

Memprediksi Jumlah Siswa Baru ERS Indonesia Menggunakan Metode Algoritma Regresi Linier

May Rani Wanisah Lubis¹, Novryenni², Marto Sihombing³

STMIK Kaputama, Binjai, Indonesia

e-mail : ¹wanisahlubis@gmail.com, ²novriyenni.sikumbang@gmail.com,
³martosihombing45@gmail.com

Abstrak

ERS (Education Resources Service) Indonesia merupakan sebuah lembaga pendidikan yang menerima pendaftaran siswa baru setiap tahunnya. Dalam menjalankan kegiatan penerimaan siswa baru, ERS Indonesia perlu melakukan perencanaan dan estimasi jumlah siswa yang akan mendaftar agar dapat melakukan persiapan yang tepat dalam hal fasilitas dan tenaga pengajar. Dalam hal ini, metode algoritma regresi linier dapat digunakan untuk memprediksi jumlah siswa baru yang akan mendaftar di ERS Indonesia. Dengan menggunakan metode regresi linier, ERS Indonesia dapat memperkirakan jumlah siswa baru yang akan mendaftar pada tahun ajaran berikutnya berdasarkan data historis dan faktor-faktor yang mempengaruhi jumlah siswa baru. Memprediksi jumlah siswa baru di sebuah institusi pendidikan memiliki peranan yang sangat penting dalam perencanaan dan pengelolaan pendidikan. Selain itu, dengan memprediksi jumlah siswa baru dimasa depan, institusi pendidikan juga dapat meningkatkan daya saingnya dengan mengembangkan strategi pemasaran yang lebih efektif.

Kata kunci: *ERS (Education Resources Service), Regresi Linier, Estimasi Jumlah Siswa*

Abstract

ERS (Education Resources Service) Indonesia is an educational institution that accepts new student registrations every year. In carrying out new student admission activities, ERS Indonesia needs to plan and estimate the number of students who will register so that they can make appropriate preparations in terms of facilities and teaching staff. In this case, the linear regression algorithm method can be used to predict the number of new students who will register at ERS Indonesia. By using the linear regression method, ERS Indonesia can estimate the number of new students who will register in the next academic year based on historical data and factors that influence the number of new students. Predicting the number of new students at an educational institution has a very important role in educational planning and management. In addition, by predicting the number of new students in the future, educational institutions can also increase their competitiveness by developing more effective marketing strategies.

Keywords: *Number of new students, linear regression*

1. Pendahuluan

ERS (Education Resources Service) Indonesia merupakan sebuah lembaga pendidikan yang menerima pendaftaran siswa baru setiap tahunnya. Dalam menjalankan kegiatan penerimaan siswa baru, ERS Indonesia perlu melakukan perencanaan dan estimasi jumlah siswa yang akan mendaftar agar dapat melakukan persiapan yang tepat dalam hal fasilitas dan tenaga pengajar [1].

Dalam hal ini, metode algoritma regresi linier dapat digunakan untuk memprediksi jumlah siswa baru yang akan mendaftar di ERS Indonesia. Regresi linier adalah metode statistik yang digunakan untuk mempelajari hubungan antara variabel dependen (dalam hal ini, jumlah siswa baru) dan satu atau lebih variabel independen (seperti tahun ajaran, jumlah pengunjung situs web atau promosi yang dilakukan). Dalam hal ini, tahun ajaran dapat dianggap sebagai variabel independen karena dapat mempengaruhi jumlah siswa baru yang mendaftar [2].

Dengan menggunakan metode regresi linier, ERS Indonesia dapat memperkirakan jumlah siswa baru yang akan mendaftar pada tahun ajaran berikutnya berdasarkan data historis dan faktor-faktor yang mempengaruhi jumlah siswa baru [3]. Dengan demikian, ERS Indonesia dapat melakukan perencanaan yang lebih efektif dalam hal fasilitas dan tenaga pengajar yang dibutuhkan untuk menampung siswa baru tersebut.

Memprediksi jumlah siswa baru di sebuah institusi pendidikan memiliki peranan yang sangat penting dalam perencanaan dan pengelolaan pendidikan. Dalam konteks Indonesia, sistem pendidikan mengalami perkembangan pesat dengan bertambahnya jumlah institusi pendidikan baik di tingkat sekolah dasar, menengah, maupun tinggi. Namun, dengan semakin ketatnya persaingan diantara institusi pendidikan, setiap institusi perlu memastikan bahwa mereka dapat menarik dan mempertahankan jumlah siswa yang memadai [4]. Oleh karena itu, institusi pendidikan perlu dapat memprediksi jumlah siswa baru di masa depan agar dapat mempersiapkan diri dengan lebih baik dalam menghadapi tahun ajaran yang akan datang.

Regresi Linier merupakan metode statistika yang menjelaskan tentang pola model untuk membentuk hubungan antara variabel terikat dan variabel bebas atau lebih (Nur Kholida Afkarina, Agus Wahyu Widodo, Muhammad Tanzil Furqon, n.d). Metode yang sering digunakan dalam memprediksi jumlah siswa baru adalah menggunakan algoritma regresi linier. Regresi linier merupakan salah satu metode statistik yang dapat mengestimasi hubungan antara variabel independen dan variabel dependen. Dalam konteks ini, variabel independen adalah tahun ajaran, sedangkan variabel dependen adalah jumlah siswa baru. Dengan menggunakan algoritma regresi linier, institusi pendidikan dapat memprediksi jumlah siswa baru di masa depan berdasarkan data historis yang tersedia. Hal ini dapat membantu institusi pendidikan dalam merencanakan dan mengelola sumber daya yang tersedia, seperti fasilitas, tenaga pengajar, dan keuangan.

Menurut (Syahputra et al., 2020) salah satu tujuan didakannya proses pendidikan adalah untuk meningkatkan kualitas kehidupan warga negaranya dan daya saing sebuah bangsa, sarana dan prasarana yang memadai sangat mendukung tercapainya proses pendidikan yang berkualitas.

Selain itu, dengan memprediksi jumlah siswa baru dimasa depan, institusi pendidikan juga dapat meningkatkan daya saingnya dengan mengembangkan strategi pemasaran yang lebih efektif. Institusi pendidikan dapat menggunakan hasil prediksi untuk mengembangkan program pemasaran yang lebih terarah dan efektif untuk menarik calon siswa baru. Oleh karena itu, memprediksi jumlah siswa baru disebuah institusi pendidikan menjadi sebuah informasi yang sangat berharga dalam perencanaan dan pengelolaan pendidikan. Dalam konteks Indonesia, metode algoritma regresi linier dapat menjadi salah satu solusi yang efektif dalam memprediksi jumlah siswa baru di masa depan.

Hal ini dapat membantu institusi pendidikan dalam membuat keputusan yang lebih baik dan mempersiapkan diri dengan lebih baik dalam menghadapi tahun ajaran yang akan datang.

Dengan demikian, memprediksi jumlah siswa baru di sebuah institusi pendidikan menjadi sebuah kebutuhan yang sangat penting dan dapat membantu institusi pendidikan dalam mengoptimalkan kinerjanya serta meningkatkan daya saingnya.

Berdasarkan uraian di atas, maka perlu memprediksi jumlah siswa baru ERS Indonesia menggunakan metode algoritma regresi linier yang diharapkan dapat membantu ERS Indonesia agar mudah memprediksi jumlah siswa.

2. Metode Penelitian

2.1 Regresi Linier

Regresi linier adalah suatu metode untuk menguji pengaruh variable independen terhadap variable dependent, dimana peubah bebasnya hanya satu peubah. Regresi linier adalah metode permalan statistika yang digunakan untuk membentuk model hubungan antara variabel terikat (dependen; respon; Y) dengan satu atau lebih variabel bebas (independen, prediktor, X). Pola yang ditunjukkan dengan analisa regresi yang sederhana mengasumsikan bahwa hubungan diantara 2 variabel dapat dinyatakan dengan suatu garis lurus (Pamungkas, 2021).

Regresi merupakan metode pengambilan keputusan yang sering digunakan dalam pengembangan model matematis. Perkembangan teknologi yang cepat membutuhkan model matematis yang dapat diandalkan yang akan memudahkan proses identifikasi variable utama, estimasi, dan dalam perancangan proyek simulasi (Muriyatmoko, 2018). Linier Regresi termasuk dalam time series model yang menggunakan metode kuantitatif dengan menggunakan waktu sebagai dasar peramalan. Linier Regresi mempunyai persamaan dasar sebagai berikut :

$$Y = a + bx \quad (1)$$

dimana :

Y : nilai ramalan periode ke-t

a : intercept

b : slope dari garis kecenderungan, merupakan tingkat perubahan

x : indeks waktu (t=1, 2, 3, ..., n); n adalah banyaknya periode waktu

Komponen pada Linier Regresi ada tiga yaitu a sebagai intercept, b sebagai slope dan x sebagai indeks waktu. Perasamaan untuk mendapatkan nilai a dan b adalah :

$$a = \frac{(\sum y)(\sum x^2) - (\sum x) \sum xy}{n \sum x^2 - (\sum x)^2} \quad (2)$$

$$n \sum x^2 - (\sum x)^2 \quad (3)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (4)$$

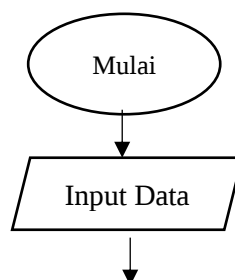
$$n(\sum x^2) - (\sum x)^2 \quad (5)$$

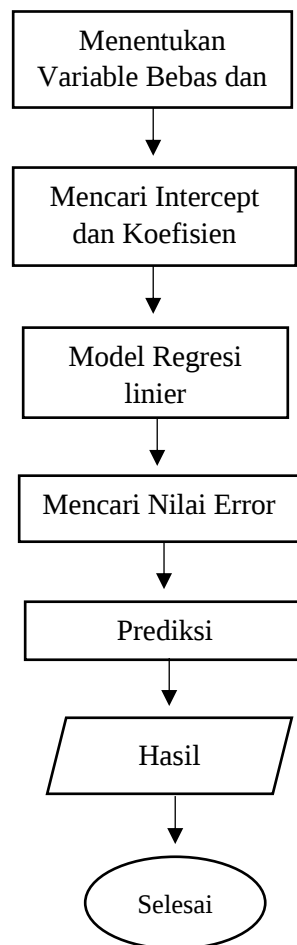
Sedangkan nilai a dan b dapat dicari melalui:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n(\sum X^2) - (\sum X)^2} \quad (6)$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} \quad (7)$$

Adapun tahapan metode regresi linier sederhana adalah sebagai berikut :





Gambar 1. Diagram Aliran tahapan Penelitian

Langkah-langkah metode yang diusulkan berdasarkan Linier Regresi sebagai berikut :

1. Pembuatan dataset yang terdiri dari data training dan data testing
2. Pembentukan model Linier Regresi (model dibuat berdasarkan data training).

Langkah pembentukan model sebagai berikut:

- a. Langkah 1: Hitung X^2 , Y^2 , XY dan total dari masing-masingnya
 - b. Langkah 2: Hitung a dengan menggunakan persamaan dan b menggunakan persamaan .
 - c. Langkah 3: Buatlah model persamaan Linier Regresi sederhana.
 - d. Langkah 4: Lakukan prediksi atau peramalan terhadap variabel faktor penyebab atau variabel akibat
3. Pengujian performa berdasarkan model prediksi yang telah dibuat dengan input data testing. Mean Absolute Percentage Error (MAPE) merupakan metode yang digunakan untuk menilai tingkat keakuratan.

$$MAPE = \frac{\sum \frac{|y - y'|}{y} \times 100\%}{n} \quad (8)$$

dimana :

y' : hasil prediksi
 y : Data aktual
 n : Jumlah data

2.2 Jumlah Siswa Baru

Penerimaan siswa baru merupakan proses awal yang dilakukan oleh pihak kursus guna untuk memenuhi jumlah siswa yang akan menjadi siswa di sebuah kursus. Proses rekrutmen siswa baru sudah menjadi agenda rutin yang dilaksanakan setiap kursus. Agenda tahunan yang dilakukan bertujuan untuk menerima siswa/siswi yang ingin belajar dan di didik dengan berbagai macam pengetahuan dan diberikan bimbingan serta pembinaan untuk mengembangkan kepribadian peserta didik (Baihaqi & Junaedi, 2022).

2.3 Prediksi

Prediksi digunakan untuk mendapatkan gambaran yang jelas tentang masa depan meneruskan dan menggabungkan data besar untuk meningkatkan pekerjaan efektif sederhana. Prediksi memberikan kesempatan untuk mendapatkan informasi tentang masa lalu untuk melihat informasi terbaru dan memprediksi hasilnya. Prediksinya adalah proses peramalan kejadian yang paling mungkin terjadi secara sistematis akan terjadi di masa depan berdasarkan informasi masa lalu dan sekarang, dalam urutan kesalahan (perbedaan antara apa yang terjadi dan apa yang diharapkan). Berkurang Prediksi tidak harus memberikan jawaban atas apa yang terjadi (Huda et al., 2020).

2.3 Metode

Metode merujuk pada cara atau pendekatan yang digunakan untuk mencapai tujuan tertentu dalam suatu aktivitas atau proses. Dalam konteks pendidikan atau penelitian, metode mengacu pada langkah-langkah yang diambil untuk mengumpulkan data, menganalisis informasi, atau mengajarkan materi kepada peserta didik.

2.4 Proses perhitungan

Adapun data penelitian yang digunakan merupakan data yang didapat dari ERS INDONESIA. Data yang digunakan sebagai pendukung penelitian yaitu berjumlah 8 data dengan variabel yang digunakan yaitu, tahun ajaran, semester, jumlah siswa (orang).

Tahap 1. Pembuatan/Pengolahan Dataset Penelitian

Berikut dataset Siswa ERS Indonesia menggunakan metode regresi linier :

Tabel 1. Dataset Penelitian ERS INDONESIA

Tahun Ajaran	Semester	Jumlah Siswa (Orang)
2019/2020	Ganjil	60
2019/2020	Genap	95
2020/2021	Ganjil	108
2020/2021	Genap	58
2021/2022	Ganjil	52
2021/2022	Genap	80
2022/2023	Ganjil	105
2022/2023	Genap	111

Tahap 2. Pembentukan model Linier Regresi yang terdiri dari :

a. Hitung X^2 , XY dan total dari masing-masing kolom.

Tabel 2. Perhitungan X^2 dan XY ERS INDONESIA

Tahun Ajaran	Semester	Jumlah Siswa Semester Ini (X)	Jumlah siswa Semester Berikutnya (Y)	XY	X^2
2019/2020	Ganjil	60	95	5700	3600
2019/2020	Genap	95	108	10260	9025
2020/2021	Ganjil	108	58	6264	11664
2020/2021	Genap	58	52	3016	3364
2021/2022	Ganjil	52	80	4160	2704
2021/2022	Genap	80	105	8400	6400
2022/2023	Ganjil	105	111	11.655	11025
Total		558	609	49455	47782

b. Hitung nilai a dengan menggunakan persamaan (2) dan nilai b menggunakan persamaan (3).

$$a = \frac{(609)(47782) - (558)(49455)}{7(47782) - (558)^2}$$

$$a = \frac{29099238 - 27595890}{334474 - 311364} = \frac{1503348}{23110} = 65.051839$$

$$b = \frac{7(49455) - (558)(609)}{7(47782) - (558)^2}$$

$$b = \frac{346185 - 339822}{334474 - 311364} = \frac{6363}{23110} = 0.2753353$$

c. Buatlah model persamaan Linier Regresi sederhana dengan menggunakan persamaan (1).

$$Y = 65.051839 + 0.275353 x$$

d. Lakukan prediksi atau peramalan terhadap variabel faktor penyebab atau variabel akibat. Misalnya akan mencari nilai Y (semester Ganjil 2023/2024) dengan x (semester Genap 2022/2023) = 110

$$Y = 65.051839 + 0.275353 (110) = 95$$

Tahap 3. Pengujian performa berdasarkan model prediksi yang telah dibuat dengan input data testing dengan output MAPE dengan menggunakan persamaan (4).

Untuk melakukan pengujian performa model yang sudah dibuat, maka dilakukan prediksi dengan menggunakan dataset pada Jumlah Siswa Baru mulai dari Semester Genap Tahun Ajaran 2019/2020 s/d Semester Genap 2022/2023, dimana data untuk variabel x adalah data pada semester sebelumnya dengan hasil perhitungan sebagai berikut :

Tabel 3. Hasil MAPE

Tahun Ajaran	Semester	Dataset Semester Sebelumnya (x)	Data Aktual (y)	Data Prediksi (y')	Selisih (y-y')	Error MAPE (%)
2019/2020	Genap	60	95	81.571957	13.428043	14.1
2020/2021	Ganjil	95	108	91.2086925	16.7913075	15.5
2020/2021	Genap	108	58	94.7880514	-36.7880514	63.4
2021/2022	Ganjil	58	52	81.0212864	-29.021264	55.8
2021/2022	Genap	52	80	79.3692746	0.6307254	0.78
2022/2023	Ganjil	80	105	87.078663	17.921337	17.1
2022/2023	Genap	105	111	93.9620455	17.0379545	15.3
Total						182.0

$$MAPE = \frac{182.0 \times 100\%}{7} = 26\%$$

Berdasarkan hasil pengujian tingkat error prediksi siswa registrasi ERS INDONESIA didapatkan hasil 26% atau tingkat akurasi sebesar 74 %.

Tabel 4. Hasil Jumlah Siswa

Tahun Ajaran	Semester	Jumlah siswa sebelumnya	Jumlah siswa berikutnya
2022/2023	Genap	105	111
2023/2024	Ganjil	111	205

Jadi untuk hasil akhir untuk semester selanjutnya akan ada kenaikan dengan 94 siswa dari siswa sebelumnya 111 siswa. Kemungkinan kenaikan prediksi jumlah siswa baru untuk semester berikutnya 205 siswa.

3. Hasil dan pembahasan

3.1 Load Data

Data pendukung disini menampilkan tahun ajaran, semester, jumlah siswa persemester sekarang dan jumlah siswa semester berikutnya.

	Tahun Ajaran	Semester	Jumlah Siswa Persemester Sekarang \
0	2019/2020	Ganjil	60
1	2019/2020	Genap	95
2	2020/2021	Ganjil	108
3	2020/2021	Genap	58
4	2021/2022	Ganjil	52
	Jumlah Siswa	Semester	Berikutnya
0			95
1			108
2			58
3			52
4			80
(7, 4)			

Gambar 2. Tampilan Data Pendukung

3.2 Data Set Siswa

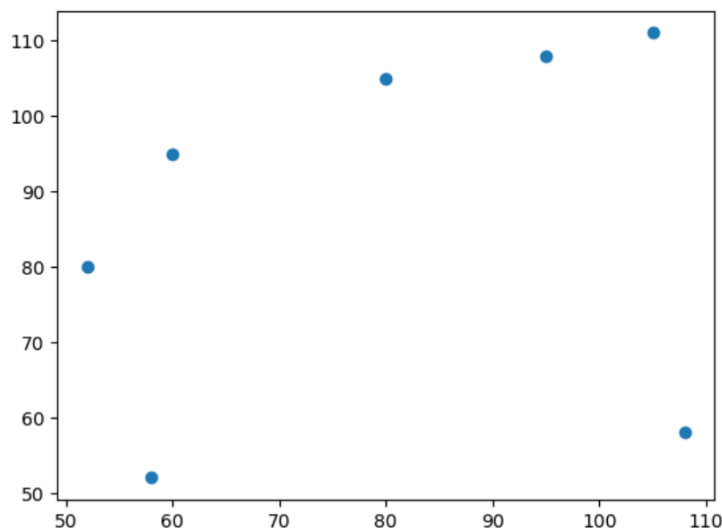
Data set akan menampilkan data pendukung prediksi dengan data x dan y. Ketika program memanggil data maka akan muncul jumlah siswa x dan y.

	x	y
0	60	95
1	95	108
2	108	58
3	58	52
4	52	80
(7, 2)		

Gambar 3. Tampilan Data Set

3.3 Grafik Jumlah Siswa

Pada grafik akan muncul jumlah siswa semester sebelumnya dan semester berikutnya. Ketika memasukkan program data sebab dan akibat maka akan muncul grafik jumlah siswa semester sebelumnya dan semester berikutnya.



Gambar 4. Grafik Jumlah Siswa Semester Sebelumnya

3.4 Nilai Korelasi

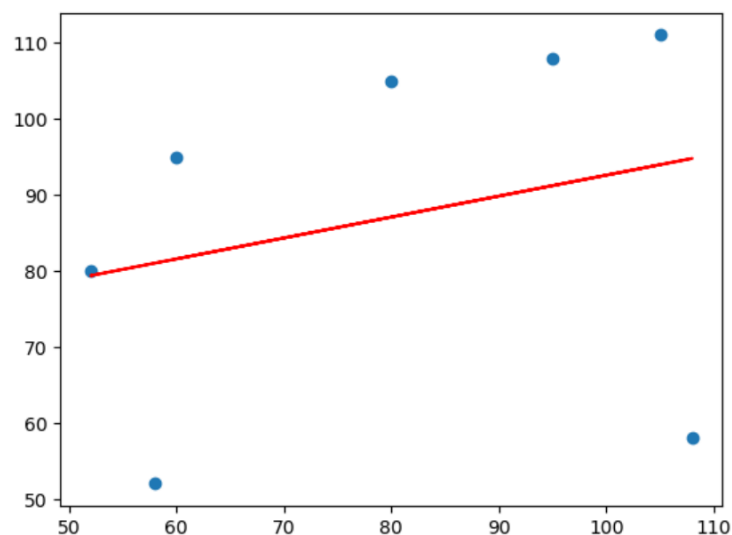
Dari output dapat diketahui bahwa nilai korelasi antara x dan y adalah 0.26665. Nilai korelasi tersebut menunjukkan adanya hubungan yang cukup tinggi antara kedua variabel.

	x	y
x	1.00000	0.26665
y	0.26665	1.00000

Gambar 5. Korelasi X dan Y

3.5 Hasil Prediksi Jumlah Siswa Berikutnya

Pada grafik ini akan menampilkan prediksi jumlah siswa selanjutnya. Pada garis menunjukkan perkiraan prediksi jumlah siswa yang akan mendaftar untuk semester kedepan. Dan pada grafik ini bahwasannya ada kenaikan di beberapa semester.



Gambar 6. Grafik Jumlah Siswa Semester Berikutnya

4 Kesimpulan

Berdasarkan hasil penelitian di atas dapat ditarik kesimpulan, hasil prediksi ERS INDONESIA dengan hasil pengujian tingkat error sebesar 26% atau tingkat akurasi sebesar 74%. Metode Linier Regresi dapat digunakan untuk memprediksi siswa registrasi secara tepat dan aplikasi yang sudah dibangun dapat digunakan karena memiliki tingkat akurasi di atas 70%. Disarankan kepada peneliti berikutnya bahwa penelitian ini dapat diteliti kembali dengan menggunakan metode prediksi yang lain dan jumlah dataset yang lebih banyak lagi, agar hasilnya bisa dibandingkan mana yang lebih baik.

DaftarPustaka

- [1] Baihaqi, H. A., & Junaedi, L. (2022). SISTEM PAKAR PENERIMAAN SISWA BARU
Baihaqi, H. A., & Junaedi, L. (2022). SISTEM PAKAR PENERIMAAN SISWA BARU
SEKOLAH DASAR BERDASARKAN TINGKAT IQ MENGGUNAKAN METODE
DEMPSTER SHAFER (STUDI KASUS: SEKOLAH DASAR LUQMAN AL HAKIM

- SURABAYA). *JUST IT: Jurnal Sistem Informasi, Teknologi Informasi Dan Komputer*, 12(2).
- [2] Bengnga, A., & Ishak, R. (2018). Prediksi Jumlah Mahasiswa Registrasi Per Semester Menggunakan Linier Regresi Pada Universitas Ichsan Gorontalo. *ILKOM Jurnal Ilmiah*, 10(2), 136–143.
 - [3] Harliana, H., & Syafrianto, A. (2017). Prediksi jumlah pendaftaran calon mahasiswa baru dengan metode regresi linier. *Data Manajemen Dan Teknologi Informasi (DASI)*, 18(3), 1–5.
 - [4] Huda, A. S., Awangga, R. M., & Fathonah, R. N. S. (2020). *Prediksi Penerimaan Pegawai Baru Dengan Metode Naive Bayes* (Vol. 1). Kreatif.
 - [5] Muriyatmoko, D. (2018). Analisa Volume Terhadap Sitasi Menggunakan Regresi Linier Pada Jurnal Bereputasi di Indonesia. *Jurnal Simantec*, 6(3).
 - [6] Nainggolan, R. Z., Ibnutama, K., & Suryanata, M. G. (2021). Implementasi Data Mining Dengan Metode Regresi Linier Berganda Dalam Estimasi Mahasiswa Baru Pada Sekolah Tinggi Agama Islam Raudhatul Akmal Batang Kuis. *Jurnal Cyber Tech*, 1(1).
 - [7] Pamungkas, N. B. (2021). PREDIKSI JUMLAH PENERIMAAN SISWA BARU PADA SMAN 1 PAGA DEWA DENGAN METODE REGRESI LINIER. *Jurnal Teknologi Dan Sistem Informasi*, 2(1), 8–14.
 - [8] Purwadi, P., Ramadhan, P. S., & Safitri, N. (2019). Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 18(1), 55–61.
 - [9] SAKIT, P. R. I. D. I. R., & ETAHAM, E. (n.d.). *YAYASAN LEMBAGA PENDIDIKAN ISLAM RIAU UNIVERSITAS ISLAM RIAU FAKULTAS ILMU SOSIAL DAN ILMU POLITIK*.
 - [10] Setyoningrum, N. R., & Rahimma, P. J. (2022). Implementasi Algoritma Regresi Linear Dalam Sistem Prediksi Pendaftar Mahasiswa Baru Sekolah Tinggi Teknologi Indonesia Tanjungpinang. *Prosiding Seminar Nasional Ilmu Sosial Dan Teknologi (SNISTEK)*, 4, 13–18.
 - [11] Syahputra, H., Syahril, M., Studi Mahasiswa, P., Triguna Dharma, S., & Studi Dosen Pembimbing, P. (2020). Prediksi Jumlah Murid Baru Dengan Menggunakan Metode Regresi Linear Berganda. *Jurnal CyberTech*, 3(4), 671–679. <https://ojs.trigunadharma.ac.id/>
 - [12] Syahrudin, A. N., & Kurniawan, T. (2018). Input dan output pada bahasa pemrograman python. *Jurnal Dasar Pemograman Python STMIK*, 20, 1–7.
 - [13] Syahza, A. (n.d.). *Metodologi Penelitian Edisi Revisi*.