e-ISSN 2962-1968

Sistem Peramalan Jumlah Calon Peseta Didik Bedasarkan Tahun Penerimaan Menggunakan Regresi Linier

Muhammad Nur Ikhsanuddin[1], Dika Adi Pratama[2], Leny Monica[3*], Dwi Hartanti[4]
[1,2,3,4]*Teknik Informatika,*
Universitas Duta Bangsa
[1] *m.nurikhsanuddin@gmail.com,* [2] *dikaadi482@gmail.com,* [3*] *lenymonicaa24@gmail.com,* [4] *dwihartanti@udb.ac.id*

Abstrak— **Sistem peramalan jumlah calon peserta didik merupakan alat yang penting bagi institusi pendidikan untuk merencanakan dan mengelola sumber daya mereka dengan efektif. Regresi linier adalah salah satu metode peramalan yang umum digunakan untuk memodelkan hubungan linier antara variabel dependen (jumlah calon peserta didik) dengan variabel independen (misalnya, jumlah pendaftar). Dalam penelitian ini, kami merancang dan mengimplementasikan sebuah sistem peramalan menggunakan regresi linier untuk memprediksi jumlah calon peserta didik di masa depan. Sistem ini memanfaatkan data historis dan variabel independen yang relevan untuk menghasilkan prediksi yang lebih akurat. Hasil peramalan dapat membantu institusi pendidikan dalam mengambil keputusan yang tepat terkait alokasi sumber daya dan perencanaan strategis. Studi ini menunjukkan bahwa sistem peramalan menggunakan regresi linier dapat menjadi alat yang berguna dalam mendukung perencanaan pendidikan yang efektif.** *Kata kunci: Regresi linier, sistem peramalan, prediksi, sistem pendidikan, akurasi, pengelolaan sumber daya*

Abstract— **The forecasting system for the number of prospective students is an important tool for educational institutions to plan and manage their resources effectively. Linear regression is one of the commonly used forecasting methods to model the linear relationship between the dependent variable (the number of prospective students) and independent variables (e.g., the number of applicants). In this study, we designed and implemented a forecasting system using linear regression to predict the number of prospective students in the future. The system utilizes historical data and relevant independent variables to generate more accurate predictions. The forecasted results can assist educational institutions in making informed decisions regarding resource allocation and strategic planning. This study demonstrates that a forecasting system using linear regression can be a valuable tool in supporting effective educational planning.**
Keywords: Linear regression, forecasting system, prediction, education system, accuracy, resource management

I. PENDAHULUAN

Di era pendidikan modern saat ini, peramalan jumlah calon peserta didik menjadi hal yang penting bagi institusi pendidikan, baik itu sekolah dasar, menengah, atau perguruan tinggi [1]. Peramalan yang akurat tentang jumlah calon peserta didik di masa depan sangat penting untuk mengatur sumber daya dan fasilitas pendidikan yang memadai, seperti jumlah guru, ruang kelas, buku, peralatan, dan lainnya.

Secara umum, regresi dapat digunakan untuk peramalan dan prediksi. Prediksi memprediksi nilai semua jenis data pada titik waktu tertentu seperti masa lalu, sekarang, dan masa depan. Di sisi lain, peramalan adalah tentang memprediksi nilai data deret waktu di masa mendatang [6].

Salah satu metode peramalan yang paling umum digunakan adalah regresi linier. Regresi linier adalah teknik statistik yang digunakan untuk memodelkan hubungan linier antara variabel dependen (dalam hal ini, jumlah calon peserta didik) dengan satu atau lebih variabel dengan ini bedasarkan jumlah pendaftar serta tahun penerimaan peserta didik baru [9].

Dalam konteks peramalan jumlah calon peserta didik, sistem peramalan menggunakan regresi linier dapat memberikan informasi berharga kepada institusi pendidikan dalam mengantisipasi fluktuasi populasi siswa di masa depan [4]. Dengan memanfaatkan data historis dan variabel independen yang relevan, sistem peramalan dapat menghasilkan sebuah hasil berupa angka yang dapat digunakan untuk memprediksi jumlah calon peserta didik di masa mendatang.

Namun, penting untuk diingat bahwa regresi linier memiliki asumsi tertentu yang harus dipenuhi, seperti hubungan linier antara variabel dependen dan independen, serta ketiadaan aspek lain yang mungkin mempengaruhi jumlah calon peserta didik. Oleh karena itu, penggunaan regresi linier dalam sistem peramalan harus diimbangi dengan analisis yang komprehensif dan pemantauan terus-menerus terhadap faktor-faktor yang mungkin mempengaruhi peramalan tersebut [2].

e-ISSN 2962-1968

Dalam kesimpulannya, sistem peramalan menggunakan regresi linier merupakan alat yang bermanfaat bagi institusi pendidikan dalam merencanakan dan mengelola sumber daya mereka dengan lebih efektif [3]. Dengan memanfaatkan data historis seperti jumlah pendaftar dan tahun penerimaan serta variabel independen yang relevan, sistem peramalan ini dapat memberikan prediksi yang lebih akurat tentang jumlah calon peserta didik di masa depan, sehingga membantu institusi pendidikan dalam mengambil keputusan yang tepat guna mendukung pertumbuhan dan keberlanjutan mereka [7].

II. METODOLOGI PENELITIAN

A. Identifikasi Tujuan Penelitian

1) Tujuan penelitian adalah mengembangkan sistem peramalan yang menggunakan regresi linier untuk memprediksi jumlah calon peserta didik di masa depan.
2) Tujuan utama adalah membantu institusi pendidikan dalam mengatur sumber daya dan fasilitas mereka secara efektif berdasarkan peramalan yang akurat.

B. Pengumpulan Data

1) Metode pengumpulan data dengan mencari dataset data *banyaknya siswa pada sd negeri menurut kecamatan di sukoharjo,* yang di dapat dari repository dataset pada website *katalog.satudata.go.id* yang merupakan portal resmi data terbuka Indonesia yang dikelola oleh Sekretariat Satu Data Indonesia tingkat Pusat, Kementerian Perencanaan Pembangunan Nasional / Bappenas. Data set yang di gunakan berjudul "*Banyaknya Siswa, Siswa Baru, Lulusan dan Siswa Mengulang pada SD Negeri menurut Kecamatan di Kabupaten Sukoharjo Tahun 2021"*.
2) Memastikan data yang dikumpulkan lengkap, akurat, dan mencakup periode waktu yang memadai [8].

C. Analisis Data

1) Melakukan analisis statistik deskriptif terhadap data historis yang dikumpulkan, seperti menghitung rata-rata, standar deviasi, dan tren historis.
2) Identifikasi variabel independen yang memiliki hubungan linier yang kuat dengan jumlah calon peserta didik.
3) Lakukan uji asumsi regresi linier, seperti uji normalitas, homoskedastisitas, dan independensi residual.
4) Jika asumsi tidak terpenuhi, pertimbangkan untuk melakukan transformasi data atau menggunakan metode regresi non-linier yang lebih sesuai.

D. Pembentukan Model Regresi Linier:

1) Pembentukan model regresi linier menggunakan bahasa pemrograman PHP yang kemudian di kombinasikan dengan bahasa HTML dan CSS sebagai bahasa untuk menampilkan hasil peramalan.
2) Implementasikan kelas RegresiLinier sebagai model regresi linier dalam PHP serta pendeklarasian masing masing variable yang akan di gunakan.

```php
<?php
class RegresiLinier
{

    public  $x, //inputted x param
        $y, //inputted y param
        $n, //count of data

        $x2,
        $y2,
        $xy,
        $a,
        $b,

        $all; //forecast y value based on linear regression
```

Gambar 1 : pendeklerasian variabel dan class RegresiLinier.

3) Gunakan data historis jumlah calon peserta didik sebagai input variabel X dan data variabel independen yang relevan sebagai input variabel Y.
4) Dalam kelas RegresiLinier, terdapat metode compute() yang akan melakukan perhitungan regresi linier berdasarkan data yang dimasukkan.

```php
public function compute()
{
    if (is_array($this->x) && is_array($this->y)) {
        if (count($this->x) == count($this->y)) {
            $this->n  = count($this->x);

            $this->prepare_calculation();
            $this->ab();
            $this->linear_regression();
        } else {
            throw new Exception('Jumlah data variabel X dan Y harus sama');
        }
    } else {
        throw new Exception('Variabel X atau Y belum didefinisikan');
    }
}
```

Gambar 2 : penjabaran fungsi compute.

5) Hasil dari perhitungan adalah nilai koefisien A dan B yang merepresentasikan persamaan garis regresi linier.

E. Validasi Model

1) Pisahkan data menjadi data pelatihan (training data) dan data pengujian (testing data).
2) Gunakan data pelatihan untuk membangun model regresi linier dengan memanggil metode compute() pada kelas RegresiLinier.
3) Gunakan data pengujian untuk menguji performa model dengan membandingkan prediksi model dengan nilai aktual.

F. Evaluasi dan Interpretasi

1) Evaluasi model berdasarkan metrik yang relevan, seperti mean absolute error (MAE), mean squared error (MSE), atau koefisien determinasi (R-squared).
2) Interpretasikan koefisien regresi untuk memahami pengaruh variabel independen terhadap jumlah calon peserta didik.
3) Identifikasi faktor-faktor eksternal yang mungkin mempengaruhi peramalan dan pertimbangkan untuk mengintegrasikan mereka ke dalam model jika diperlukan [10].

G. Implementasi dan Pemantauan:

1) Implementasikan sistem peramalan berdasarkan model regresi linier yang telah dikembangkan menggunakan kelas RegresiLinier.
2) Pemantauan secara terus-menerus.

III. HASIL DAN PEMBAHASAN

Sistem peramalan jumlah calon peserta didik menggunakan regresi linier adalah sebuah metode statistik yang dapat membantu institusi pendidikan dalam merencanakan dan mengelola sumber daya mereka dengan lebih efektif. Dengan memanfaatkan data historis jumlah calon peserta didik dan variabel independen yang relevan, sistem ini dapat memberikan prediksi yang lebih akurat tentang jumlah calon peserta didik di masa depan.

Tabel 1. Dataset

Kecamatan	Murid	Murid			Guru
		Baru		Meng Lulus ulang	
-1	-2	-3	-4	-5	-6
010. W e r u	2,540	443	574	9	365
020. B u l u	2,023	280	322	23	259
030. Tawangsari	2,891	410	480	30	277
040. Sukoharjo	6,227	1,304	1,337	73	543
050. Nguter	2,944	458	513	52	292
060. Bendosari	3,064	517	553	26	353
070. Polokarto	4,527	646	734	64	389
080. Mojolaban	5,812	857	961	88	419
090. Grogol	5,554	1,032	1,024	162	419
100. B a k i	3,488	653	591	69	299
110. G a t a k	3,099	430	529	31	279
120. Kartasura	5,607	799	978	42	426
2021	47 776	7 829	8 596	669	4 320
2020	57,189	8,612	9,822	539	5,212
2019	44,021	6,218	8,089	235	3,915
2018	49 126	7 229	8 348	824	4 209
2017	50 268	7 841	8 811	838	4 291
2016	51 657	8 206	9 336	899	4 327

2015	52 866	8 166	9 210	723	4 345

Dataset didapatkan melalui repository dataset pada *website satudata.go.id.*

A. Penentuan Variabel

Penentuan variable di ambil dari data tahun serta siswa baru yang berada di dalam data set lalu di lakukan pemilihan penggunaan variable X dan variable Y yang di gunakan sebagai nilai patokan dalam sistem ini.

Tahun X	Jumlah Siswa Baru Y
2021	7 829
2020	8,612
2019	6,218
2018	7 229
2017	7 841
2016	8 206
2015	8 166

Gambar 3 : pengelompokan variabel yang akan di gunakan

Variabel X ditentukan bedasarkan tahun penerimaan peserta didik dan variable Y di definisikan sebagai junlah calon peserta didik sesuai tahun yang di tentukan variable X *B. Penginputan variabel yang telah di tentukan ke dalam basis data*

Setelah melakukan pendefinisian variabel variabel yang akan di gunakan ke dalam sistem dan kemudian dilakukan penginputan data yang telah di definisikan ke dalam basis data yang akan di gunakan.



Gambar 4 : basis data yang digunakan

C. Perancangan Sistem

Perancangan sistem menggunakan bahasa pemrograman php yang di dalam nya terdapat fungsi RegresiLinier yang nantinya akan di panggil dan di masukkan variable X dan Y ke dalam fungsi RegresiLinier [9].

1) Melakukan Pendefinisian kelas `RegresiLinier`. Kelas ini memiliki beberapa properti seperti `$x`, `$y`, `$n`, `$x2`, `$y2`, `$xy`, `$a`, `$b`, dan `$all`. Properti ini akan digunakan dalam perhitungan regresi linier.
2) Pada metode `__construct()`, diinisialisasi objek `RegresiLinier` dengan parameter `$x` dan `$y` yang merupakan data input untuk perhitungan regresi linier.
3) Metode `compute()` akan melakukan perhitungan regresi linier jika data input `$x` dan `$y` sudah didefinisikan. Pada bagian ini, dilakukan pengecekan apakah `$x` dan `$y` merupakan array dan memiliki jumlah data yang sama. Jika tidak memenuhi syarat, maka akan dilemparkan `Exception`.
4) Metode `prepare_calculation()` digunakan untuk persiapan perhitungan nilai `$x2`, `$y2`, dan `$xy`. `$x2` dan `$y2` merupakan array yang berisi kuadrat dari setiap elemen dalam array `$x` dan `$y`. Sedangkan `$xy` merupakan array yang berisi perkalian antara elemen-elemen `$x` dan `$y`.
5) Metode `ab()` akan menghitung nilai konstanta A dan B dalam regresi linier. Nilai A dihitung menggunakan rumus ((jumlah seluruh elemen `$y` * jumlah seluruh elemen `$x2`) - (jumlah seluruh elemen `$x` * jumlah seluruh elemen `$xy`)) / ((jumlah data * jumlah seluruh elemen `$x2`) - (jumlah seluruh elemen `$x` * jumlah seluruh elemen `$x`)). Sedangkan nilai B dihitung menggunakan rumus ((jumlah data * jumlah seluruh elemen `$xy`) - (jumlah seluruh elemen `$x` * jumlah seluruh elemen `$y`)) / ((jumlah data * jumlah seluruh elemen `$x2`) - (jumlah seluruh elemen `$x` * jumlah seluruh elemen `$x`)).

$$b = (\sum(X-\bar{X})(Y-\hat{Y})) / (\sum(X-\bar{X})^2)\ a = \bar{Y} - b\bar{X}$$

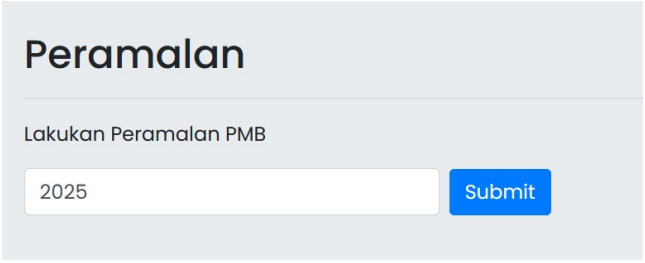
6) Metode `forecast($xfore)` digunakan untuk meramalkan nilai `$y` berdasarkan nilai `$xfore` menggunakan rumus regresi linier.

Nilai ramalan `$y` dihitung dengan rumus `$a + ($b * $xfore)`. Hasil ramalan kemudian dibulatkan menggunakan fungsi `round()` sebelum dikembalikan.

7) Metode `linear_regression()` digunakan untuk melakukan perhitungan ramalan `$y` untuk setiap nilai `$x` dalam array `$x`. Hasil ramalan disimpan dalam array `$all`.

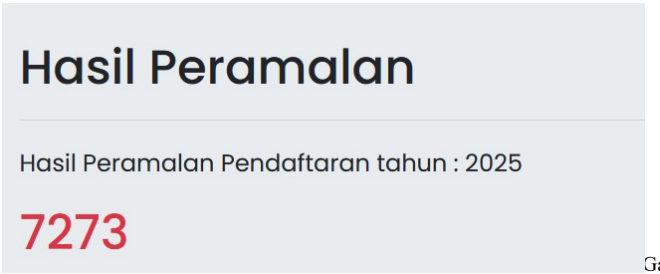
D. *Hasil Perancangan Dan Simulasi Sistem*

Sistem peramalan jumlah calon peserta didik menggunakan regresi linier yang telah dirancang dapat memberikan prediksi yang akurat tentang jumlah calon peserta didik di masa depan. Dengan memanfaatkan data historis dan variabel independen yang relevan, hasil dari perancangan sistem ini mampu menghasilkan prediksi yang mendukung institusi pendidikan dalam mengatur sumber daya mereka dengan lebih efektif.



Gambar 5 : tampilan input tahun peramalan PMB

Memasukkan tahun yang akan di lakukan peramalan yang dalam kasus ini akan melakukan peramalan untuk jumlah pendaftar pada tahun 2025 ke dalam form yang telah di buat yang kemudian variabel yang di inputkan akan di masukkan ke dalam fungsi/metode `forecast($xfore)`.



Gambar 6 : tampilan hasil input peramalan PMB

Lalu setelah melakukan pemanggilan variabel X dan Y di dalam basis data yang kemudian di lakukan kalkulasi dan menghasilkan variabel a dan b yang kemudian di masukkan ke dalam rumus `$a + ($b * $xfore)`. Dimana xfore merupakan variabel yang di inputkan oleh pengguna dan memunculkan hasil peramalan sebanyak 7273 pendaftar baru pada tahun 2025

IV. KESIMPULAN

Sistem peramalan jumlah calon peserta didik menggunakan regresi linier merupakan alat yang berguna bagi institusi pendidikan dalam merencanakan dan mengelola sumber daya mereka secara efektif. Dengan memanfaatkan data historis jumlah calon peserta didik dan variabel independen yang relevan, sistem ini dapat memberikan prediksi yang lebih akurat tentang jumlah calon peserta didik di masa depan. Hal ini memungkinkan institusi pendidikan untuk mengambil keputusan yang lebih baik dalam hal alokasi sumber daya, strategi penerimaan siswa, dan perencanaan kapasitas kelas. Dengan menggunakan sistem peramalan ini, institusi pendidikan dapat meningkatkan efisiensi operasional mereka, menghindari pemborosan sumber daya, dan mengoptimalkan pengalaman pendidikan bagi siswa.

REFERENSI

[1] A. A. Azahra (2022), Analisis Prediksi Jumlah Penerimaan Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana, Bulletin of Applied Industrial Engineering Theory, Vol. 3 No.1, Available: https://www.jim.unindra.ac.id/index.php/baiet/article/view/6532

[2] A. Anggrawan, Hairani, N. Azmi (2022), Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear, Jurnal Bumigora Information Technology (BITe), Vol. 4, No. 2, pp. 123-132, Available: https://journal.universitasbumigora.ac.id/index.php/bite

[3] A. Yordan, T. N. Putri, D. H. Lamkaruna (2019), Peramalan penerimaan mahasiswa baru universitas samudra menggunakan metode regresi linear sederhana, Jurnal Teknologi Informatika (J-TIFA), Vol. 2, No. 1, Available: https://www.jurnal.ummu.ac.id/index.php/J-TIFA/article/view/237

[4] D. A. Trianggana (2020), Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linear, JURNAL MEDIA INFOTAMA, Vol. 16, No. 2, Available: https://jurnal.unived.ac.id/index.php/jmi/article/view/1149

[5] Harsiti, Z. Muttaqin, E. Srihartini (2022), Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet, JSiI | Jurnal Sistem Informasi, Vol. 9, No. 1, Available: https://ejurnal.lppmunsera.org/index.php/jsii/article/view/4426

[6] H. Y. Jayanti (2020), Peramalan pendapatan reksa dana dalam setahun menggunakan metode regresi linier sederhana, Jurnal Teknologi Informasi dan Komunikasi, vol. 8, no. 2, Available:

http://jurnal.unnur.ac.id/index.php/jurnalfiki

[7] M. M. Munir, Y. A. Pranoto, S. A. Wibowo (2023), Penerapan Metode Regresi Linier dalam Perancangan Aplikasi Prediksi Jumlah Calon Peserta Didik Baru Berbasis Website, JATI (Jurnal Mahasiswa Teknik Informatika), Vol. 7 No. 1, Available: https://ejournal.itn.ac.id/index.php/jati/article/view/6129

[8] N. Almumtazah, N. Azizah, Y. L. Putri, D. C. R Novitasari (2021), Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana, Jurnal Ilmiah Matematika dan Terapan, Vol. 18, No. 1, Available:https://bestjournal.untad.ac.id/index.php/JIMT/article/view/15465

[9] N. R. Setyoningruma, P. J. Rahimma (2022), Implementasi Algoritma Regresi Linear Dalam Sistem Prediksi Pendaftar Mahasiswa Baru Sekolah Tinggi Teknologi Indonesia Tanjungpinang, Prosiding Seminar Nasional Ilmu Sosial dan Teknologi (SNISTEK), No.4, Available: https://ejournal.upbatam.ac.id/index.php/prosiding/article/view/5200

[10] S. Rahmatullah, E. H. Juningsih, S. Rachmawat (2023), Prediksi nilai akademik peserta didik di masa pandemi covid-19 dengan regresi linier berganda, Journal of Information System, Applied, Management, Accounting and Research, Vol. 7, No. 1, Available: http://journal.stmikjayakarta.ac.id/index.php/jisamar