

Implementasi Metode Regresi Linear Berganda Dalam Memprediksi Santri Baru Di Pondok Pesantren Al-Falah Sumber Gayam

Khalishatus Shafariyah¹, Siti Khotijah²

^{1,2}Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Annuqayah Guluk-Guluk, Sumenep, Jawa Timur, Indonesia

Article Info

Article history:

Received January 2, 2025

Revised January 15, 2025

Accepted January 30, 2025

Keywords:

Al-Falah Islamic Boarding School

Source of Style

Prediction

Multiple Linear Regression

ABSTRACT

Al-Falah Islamic Boarding School Sumber Gayam, Kec. Kadur, Kab. Pamekasan is one of the boarding schools that fluctuates in terms of the number of new students each year. Therefore, predictions are very necessary to find out the number of new students who will register in the following year. The aim is to help the committee for accepting new students to be ready to prepare the boarding facilities as well as possible. The multiple linear regression method is the method used in this research to predict future conditions using previous data and to determine whether or not there is an influence exerted by the independent variable on the dependent variable. Annual student alumni data, boarding school facilities, annual fees and data on outstanding students are variables that are thought to influence the number of new students in 2023-2024 by calculating MAPE and RMSE errors. The results of the research obtained as many as 98 new students based on testing the r^2 , r , F test and t test based on the sig value. < 0.05 which was calculated manually and SPSS version 26 calculations. The results show that the variables studied have a significant effect on the number of new students registering in 2023-2024 with a MAPE accuracy of 0.02% and an RMSE accuracy of 0.66%, this shows that the ability of the forecasting model formed can be said to be good or accurate.

Corresponding Author:

Siti Khotijah

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam

Institut Sains dan Teknologi Annuqayah Guluk-Guluk, Sumenep, Jawa Timur, Indonesia

Email: siti@istannuqayah.ac.id

1. PENDAHULUAN

Pendidikan memiliki peran penting untuk mencerdaskan setiap warga negara Indonesia, karena pendidikan merupakan prioritas utama suatu negara. Sistem pendidikan diharapkan memiliki sarana dan prasarana yang memadai sehingga terbentuklah suatu pendidikan yang berkualitas [1]. Pada dasarnya, pendidikan berbasis agama berkaitan dengan pendidikan berbasis umum. Berbicara tentang pendidikan agama terutama agama Islam, erat kaitannya dengan keberadaan pondok pesantren yang dapat direpresentasikan sebagai tempat kehidupan santri dimana pendidikan di pesantren meliputi pendidikan Islam yang menyeluruh baik ilmu tentang keislaman maupun ilmu yang bersifat umum [2]. Penerimaan santri baru merupakan kegiatan rutin yang dilakukan oleh seluruh lembaga pendidikan di Indonesia setiap tahunnya, salah satunya adalah Pondok Pesantren Al-Falah Sumber Gayam Kec. Kadur, Kab. Pamekasan. Jumlah santri baru di Pondok Pesantren Al-Falah Sumber Gayam berfluktuasi setiap tahunnya sehingga hal ini akan menimbulkan suatu permasalahan terkait banyaknya jumlah santri baru yang akan mendaftar untuk tahun berikutnya. Dari kondisi yang fluktuatif tersebut, prediksi menjadi hal yang sangat dibutuhkan untuk mengetahui gambaran masa depan apakah santri yang mendaftar akan mengalami kenaikan ataupun penurunan.

Prediksi atau peramalan merupakan proses perkiraan atau dugaan mengenai kemungkinan suatu peristiwa di masa depan dengan memeriksa keadaan masa lalu [3].

Ada beberapa faktor yang diduga dapat mempengaruhi jumlah santri baru yaitu santri alumni pertahun, fasilitas, biaya pondok, santri berprestasi dan santri baru pada 10 tahun terakhir yang merupakan data Pondok Pesantren Al-Falah Sumber Gayam sebagai data acuan untuk melakukan prediksi terhadap santri baru yang akan mendaftar pada tahun selanjutnya.

Adapun metode yang digunakan dalam penelitian ini untuk meramalkan jumlah santri baru yang akan mendaftar pada tahun ajaran berikutnya adalah metode regresi linear berganda, karena selain dapat memprediksi metode ini juga dapat mengetahui pengaruh antara variabel *independent* terhadap variabel *dependent*.

Berdasarkan penelitian sebelumnya, Nainggolan, dkk. (2021) melakukan penelitian yang berjudul *Implementasi Data Mining dengan Metode Regresi Linear Berganda dalam Estimasi Mahasiswa Baru pada Sekolah Tinggi Agama Islam Raudhatul Akmal Batang Kuis* dimana hasil penelitiannya diperoleh hasil sebanyak 283 orang mahasiswa pada tahun 2019-2020 yang dihitung secara manual dan berupa tampilan form prediksi.

Penelitian Aryani (2020) juga melakukan estimasi atau prediksi terhadap pendapatan perusahaan dengan melakukan uji t, uji F dan R^2 berdasarkan nilai $\text{sig} < 0,05$ melalui perhitungan SPSS versi 21, hasil yang diperoleh menunjukkan bahwa variabel yang diteliti berpengaruh positif dan signifikan terhadap pendapatan perusahaan. Oleh karena itu, peneliti tertarik untuk melakukan penelitian yang berjudul "Implementasi Metode Regresi Linear Berganda dalam Memprediksi Santri Baru di Pondok Pesantren Al Falah Sumber Gayam" dengan tujuan agar pihak panitia penerimaan santri baru memiliki kesiapan dalam menyediakan fasilitas pondok dengan sebaik baiknya.

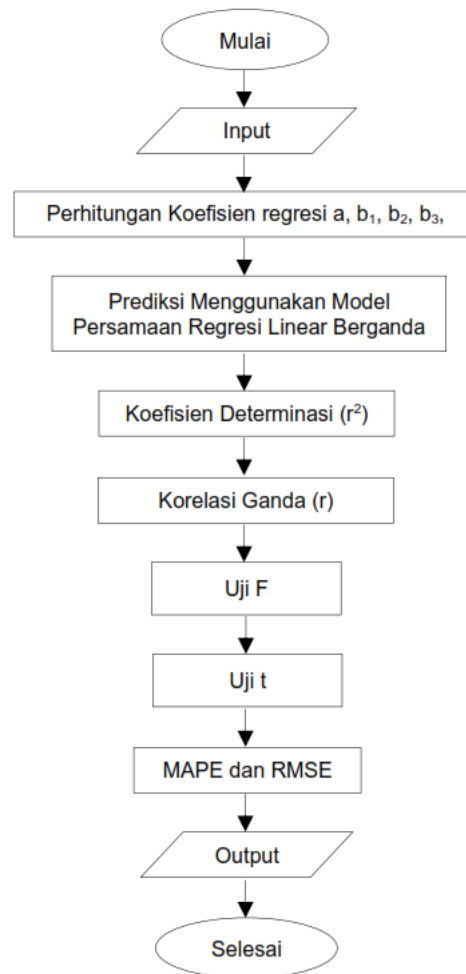
2. METODE

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif yang bertujuan untuk menguji hipotesis dengan menggunakan teori-teori yang sudah ada. Adapun jenis pendekatannya adalah pendekatan kuantitatif deskriptif yang digunakan untuk membuat gambaran tentang suatu keadaan secara objektif yang menggunakan angka, penafsiran terhadap data tersebut serta penampilan dan hasilnya [4].

Adapun pelaksanaan penelitian ini bertempat di Pondok Pesantren Al-Falah Sumber Gayam, Jl. Sumber Gayam, Kadur Timur, Kadur, Kec. Kadur, Kabupaten Pamekasan, Jawa Timur. Sedangkan waktu penelitian dilakukan dalam jangka waktu 3 bulan yaitu mulai dari bulan Mei 2023 sampai Juli 2023. Metode yang digunakan dalam penelitian ini adalah metode regresi linear berganda, Dimana model dari persamaan tersebut akan digunakan untuk melakukan prediksi. Tahapan-tahapannya adalah sebagai berikut:

1. Masukkan variabel bebas dan variabel terikat yaitu jumlah santri alumni pertahun sebagai X_1 , fasilitas pondok sebagai X_2 , biaya pertahun sebagai X_3 , santri berprestasi sebagai X_4 dan santri baru sebagai Y .
2. Mencari koefisien regresi yaitu nilai a, b_1, b_2, b_3 dan b_4 yang digunakan untuk mendapatkan model persamaan regresi linear berganda.
3. Menghitung hasil prediksi menggunakan model persamaan regresi linear berganda yang sudah diperoleh. Prediksi ini dilakukan dengan cara menggunakan variabel bebas pada tahun terakhir yaitu pada tahun 2022.
4. Menghitung koefisien determinasi untuk mengetahui seberapa besar kemampuan variable independen X_1, X_2, X_3 dan X_4 dalam model regresi dalam menjelaskan variabel dependen Y .
5. Menghitung korelasi ganda, hal ini dilakukan untuk mengetahui arah dan kuatnya hubungan antar dua variabel atau lebih.
6. Melakukan uji F. Uji ini dilakukan untuk menguji dan melihat pengaruh antara variabel bebas secara simultan terhadap variabel tak bebas secara signifikan.
7. Melakukan uji-t dengan tujuan untuk melihat pengaruh dari variabel bebas secara parsial terhadap variabel tak bebas secara signifikan.
8. Menghitung nilai akurasi yaitu dengan menggunakan MAPE dan RMSE untuk mengetahui selisih atau kesalahan hasil prediksi terhadap nilai yang sebenarnya.
9. Menampilkan hasil prediksi.

Adapun tahapan untuk memprediksi dengan menggunakan metode regresi linear berganda dapat dilihat pada diagram alir berikut:



Gambar 1. Diagram alir metode regresi linier berganda

3. HASIL DAN PEMBAHASAN

3.1 Data Santri

Data yang digunakan dalam penelitian ini adalah data santri Pondok Pesantren Al-Falah Sumber Gayam Kec. Kadur, Kab. Pamekasan berdasarkan data historis 10 tahun terakhir yaitu dari tahun 2013-2022 dimana variabel bebas X_1 sebagai Santri Alumni Pertama, X_2 sebagai Fasilitas Pondok, X_3 sebagai Biaya Pertama, X_4 sebagai Santri Berprestasi dan variabel terikat Y sebagai Santri Baru. Adapun data data tersebut akan disajikan dalam bentuk tabel berikut:

Tabel 1. Data santri pondok pesantren Al-Falah Sumber Gayam

Periode	Santri Alumni	Fasilitas Pondok	Biaya Pertama	Santri Berprestasi	Santri Baru
	X_1	X_2	X_3	X_4	Y
2013	55	5	285000	3	70
2014	61	5	300000	5	56
2015	53	5	300000	3	74
2016	43	5	300000	4	81
2017	57	6	345000	2	86
2018	55	6	345000	5	77
2019	71	7	345000	3	73
2020	53	7	370000	4	92
2021	50	7	370000	5	90
2022	43	7	370000	5	97
Σ	541	60	3330000	39	796

3.2 Menghitung konstanta dan koefisien regresi

Berdasarkan penentuan variabel bebas dan variabel tak bebas di atas, maka langkah selanjutnya adalah menghitung nilai konstanta dan koefisien regresi dengan cara mencari nilai $\sum X_1^2, \sum X_2^2, \sum X_3^2, \sum X_4^2, \sum Y^2, \sum X_1Y, \sum X_2Y, \sum X_3Y, \sum X_4Y, \sum X_1X_2, \sum X_1X_3, \sum X_1X_4, \sum X_2X_3, \sum X_2X_4$ dan $\sum X_3X_4$ menggunakan tabel bantu koefisien regresi.

Tabel 2. Tabel bantu koefisien regresi

No	X_1^2	X_2^2	X_3^2	X_4^2	Y^2	X_1Y	X_2Y
1	3025	25	81225000000	9	4900	3850	350
2	3721	25	90000000000	25	3136	3416	280
3	2809	25	90000000000	9	5476	3922	370
4	1849	25	90000000000	16	6561	3483	405
5	3249	36	119025000000	4	7396	4902	516
6	3025	36	119025000000	25	5929	4235	462
7	5041	49	119025000000	9	5329	5183	511
8	2809	49	136900000000	16	8464	4876	644
9	2500	49	136900000000	25	8100	4500	630
10	1849	49	136900000000	25	4171	4171	679
Σ	29877	368	1119000000000	163	42538	42538	4847

Tabel 2. Tabel bantu koefisien regresi

No	X_3Y	X_4Y	X_1X_2	X_1X_3	X_1X_4	X_2X_3	X_2X_4	X_3X_4
1	19950000	210	275	15675000	165	1425000	15	855000
2	16800000	280	305	18300000	305	1500000	25	1500000
3	22200000	222	265	15900000	159	1500000	15	900000
4	24300000	324	215	12900000	172	1500000	20	1200000
5	29670000	172	342	19665000	114	2070000	12	690000
6	26565000	385	330	18975000	275	2070000	30	1725000
7	25185000	219	497	24495000	213	2415000	21	1035000
8	34040000	368	371	19610000	212	2590000	28	1480000
9	33300000	450	350	18500000	250	2590000	35	1850000
10	35890000	445	301	15910000	215	2590000	35	1850000
Σ	267900000	3115	3251	179930000	2080	20250000	236	13085000

Kemudian untuk memperoleh koefisien regresi a, b_1, b_2, b_3 dan b_4 adalah dengan melakukan proses perhitungan regresi linear berganda yaitu dengan cara eliminasi menggunakan persamaan berikut [5]:

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 + b_4 \sum X_4 \quad (1)$$

$$\sum X_1Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1X_2 + b_3 \sum X_1X_3 + b_4 \sum X_1X_4 \quad (2)$$

$$\sum X_2Y = a \sum X_2 + b_1 \sum X_1X_2 + b_2 \sum X_2^2 + b_3 \sum X_2X_3 + b_4 \sum X_2X_4 \quad (3)$$

$$\sum X_3Y = a \sum X_3 + b_1 \sum X_1X_3 + b_2 \sum X_2X_3 + b_3 \sum X_3^2 + b_4 \sum X_3X_4 \quad (4)$$

$$\sum X_4Y = a \sum X_4 + b_1 \sum X_1X_4 + b_2 \sum X_2X_4 + b_3 \sum X_3X_4 + b_4 \sum X_4^2 \quad (5)$$

Berdasarkan tabel 2 dan tabel 3, maka diperoleh persamaan:

$$10a + 54b_1 + 60b_2 + 3330000b_3 + 39b_4 = 796 \quad (6)$$

$$541a + 29877b_1 + 3251b_2 + 179930000b_3 + 2080b_4 = 796 \quad (7)$$

$$3330000a + 179930000b_1 + 20250000b_2 + 1119000000000b_3 + 13085000b_4 = 267900000 \quad (8)$$

$$39a + 2080b_1 + 236b_2 + 13085000b_3 + 163b_4 = 796 \quad (9)$$

Untuk memperoleh nilai a, b_1, b_2, b_3 dan b_4 adalah dengan cara mengeliminasi persamaan (6), (7), (8), (9) dan (10). Dengan mengaplikasikan seluruh data, maka hasil yang diperoleh adalah sebagai berikut:

$$a = 70,7627885505303 \approx 70,763$$

$$b_1 = -1,06584933412433 \approx -1,066$$

$$b_2 = 5,546495227196812 \approx 5,546$$

$$b_3 = 0,0001500367009824364 \approx 0,000$$

$$b_4 = -4,32090389977076 \approx -4,321$$

Sehingga diperoleh nilai konstanta $a = 70,763, b_1 = -1,066, b_2 = 5,546, b_3 = 0$ dan $b_4 = -4,321$.

3.3 Prediksi dengan metode Regresi Linier Berganda

Berdasarkan proses perhitungan nilai konstanta dan koefisien regresi di atas, maka langkah selanjutnya adalah membentuk model persamaan regresi dengan 4 variabel yaitu sebagai berikut:

$$\hat{Y} = 70,7627885505303 + (-1,06584933412433)X_1 + 5,546495227196812X_2 + 0,0001500367009824364X_3 + (-4,32090389977076)X_4$$

Dari persamaan regresi yang telah dibentuk, maka untuk memperhitungkan jumlah santri baru dapat dihasilkan dengan memasukkan nilai X_1, X_2, X_3 dan X_4 Pada periode terakhir (tahun 2022) yaitu dengan nilai $X_1 = 43, X_2 = 7, X_3 = 370000$ dan $X_4 = 5$:

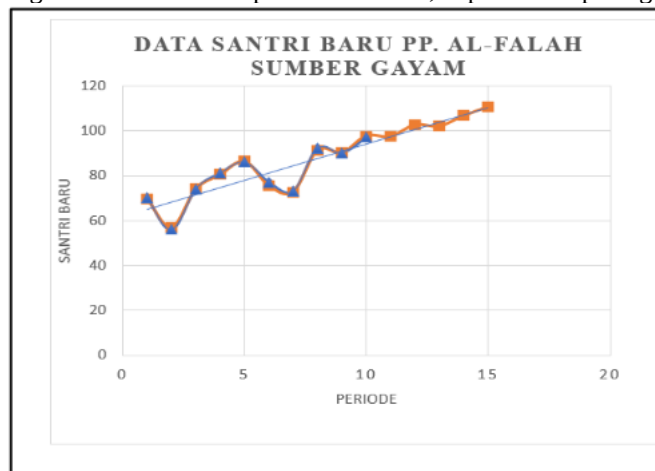
$$\hat{Y} = 70,7627885505303 + (-1,06584933412433)X_1 + 5,546495227196812X_2 + 0,0001500367009824364X_3 + (-4,32090389977076)X_4$$

$$\hat{Y} = 70,7627885505303 + (-1,06584933412433) \times 43 + 5,546495227196812 \times 7 + 0,0001500367009824364 \times 370000 + (-4,32090389977076) \times 5$$

$$\hat{Y} = 70,7627885505303 - 45,83152136734619 + 38,82546659037768 + 55,635793635028 - 21,6045194988538$$

$$\hat{Y} = 97,7880079097359 \approx 98$$

Jadi, hasil prediksi jumlah santri baru di Pondok Pesantren Al-Falah Sumber Gayam pada tahun 2023 sebanyak 98 santri baru yang berarti mengalami peningkatan jumlah santri baru dibanding tahun 2022 yaitu bertambah satu orang. Berdasarkan hasil prediksi tersebut, dapat dilihat pada gambar berikut:



Gambar 2. Plot data aktual dan data prediksi PP. Al-Falah Sumber Gayam

3.4 Prediksi dengan metode Regresi Linier Berganda

Koefisien determinasi digunakan untuk mengetahui seberapa besar kemampuan variabel independen X_1, X_2, X_3 dan X_4 dalam model regresi dalam menjelaskan variabel dependen yaitu [6]:

$$\sum X_1 Y = \sum X_1 Y - \frac{\sum X_1 \times \sum Y}{n} = 42538 - \frac{541 \times 796}{10} = -525,6$$

$$\sum X_2 Y = \sum X_2 Y - \frac{\sum X_2 \times \sum Y}{n} = 4847 - \frac{60 \times 796}{10} = 71$$

$$\sum X_3 Y = \sum X_3 Y - \frac{\sum X_3 \times \sum Y}{n} = 267900000 - \frac{3330000 \times 796}{10} = 2832000$$

$$\sum X_4 Y = \sum X_4 Y - \frac{\sum X_4 \times \sum Y}{n} = 3115 - \frac{39 \times 796}{10} = 10,6$$

$$\sum Y^2 = \sum Y^2 - \frac{(\sum Y)^2}{n} = 64700 - \frac{796 \times 796}{10} = 1338,4$$

$$r^2 = \frac{(b_1 \sum X_1 Y) + (b_2 \sum X_2 Y) + (b_3 \sum X_3 Y) + (b_4 \sum X_4 Y)}{\sum Y^2}$$

$$= \frac{(-1,066) \times (-525,6) + 5,55 \times 71 + 0,00015 \times 2832000 + (-4,32) \times 10,6}{1338,8}$$

$$= \frac{560,21 + 393,8 + 425,8393 - 45,8}{1338,8}$$

$$= \frac{1334,049}{1338,8}$$

$$= 0,9967 \approx 0,997$$

Sehingga dapat disimpulkan bahwa nilai r^2 adalah 0.997 yang berarti 99.7% variabel santri baru dipengaruhi oleh variabel santri alumni pertahun, fasilitas pondok, biaya pertahun dan santri berprestasi, sedangkan untuk 0.3% dipengaruhi oleh variabel yang lain. Hal ini menunjukkan nilai koefisien determinasinya mendekati nilai 1 yang berarti variabel-variabel *independent* memiliki kemampuan untuk memberikan semua informasi yang dibutuhkan untuk memprediksi variabel *dependent*.

3.5 Koefisien korelasi ganda (r)

Koefisien korelasi ganda digunakan untuk mengetahui arah dan kuatnya hubungan antar dua variabel atau lebih. Adapun perhitungannya adalah sebagai berikut [7]:

$$\begin{aligned}
 r = \sqrt{r^2} &= \sqrt{\frac{(b_1 \sum X_1 Y) + (b_2 \sum X_2 Y) + (b_3 \sum X_3 Y) + (b_4 \sum X_4 Y)}{\sum Y^2}} \\
 &= \sqrt{\frac{(-1,066) \times (-525,6) + 5,55 \times 71 + 0,00015 \times 2832000 + (-4,32) \times 10,6}{1338,8}} \\
 &= \sqrt{\frac{1334,049}{1338,8}} \\
 &= \sqrt{0,9967} \\
 &= 0,9987 \approx 0,998
 \end{aligned}$$

Diperoleh nilai r adalah 0.998 yang berarti 99.8% variabel santri baru dipengaruhi oleh variabel santri alumni pertahun, fasilitas pondok, biaya pertahun dan santri berprestasi, sedangkan untuk 0.2% dipengaruhi oleh variabel yang lain. Hal ini menunjukkan bahwa nilai koefisien determinasinya mendekati nilai 1. Jadi, dapat disimpulkan bahwa terdapat pengaruh positif dan korelasi variabelvariabel yang diuji sangat kuat.

3.6 Uji simultan (uji F)

Koefisien korelasi ganda Uji F digunakan untuk mengetahui pengaruh variabel *independent* secara simultan terhadap variabel *dependent* yang terdapat pada model regresi. Kriteria keputusan yang menjadi penilaian untuk uji F adalah dengan taraf signifikan $\alpha = 0.05$. Apabila $F_{hitung} < F_{tabel}$ maka H_0 diterima dan H_1 ditolak, artinya tidak terdapat pengaruh yang signifikan antara variabel bebas secara simultan terhadap variabel terikat. Begitu juga sebaliknya, apabila $F_{hitung} > F_{tabel}$ maka H_0 ditolak dan H_1 diterima, artinya terdapat pengaruh yang signifikan antara variabel bebas secara simultan terhadap variabel terikat. Adapun perhitungannya adalah sebagai berikut [6]:

$$\begin{aligned}
 F_{hitung} &= \frac{r^2(n-k-1)}{k(1-r^2)} \\
 &= \frac{0,998(10-4-1)}{4(1-0,998)} \\
 &= \frac{0,998(5)}{4(0,00325)} \\
 &= \frac{4,9837}{0,013} \\
 &= 383,2912
 \end{aligned}$$

Berdasarkan jumlah variabel bebas (df1) sebagai derajat kebebasan pembilang dan (df2) : $(n - k - 1)$ sebagai derajat penyebut diperoleh $F_{tabel} = 5,1922$ yang dapat dilihat pada tabel 2 berikut :

Tabel 4. F-Tabel

Nilai alpha		0,05								
		dk1=dk								
dk2=dk		pembilang								
Penyebut	1	2	3	4	5	6	7	8	9	10
1	161,45	199,5	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88
2	18,51	19	19,16	19,25	19,29	19,33	19,35	19,37	19,38	19,39
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,84	8,81	8,78
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	5,99	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,87	4,82	4,77	4,73
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,09	4,06
7	5,59	4,74	4,35	4,12	3,97	3,86	3,79	3,72	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,32	3,22	3,13	3,07	3,02	2,98

Dimana untuk derajat kebebasan $df1: k$ (jumlah variabel bebas) = 4 dan $df2: n - k - 1 = 10 - 4 - 1 = 5$.

Diperoleh nilai F hitung adalah 383,291 dengan taraf sig. = 0.000 < 0.05 dimana df *regression* sebesar 4 dan df *residual* sebesar 5. Dari tabel F didapatkan nilai F table sebesar 5.1921. Hal ini menunjukkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan signifikan secara simultan atau bersama-sama antara variabel bebas terhadap variabel terikat dikarenakan nilai $F_{hitung} > F_{tabel}$ yaitu $383.291 > 5.1921$.

3.7 Uji Parsial (uji t)

Adapun tujuan yang ingin dicapai dari uji yang dilakukan terhadap uji t adalah untuk melihat seberapa besar pengaruh yang diberikan variable bebas secara parsial atau masing-masing terhadap variabel terikat yang ada di dalam model regresi. Kriteria keputusan untuk uji t adalah dengan taraf signifikan $\alpha = 0,05$. Apabila $t_{hit} > t_{tab}$ atau $-(t_{hit}) < -(t_{tab})$ maka H_0 ditolak dan H_1 diterima, artinya terdapat hubungan yang signifikan secara parsial antara variabel bebas terhadap variabel terikat. Begitu jugasebaliknya, apabila $t_{hit} < t_{tab}$ atau $-(t_{hit}) > -(t_{tab})$ maka H_0 diterima dan H_1 ditolak, artinya tidak ada hubungan yang signifikan secara parsial antara variabel bebas terhadap variabel terikat. Perhitungan uji t dapat dilakukan sebagai berikut [6]:

- a. Membentuk matriks A dan matriks H pada persamaan (6), (7), (8), (9) dan (10)

$$A = \begin{bmatrix} 10 & 541 & 60 & 3330000 & 39 \\ 541 & 29877 & 3251 & 179930000 & 2080 \\ 60 & 3251 & 368 & 20250000 & 236 \\ 3330000 & 179930000 & 20250000 & 1119000000000 & 13085000 \\ 39 & 2080 & 236 & 13085000 & 163 \end{bmatrix}$$

$$H = \begin{bmatrix} 796 \\ 42538 \\ 4847 \\ 267900000 \\ 3115 \end{bmatrix}$$

- b. Mencari nilai invers dari matriks A

Adapun nilai invers dari matriks A adalah sebagai berikut:

$$A^{-1} = \begin{bmatrix} 45,36994 & -0,23815 & 6,72515 & -0,00021475 & -0,314297 \\ -0,23815 & 0,0024207 & -0,029766 & 0,00000080081 & 0,00490201 \\ 6,72515 & -0,029766 & 1,69750923 & -0,000046207 & 0,022318 \\ -0,00021475 & 0,00000080081 & -0,000046207 & 0,0000000013662 & -0,000001608 \\ -0,314297 & 0,00490201 & 0,022318 & -0,000001608 & 0,115552062 \end{bmatrix}$$

- c. Mencari nilai matriks varian kovarian

Nilai ini didapatkan dengan mengalikan nilai *Mean Square Error* (MSE) dengan invers matriks A. Adapun perhitungan MSE dapat dihitung dengan menggunakan tabel berikut:

Tabel 5. Tabel bantu MSE

No	Y	$\hat{Y}$	$Y - \hat{Y}$	$(Y - \hat{Y})^2$	$\frac{Y_t - F_t}{Y_t}$
1	70	69,765	0,235	0,055019608	0,003350894
2	56	56,984	-0,984	0,968332268	-0,01757212
3	74	74,153	-0,153	0,023299344	-0,00206272
4	81	80,490	0,509	0,259864772	0,006293449
5	86	86,523	-0,523	0,273694773	-0,00608234
6	77	75,692	1,308	1,710483583	0,016985124
7	73	72,827	0,173	0,029977768	0,002371793
8	92	91,450	0,5496	0,30203986	0,005973712
9	90	90,327	-0,327	0,106969925	-0,00363403
10	97	97,788	-0,788	0,620956466	-0,00812379
Σ	796	796	2,13163E-13	4,350638368	-0,00250093

$$MSE = \frac{\Sigma(Y - \hat{Y})^2}{n - k - 1} = \frac{4,350638368}{10 - 4 - 1} = 0,870127673$$

$$A^{-1} \times MSE =$$

$$\begin{bmatrix} 45,36994 & -0,23815 & 6,72515 & -0,00021475 & -0,314297 \\ -0,23815 & 0,0024207 & -0,029766 & 0,00000080081 & 0,00490201 \\ 6,72515 & -0,029766 & 1,69750923 & -0,000046207 & 0,022318 \\ -0,00021475 & 0,00000080081 & -0,000046207 & 0,0000000013662 & -0,000001608 \\ -0,314297 & 0,00490201 & 0,022318 & -0,000001608 & 0,115552062 \end{bmatrix} \times MSE$$

$$= \begin{bmatrix} 45,36994 & -0,23815 & 6,72515 & -0,00021475 & -0,314297 \\ -0,23815 & 0,0024207 & -0,029766 & 0,00000080081 & 0,00490201 \\ 6,72515 & -0,029766 & 1,69750923 & -0,000046207 & 0,022318 \\ -0,00021475 & 0,00000080081 & -0,000046207 & 0,000000013662 & -0,000001608 \\ -0,314297 & 0,00490201 & 0,022318 & -0,000001608 & 0,115552062 \end{bmatrix} \times$$

$$0,870127673$$

$$= \begin{bmatrix} 39,47764 & -0,20722 & 5,85174 & -0,00018686 & -0,273479 \\ -0,20722 & 0,0021063 & -0,025900 & -0,025900 & 0,00426537 \\ 5,85174 & -0,025900 & 1,47704976 & -0,000040206 & 0,019419 \\ -0,00018686 & 0,00000069681 & -0,000040206 & 0,0000000011888 & -0,000001399 \\ -0,273479 & 0,00426537 & 0,019419 & -0,000001399 & 0,100545047 \end{bmatrix}$$

- d. Mencari nilai standar error (SE) koefisien regresi

Nilai Standar Error (SE) didapatkan dari akar matriks kovarian dari diagonal matriks varian kovarian yaitu sebagai berikut:

$$SE_a = \sqrt{39,47764} = 6,283123139$$

$$SE_{b_1} = \sqrt{0,00210634} = 0,045894879$$

$$SE_{b_2} = \sqrt{1,47704976} = 1,21533936$$

$$SE_{b_3} = \sqrt{0,000000001188875} = 0,000034478$$

$$SE_{b_4} = \sqrt{0,100545047} = 0,317088389$$

- e. Mencari nilai t hitung

Nilai t hitung didapatkan dengan membagi nilai koefisien regresi dengan Standar Error (SE) koefisien regresi yaitu dengan menggunakan persamaan (2.26) sebagai berikut:

$$t_a = \frac{a}{SE_a} = \frac{70,762788}{6,283123139} = 11,26236$$

$$t_1 = \frac{b_1}{SE_{b_1}} = \frac{-1,065849}{0,045894879} = -23,22371$$

$$t_2 = \frac{b_2}{SE_{b_2}} = \frac{5,546495}{1,21533936} = 4,56374$$

$$t_3 = \frac{b_3}{SE_{b_3}} = \frac{0,000150367}{0,000034478} = 4,36121$$

$$t_4 = \frac{b_4}{SE_{b_4}} = \frac{-4,320903899}{0,317088389} = -13,62681$$

Berdasarkan derajat kebebasan ($n - k - 1$) diperoleh $t_{tabel} = 2,571$ yang dapat dilihat pada tabel

4. Berikut:

Tabel 6. t-tabel

Alpha untuk uji dua pihak						
	0,5	0,2	0,1	0,05	0,02	0,01
dk	Alpha untuk uji satu pihak					
	0,25	0,1	0,05	0,025	0,01	0,005
1	1	3,077684	6,313752	12,7062	31,82052	63,65674
2	0,816497	1,885618	2,919986	4,302653	6,964557	9,924843
3	0,764892	1,637744	2,353363	3,182446	4,540703	5,840909
4	0,740697	1,533206	2,131847	2,776445	3,746947	4,604095
5	0,726687	1,475884	2,015048	2,570582	3,3693	4,032143
6	0,717558	1,439756	1,94318	2,446912	3,142668	3,707428
7	0,711142	1,414924	1,894579	2,364624	2,997952	3,499483
8	0,706387	1,396815	1,859548	2,306004	2,896459	3,355387
9	0,702722	1,383029	1,833113	2,262157	2,821438	3,249836
10	0,699812	1,372184	1,812461	2,228139	2,763769	3,169273

Penentuan untuk mengetahui nilai t tabel adalah $t_{tabel} = t\left(\frac{\alpha}{2}; n - k - 1\right) = \left(\frac{0,05}{2}; 10 - 4 - 1\right) = (0,025; 5)$. Berdasarkan hasil di atas, uji t dapat disimpulkan sebagai berikut:

1. Nilai t_{hitung} pada variabel X_1 adalah -23,224 dengan taraf $sig. = 0,000 < 0,005$ dimana derajat kebebasan sebesar 6. Dari tabel t didapatkan nilai t_{tabel} sebesar 2,571. Dikarenakan t_{hitung} merupakan negatif, maka t_{tabel} yang digunakan juga negatif yaitu sebesar -2,571. Dapat

disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan secara parsial antara variabel X_1 terhadap variabel terikat Y dikarenakan nilai $-t_{hitung} < -t_{tabel}$ yaitu $-23,224 < -2,571$ – $23,224 < -2,571$.

2. Nilai t_{hitung} pada variabel X_2 adalah 4,564 dengan taraf $sig. = 0,006 > 0,005$ dimana derajat kebebasan sebesar 6. Dari tabel t didapatkan nilai t_{tabel} sebesar 2,571. Dapat disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan secara parsial antara variabel X_2 terhadap variabel terikat Y dikarenakan nilai $t_{hitung} > t_{tabel}$ yaitu $4,564 > 2,571$.
3. Nilai t_{hitung} pada X_3 adalah 4.361 dengan taraf $sig. = 0,007 > 0,005$ dimana derajat kebebasan sebesar 6. Dari tabel t didapatkan nilai t_{tabel} sebesar 2,571. Dapat disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan secara parsial antara variabel X_3 terhadap variabel terikat Y dikarenakan nilai $t_{hitung} > t_{tabel}$ yaitu $4,361 > 2,571$.
4. Nilai t_{hitung} pada variabel X_4 adalah -13.627 dengan taraf $sig. = 0,000 < 0,005$ dimana derajat kebebasan sebesar 6. Dari tabel t didapatkan nilai t_{tabel} sebesar 2,571. Dikarenakan t_{hitung} merupakan negatif, maka t_{tabel} yang digunakan juga negatif yaitu sebesar -2,571. Dapat disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan secara parsial antara variabel X_4 terhadap variabel terikat Y dikarenakan nilai $t_{hitung} < t_{tabel}$ yaitu $-13,627 < -2,571$.

Berdasarkan t hitung dari masing-masing variabel bebas di atas, dapat disimpulkan bahwa variabel bebas yang sangat berpengaruh terhadap variabel santri baru (Y) adalah fasilitas pondok (X_2) dan biaya pertahun (X_3).

3.8 Perhitungan Nilai Akurasi Peramalan

3.8.1 MAPE (Mean Absolute Percentage Error)

MAPE digunakan untuk mengetahui persentase kesalahan dari hasil prediksi yang dilakukan [8]. Perhitungannya adalah sebagai berikut:

$$\begin{aligned}
 MAPE &= \frac{\sum \left| \frac{Y_t - F_t}{Y_t} \right|}{n} \times 100\% \\
 &= \frac{|-0,00250093|}{10} \times 100\% \\
 &= \frac{0,00250093}{10} \times 100\% \\
 &= 0,000250093 \times 100\% \\
 &= 0,0250093\% \\
 &= 0,02\%
 \end{aligned}$$

Berdasarkan hasil perhitungan *mean absolut percentage error*, tingkat *error* yang dihasilkan metode Regresi Linear Berganda adalah sebesar 0,02%. Jadi, tingkat akurasi MAPE yaitu $0,02\% \leq 10\%$. Jadi dapat disimpulkan bahwa kemampuan dari model peramalan yang digunakan dapat dikatakan sangat baik atau sangat akurat.

3.8.2 RMSE (Root Mean Squared Error)

Saat membandingkan nilai *forecasting* dan nilai aktual yang diperoleh dari model peramalan, RMSE digunakan sebagai pembeda yang dihitung sebagai berikut [9]:

$$\begin{aligned}
 RMSE &= \sqrt{\frac{\sum (Y_t - F_t)^2}{n}} \\
 &= \sqrt{\frac{4,350638368}{10}} \\
 &= \sqrt{0,4350638368} \\
 &= 0,65959369 \\
 &= 0,66
 \end{aligned}$$

Berdasarkan hasil perhitungan *root mean Squared Error*, tingkat *error* yang dihasilkan metode Regresi Linear Berganda adalah sebesar 0.66. Jadi, dapat disimpulkan bahwa kemampuan dari model peramalan yang digunakan dapat dikatakan baik atau akurat.

4. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan sebelumnya, maka pada penelitian ini dapat diambil kesimpulan yaitu:

1. Diketahui bahwa metode regresi linear berganda dapat digunakan untuk memprediksi jumlah santri baru di Pondok Pesantren Al-Falah Sumber Gayam Kec. Kadur, Kab. Pamekasan pada tahun yang akan datang dengan tingkat *error* yang sangat kecil, yaitu pada tahun 2023-2024 dengan hasil prediksi sebanyak 98 santri baru.
2. Model dari metode regresi linear berganda yang sudah dibentuk adalah $\hat{Y} = 70,7627885505303 + (-1,06584933412433)X_1 + 5,546495227196812X_2 + 0,0001500367009824364X_3 + (-4,32090389977076)X_4$
3. Ukuran kesalahan peramalan digunakan pada penelitian ini adalah MAPE dan RMSE. Tingkat akurasi yang diperoleh MAPE sebesar 0.02%, dan tingkat akurasi yang diperoleh RMSE sebesar 0.66%. Berdasarkan tingkat keakuratan yang diperoleh, hasil menunjukkan bahwa kemampuan dari model peramalan yang digunakan dapat dikatakan baik atau akurat.
4. Berdasarkan hasil dari koefisien determinasi (r^2) diperoleh hasil sebesar 99.7%, hal ini dapat dikatakan bahwa variabel santri baru dipengaruhi oleh variabel santri alumni pertahun, fasilitas pondok, biaya pertahun dan santri berprestasi dan untuk 0.3% dipengaruhi oleh variabel yang lain, hal ini dapat disimpulkan bahwa variabel-variabel *independent* memiliki kemampuan untuk memberikan semua informasi yang dibutuhkan untuk memprediksi variabel *dependent*. Hasil yang diperoleh koefisien korelasi ganda (r) sebesar 0.998 yang berarti 99.8% variabel santri baru dipengaruhi oleh variabel santri alumni pertahun, fasilitas pondok, biaya pertahun dan santri berprestasi, sedangkan untuk 0.2% dipengaruhi oleh variabel yang lain, hal ini dapat disimpulkan bahwa terdapat pengaruh positif dan korelasi variabel-variabel yang diuji sangat kuat. Berdasarkan hasil uji simultan (uji F) diperoleh nilai $F_{hitung} > F_{tabel}$ yaitu $383.291 > 5.1921$ dengan taraf $sig. = 0.000 < 0.05$. Dapat disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan secara simultan antara variabel bebas terhadap variabel terikat. Sedangkan hasil uji parsial (uji-t) untuk t_1 diperoleh nilai $-t_{hitung} < -t_{tabel}$ yaitu $-23.224 < -2.571$ dengan taraf $sig. = 0.000 < 0.05$, t_2 diperoleh nilai $t_{hitung} > t_{tabel}$ yaitu $4.564 > 2.571$ dengan taraf $sig. = 0.006 < 0.05$, t_3 diperoleh nilai $t_{hitung} > t_{tabel}$ yaitu $4.361 > 2.571$ dengan taraf $sig. = 0.007 < 0.05$, dan t_4 diperoleh nilai $-t_{hitung} < -t_{tabel}$ yaitu $-13.627 < -2.571$ dengan taraf $sig. = 0.000 < 0.05$. Jadi, dapat disimpulkan bahwa H_0 ditolak dan H_1 diterima yang berarti terdapat hubungan yang signifikan antara variabel X_1, X_2, X_3 dan X_4 secara parsial terhadap variabel terikat Y . Berdasarkan hasil uji t tersebut, dapat disimpulkan bahwa variabel bebas yang sangat berpengaruh terhadap variabel Santri Baru (Y) adalah fasilitas pondok (X_2) dan biaya pertahun (X_3).

DAFTAR PUSTAKA

- [1] R. Zana Nainggolan, K. Ibnutama, and M. Gilang Suryanata, "Implementasi Data Mining Dengan Metode Regresi Linier Berganda Dalam Estimasi Mahasiswa Baru Pada Sekolah Tinggi Agama Islam Raudhatul Akmal Batang Kuis," *J. CyberTech*, vol. 1, no. 1, pp. 13–20, 2021, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct>
- [2] V. Z. Ningrum and T. Rochana, "Perilaku Sosial Santri di Pondok Pesantren Tarbiyatul Muballighin Desa Reksosari," *Solidarity*, vol. 8, no. 2, pp. 749–761, 2019.
- [3] R. Puspitasari, S. Effendi, H. Kurniawan, and M. A. E. Nasution, "Penentuan Prediksi Hasil Panen Kelapa Sawit Menggunakan Metode Regresi Linier," in *Seminar Nasional Riset Information Science*, 2022.
- [4] W. Suci, "Pengaruh Media Pembelajaran Terhadap Hasil Belajar Al-Islam Di SMA Muhammadiyah 1 Gisting Kabupaten Tanggamus Tahun Pelajaran 2019/2020," Institut Agama Islam Negeri (Iain) Metro. Metro, 2020.
- [5] A. F. Boy, "Implementasi data mining dalam memprediksi harga Crude Palm Oil (CPO) pasar domestik menggunakan algoritma regresi linier berganda (Studi kasus dinas perkebunan provinsi Sumatera Utara)," *J. Sci. Soc. Res.*, vol. 4307, no. 2, pp. 78–85, 2020.
- [6] D. L. Trenggonowati, K. Kulsum, and F. A. Salma, "Model Persamaan Regresi Pengoptimalan Waktu Pelayanan Bongkar Kargo Menggunakan Multiple Regression Analysis di PT. XYZ," *J. Ind. Serv.*, vol. 6, no. 1, p. 13, Nov. 2020, doi: 10.36055/jiss.v6i1.9484.
- [7] D. Erri, A. P. Lestari, and H. H. Asymar, "Pengaruh Gaya Kepemimpinan Terhadap Kinerja Karyawan Pada PT Melzer

- Global Sejahtera Jakarta, Jurnal Inovasi Penelitian,” *J. Inov. Penelit.*, vol. 1, no. 9, pp. 1897–1907, 2021.
- [8] M. M. Azman, “Analisa perbandingan nilai akurasi moving average dan exponential smoothing untuk sistem peramalan pendapatan pada perusahaan XYZ,” *J. Sist. dan Inform.*, vol. 13, no. 2, pp. 36–45, 2019.
- [9] M. G. Pramayudha, “Prediksi Hasil Panen Tanaman Pangan Dengan Metode Single Moving Average Dan Single Exponential Smoothing,” Universitas Negeri Maulana Malik Ibrahim, 2018.