawancara adalah merupakan cara dimana kita bisa mendapatkan sebuah informasi yang lebih detail, langsung, tidak terstruktur, serta individu untuk menghasilkan sebuah informasi yang valid.

Tujuan melakukan analisis regresi linier berganda adalah mengukur intensitas hubungan antara dua variabel atau lebih untuk memprediksi nilai prediksi X atas Y. Dalam memprediksi Y, jika semua nilai variabel independen diketahui, maka persamaan Regresi Linier Berganda digunakan. Bentuk umum Berganda secara statistik dapat disimpulkan sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$$

Dimana $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ adalah koefisien atau parameter model.

Model regresi linier berganda untuk populasi di atas dapat diestimasi berdasarkan sampel acak berukuran n dengan model regresi linier berganda untuk sampel, yaitu:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Di mana :

Y = Variabel terikat

a = konstan

$b_1 - b_k$ = Koefisien Regresi

$X_1 - X_k$ = Variabel bebas

Tabel 1
Bentuk Umum Data Pengamatan

Observation Number	Respondents (Y)	Independent Variable			
		X ₁	X ₂	...	X _k
1	Y ₁	X ₁₁	X ₂₁	...	X _{k1}
2	Y ₂	X ₁₂	X ₂₂	...	X _{k2}
.

...
N	Y _n	X _{1n}	X _{2n}	...	X _{kn}
Σ	ΣY	ΣX ₁	ΣX ₂	ΣX _{...}	ΣX _k

Rumus dapat diselesaikan dengan empat persamaan dengan empat variabel yang terbentuk:

$$\sum Y = a_n + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 + \dots + b_k \sum X_k$$

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

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

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

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

5. Flowchart

Flowchart merupakan suatu diagram yang menampilkan tahap-tahap dan tindakan untuk mengambil keputusan pada saat proses di suatu program. Setiap tahap yang digunakan dalam bentuk diagram dan dihubungkan oleh garis atau arah panah.

HASIL PENELITIAN DAN PEMBAHASAN

1. Analisa Masalah

Analisis masalah adalah cara menilai atau menganalisis masalah dengan mengamati input dan pengolahan data dan informasi output untuk membantu meningkatkan proses pemecahan masalah.

Dalam representasi masalah tersebut dapat disimpulkan bahwa diperlukan beberapa kriteria atau variabel untuk memprediksi jumlah bayi yang lahir. Berikut tabel rekapitulasi jumlah angka kelahiran di Puskesmas Sibolangit.

Tabel 2
Rekapitulasi Total Angka Kelahiran di Puskesmas Sibolangit Medan Tahun 2006-2015

No	Year	Number of Spouses of Productive Age (X1)	Number of Births (X2)	Number of Deaths (X3)	Birthrate (Y)
1	2015	26	12	2	15
2	2014	20	12	3	23
3	2013	14	16	2	13
4	2012	16	8	3	9

5	2011	10	6	5	34
6	2010	16	12	3	11
7	2009	18	8	2	14
8	2008	12	10	2	22
9	2007	12	6	2	13
10	2006	16	8	4	17
AMOUNT		160	98	28	171

Source: Data processed from BPS

2. Analisis dan Estimasi Multiple Linear Regression

Regresi Linier Berganda adalah regresi yang memprediksi hubungan antara dua variabel terikat dengan dua atau lebih variabel bebas, yang akan menentukan hubungan antara Y dan X1, X2, Xk. Metode Regresi Linier Berganda merupakan salah satu jenis regresi linier dan statistik yang mengadopsi data mining untuk mengetahui pengaruh variabel-variabel tersebut.

Untuk memprediksi Y, jika semua nilai variabel independen diketahui, digunakan persamaan regresi linier berganda. Hubungan antara Y dan X1,X2,...Xk kebenarannya adalah sebagai berikut:

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

Dimana :

Y = Variabel tak bebas

a = konstan

b₁,X₁ = Koefisien regresi

b₂,X₂ = Variabel bebas

Persamaan regresi untuk kedua variabel bebas adalah

$$Y = a + b_1X_1 + b_2X_2$$

Secara umum persamaan regresi untuk k variabel adalah :

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

Untuk mendapatkan koefisien regresi a, b1, b2, b3 dapat diperoleh secara bersamaan dari tiga persamaan berikut :

$$\Sigma Y = na + \Sigma b_1 \Sigma X_1 + \Sigma b_2 \Sigma X_2$$

$$\Sigma YX_1 = a \Sigma X_1 + b_1 \Sigma X_1^2 + b_2 \Sigma X_1X_2$$

$$\Sigma YX_2 = a \Sigma X_2 + b_1 \Sigma X_1X_2 + b_2 \Sigma X_1X_2^2$$

1. Model Regresi Linier Berganda untuk Rekapitulasi Jumlah Kelahiran.

Berikut tabel rekapitulasi jumlah pasangan suami istri usia produktif, lahir dan meninggal:

Tabel 3
Gambaran Umum Perhitungan Regresi Linier Berganda Dalam Rekapitulasi Jumlah Angka Kelahiran

Tahun	ΣY	ΣX ₁	ΣX ₂	ΣX ₃	ΣYX ₁	ΣYX ₂	ΣYX ₃	ΣY ²	ΣX ₁ ²	ΣX ₂ ²	ΣX ₃ ²	ΣX ₁ X ₂	ΣX ₁ X ₃	ΣX ₂ X ₃
2015	15	26	12	2	390	180	30	225	676	144	4	312	52	24
2014	23	20	12	3	460	180	45	529	400	144	9	240	60	36
2013	13	14	16	2	182	240	30	169	196	256	4	224	28	32
2012	9	16	8	3	144	120	45	81	256	64	9	128	48	24
2011	34	10	6	5	340	90	75	1156	100	36	25	60	50	30
2010	11	16	12	3	176	180	45	121	256	144	9	192	48	36
2009	14	18	8	2	252	120	30	196	324	64	4	144	36	16
2008	22	12	10	2	264	150	30	484	144	100	4	120	24	20
2007	13	12	6	2	156	90	30	169	144	36	4	72	24	12
2006	17	16	8	4	272	120	60	289	256	64	16	128	64	32
Amount	171	160	98	28	2636	1470	420	3419	2752	1052	88	1620	434	262

Kemudian masukkan angka-angka yang diperoleh dalam rumus persamaan sehingga hasilnya adalah sebagai berikut:

Langkah – langkah penyelesaiannya adalah sebagai berikut :

Persamaan satu dan dua dihilangkan:

$$\begin{array}{rcl}
 10b_0 + 160b_1 + 98b_2 + 28b_3 & = & 171 \quad \times 16 \\
 160b_0 + 2756b_1 + 162b_2 + 434b_3 & = & 2636 \quad \times 1 \\
 \hline
 \end{array}$$

$$\begin{array}{rcl}
 160b_0 + 2560b_1 + 1568b_2 + 448b_3 & = & 2736 \\
 160b_0 + 2756b_1 + 162b_2 + 434b_3 & = & 2636 \quad - \\
 \hline
 \end{array}$$

$$0 - 196b_1 + 1406b_2 + 14b_3 = 100 \text{ (Persamaan 5)}$$

Kemudian persamaan satu dan tiga dihilangkan:

$$\begin{array}{rcl}
 10b_0 + 160b_1 + 98b_2 + 28b_3 & = & 171 \quad \times 800 \\
 8000b_0 + 1620b_1 + 1052b_2 + 262b_3 & = & 1470 \quad \times 1 \\
 \hline
 \end{array}$$

$$\begin{array}{rcl}
 8000b_0 + 128000b_1 + 78400b_2 + 22400b_3 & = & 136800 \\
 8000b_0 + 1620b_1 + 1052b_2 + 262b_3 & = & 1470 \quad - \\
 \hline
 \end{array}$$

$$0 + 126380b_1 + 77438b_2 + 22138b_3 = 135330 \text{ (Persamaan 6)}$$

Kemudian kita hilangkan persamaan satu dan empat:

$$10b_0 + 160b_1 + 98b_2 + 28b_3 = 171 \quad \times 28$$

$$28b_0 + 434b_1 + 262b_2 + 88b_3 = 420 \quad \times 10$$

$$280b_0 + 4480b_1 + 2744b_2 + 784b_3 = 4788$$

$$280b_0 + 4340b_1 + 2620b_2 + 880b_3 = 4200 \quad -$$

$$0 + 140b_1 + 124b_2 - 96b_3 = 588 \text{ (Persamaan 7)}$$

Persamaan lima dan enam dihilangkan:

$$-196b_1 + 1406b_2 + 14b_3 = 100 \quad \times 644.79$$

$$126380b_1 + 77438b_2 + 22138b_3 = 135330 \quad \times 644.79$$

$$-126380b_1 + 906581.77b_2 + 9027.13b_3 = 64479.5$$

$$126380b_1 + 49931248.02b_2 + 14274471.71b_3 = 8726010.735 \quad -$$

$$0 + 49024666b_2 + 14265444.58b_3 = 8661537.235 \text{ (Persamaan 8)}$$

Persamaan lima dan tujuh dihilangkan :

$$-196b_1 + 1406b_2 + 14b_3 = 100 \quad \times 1,4$$

$$140b_1 + 124b_2 - 96b_3 = 588 \quad \times 1,4$$

$$-196b_1 + 1968.4b_2 + 19.6b_3 = 140$$

$$196b_1 + 173.6b_2 + 134.4b_3 = 823.2 \quad -$$

$$0 + 2142b_2 + 154b_3 = 963.2 \text{ (Persamaan 9)}$$

Persamaan delapan dan sembilan dihilangkan :

$$49024666b_2 + 14265444.58b_3 = 8661537.235 \quad \times 1$$

$$2142b_2 + 154b_3 = 963.2 \quad \times 22887.332$$

$$49024666b_2 + 14265444.58b_3 = 8661537.235$$

$$49024666b_2 + 3524649.182b_3 = 22045078.182 \quad -$$

$$10740795.452b_3 = -13383540.947$$

$$b_3 = -1.246$$

$$2142b_2 + 154b_3 = 963.2$$

$$2142b_2 + 154(-1.246) = 963.2$$

$$2142b_2 - 191.884 = 963.2$$

$$2142b_2 = 1155$$

$$b_2 = 0.539$$

substitusi ke persamaan lima

$$140b_1 + 124b_2 - 96b_3 = 588$$

$$140b_1 + 124(0.539) - 96(-1.246) = 588$$

$$140b_1 + 66.836 - 119.616 = 588$$

$$140b_1 = 588 + 52.78$$

$$140b_1 = 640.78$$

$$b_1 = 513.148/140$$

$$b_1 = 4.5$$

substitusi ke persamaan satu

$$10b_0 + 160b_1 + 98b_2 + 28b_3 = 171$$

$$10b_0 + 160(3.665) + 98(0.539) + 28(-1.246) = 171$$

$$10b_0 + 586.4 + 52.822 - 34.888 = 171$$

$$10b_0 = 415.89$$

$$b_0 = -43.334/10$$

$$b_0 = -43.333$$

maka dari perhitungan diatas diperoleh persamaan regresi linier berganda sebagai berikut:

$$Y = -43.333 + (3.665)X_1 + (0.36)X_2 + (-1.246)X_3$$

Setelah kita mendapatkan persamaan regresi linier berganda, kita dapat menentukan jumlah bayi, nilai berikut akan dimasukkan berdasarkan data sebelumnya:

$$Y = -43.333 + (4.5)X_1 + (0.539)X_2 + (-1.246)X_3$$

$$Y = -43.333 + (4.5)26 + (0.539)12 + (-1.246)2$$

$$Y = -43.333 + 117 + 6.468 + 4.32 - 2.492$$

$$Y = 86.947 = 87 \text{ (dibulatkan) jumlah bayi yang lahir}$$

Maka dari uraian diatas dapat kita tentukan bahwa angka kelahiran di Puskesmas Sibolangit dari perhitungan persamaan regresi linier berganda diatas adalah 87 orang

SIMPULAN

Berdasarkan dari apa yang telah dilakukan dalam penelitian ini, maka dapat disimpulkan sebagai berikut:

1. Merujuk pada hasil analisa dalam memprediksi jumlah kelahiran, kita mampu memprediksi jumlah kelahiran bayi berdasarkan hasil penelitian sehingga dapat membantu pemerintah setempat dalam mengatasi masalah berhubungan dengan angka kelahiran bayi secara efisien dan efektif..
2. Dalam rancangan ini untuk memprediksi jumlah kelahiran bayi sebaiknya kita menentukan variabel serta menghitung koefisien data dari persamaan regresi, untuk kemudian mendapatkan hasil persamaan regresi tersebut.
3. Hasil penelitian ini bisa diterapkan serta diimplementasikan guna memprediksi jumlah kelahiran di daerah tersebut.

DAFTAR PUSTAKA

- D. Triansyah and D. Fitriana, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," *J. Telekomun. dan Komput.*, vol. 8, no. 3, p. 163, 2018, doi: 10.22441/incomtech.v8i3.4174.
- N. R. Indraswari and Y. I. Kurniawan, "Aplikasi Prediksi Usia Kelahiran Dengan Metode Naive Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 9, no. 1, pp. 129–138, 2018, doi: 10.24176/simet.v9i1.1827.
- R. A. Ashley and C. F. Parmeter, "Sensitivity analysis of an OLS multiple regression inference with respect to possible linear endogeneity in the explanatory variables, for both modest and for extremely large samples," *Econometrics*, vol. 8, no. 1, 2020, doi: 10.3390/econometrics8010011.
- M. L. Sibuea and A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," *Jurteks*, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteks.v4i1.28.
- I. Muthahharah and I. Fatwa, "Modeling The Types of Online Learning Media Using Multiple Linear Regression Analysis," *J. Varian*, vol. 5, no. 1, pp. 39–46, 2021, doi: 10.30812/varian.v5i1.1459.
- W. Dhuhita, "Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita," *J. Inform. Darmajaya*, vol. 15, no. 2, pp. 160–174, 2015.
- C. C.L, J. D.S, and S. Sonoli, "Multiple Linear Regression Analysis for Prediction of Boiler Losses and Boiler Efficiency," *Int. J. Instrum. Control Syst.*, vol. 8, no. 2, pp. 01–09, 2018, doi: 10.5121/ijics.2018.8201.
- A. H. Nasyuha, Zulham, and I. Rusydi, "Implementation of K-means algorithm in data analysis," *Telkomnika (Telecommunication Comput. Electron. Control.)*, vol. 20, no. 2, pp. 307–313, 2022, doi: 10.12928/TELKOMNIKA.v20i2.21986.
- H. Sy, Rismayani, and A. Syam, "Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare di Kota Makassar," *SISITI Semin. Ilm. Sist. Inf. dan Teknol. Inf.*, vol. 8, no. 1, pp. 73–82, 2019.
- S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-

- Means Clustering,” *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.
- E. E. Putri, B. S. Hasugian, A. Info, and D. Mining, “Pattern Analysis of Drug Procurement System With FP-Growth Algorithm,” vol. 7, no. 1, pp. 70–79, 2022, doi: 10.15575/join.v7i1.841.
- C. Yuan and H. Yang, “Research on K-Value Selection Method of K-Means Clustering Algorithm,” *J*, vol. 2, no. 2, pp. 226–235, 2019, doi: 10.3390/j2020016.
- L. Pun, P. Zhao, and X. Liu, “A Multiple Regression Approach for Traffic Flow Estimation,” *IEEE Access*, vol. 7, no. c, pp. 35998–36009, 2019, doi: 10.1109/ACCESS.2019.2904645.
- B. D. F. Kurniatullah and Y. T. C. Pramudi, “Estimation of Students’ Graduation Using Multiple Linear Regression Method,” *J. Appl. Intell. Syst.*, vol. 2, no. 1, pp. 29–36, 2017, doi: 10.33633/jais.v2i1.1415.
- C. Farkas, L. Fenyvesi, and K. Petr oczki, “Multiple linear regression model of golden apple’s failure characteristics under repeated compressive load,” *Potravin. Slovak J. Food Sci.*, vol. 13, no. 1, pp. 793–799, 2019, doi: 10.5219/1168.
- R. Setiawan, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik Lp3i Jakarta),” *J. Lentera Ict*, vol. 3, no. 1, pp. 76–92, 2016.
- A. Triayudi and I. Fitri, “Comparison Of The Feature Selection Algorithm In Educational Data Mining,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 19, no. 6, pp. 1865–1871, 2021, doi: 10.12928/TELKOMNIKA.v19i6.21594.
- M. Lubis, D. O. D. Handayani, and A. R. Lubis, “Prediction analysis of the happiness ranking of countries based on macro level factors,” *IAES Int. J. Artif. Intell.*, vol. 11, no. 2, pp. 666–678, 2022, doi: 10.11591/ijai.v11.i2.pp666-678.
- S. Fatimah and W. Wiharto, “Multiple linear regression for the analysis of the parameters used in dyes decolourisation by ozonation techniques,” *MATEC Web Conf.*, vol. 154, pp. 2–6, 2018, doi: 10.1051/mateconf/201815401003.
- H. Kang and H. Zhao, “Description and Application Research of Multiple Regression Model Optimization Algorithm Based on Data Set Denoising,” *J. Phys. Conf. Ser.*, vol. 1631, no. 1, 2020, doi: 10.1088/1742-6596/1631/1/012063.
- H. H. Nuha and A. A. Absa, “Data Visualization of COVID-19 Vaccination Progress and Prediction Using Linear Regression,” vol. 7, no. 1, pp. 1–6, 2022, doi: 10.15575/join.v7i1.736.