Implementasi Regresi Linear Berganda Untuk Prediksi Pembayaran Premi Asuransi Kesehatan

Noviana Riza[1)*], Sri Cahyani[2)], Siti Khairiyah[3)]
[1,2,3)] Universitas Logistik dan Bisnis Internasional, Fakultas Logistik, Teknologi dan Bisnis, Program Studi Sains Data

Received: 18 December 2025 Accepted: 30 December 2025 Published: 31 December 2025

*sricahyani.15sd@gmail.com

Kata Kunci: Asuransi Kesehatan, Prediksi Premi, Regresi Linear Berganda, Streamlit, Machine Learning

DSI: Jurnal Data Science Indonesia is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Meskipun menunjukkan tren pertumbuhan yang positif, penetrasi asuransi di Indonesia masih relatif rendah dibandingkan dengan negara lain di Asia Tenggara. Oleh karena itu, penting untuk memahami variabel yang memengaruhi besaran premi asuransi kesehatan. Tujuan dari penelitian ini adalah untuk membuat model prediksi pembayaran premi asuransi kesehatan swasta dengan menggunakan teknik regresi linear berganda. Data penelitian ini diperoleh dari Kaggle, dan 1.338 observasi termasuk variabel usia, jenis kelamin, indeks massa tubuh (BMI), jumlah anak, dan status merokok. Hasil analisis menunjukkan bahwa usia, BMI, jumlah anak, dan status merokok memiliki pengaruh signifikan terhadap premi asuransi kesehatan. Sebaliknya, variabel jenis kelamin tidak memiliki pengaruh yang signifikan. Semua variabel independen memiliki dampak yang signifikan terhadap variabel premi asuransi kesehatan, menurut hasil uji simultan (uji F). Koefisien determinasi (R2) model regresi adalah 74,9%, menunjukkan kemampuan model untuk menjelaskan variasi data premi. Meskipun multikolinearitas tidak ditemukan oleh pengujian asumsi klasik, asumsi normalitas residual dan heteroskedastisitas dibatalkan. Hasil perhitungan interval prediksi menunjukkan nilai prediksi sebesar 3.652,90 dengan interval kepercayaan 95% yang sangat lebar, yaitu antara −96.877,09 dan 104.182,90. Hasil ini menunjukkan bahwa ada tingkat ketidakpastian yang tinggi dalam prediksi.

Pendahuluan

Kesehatan adalah kebutuhan dasar setiap orang. Pemerintah sebagai pemerintah negara bertanggung jawab untuk memastikan bahwa semua orang, termasuk mereka yang miskin dan rentan, memiliki akses ke layanan kesehatan[1]. Semakin banyak kemajuan yang dilakukan di bidang asuransi jiwa akhir-akhir ini telah meningkatkan kesadaran masyarakat akan pentingnya memiliki asuransi. Pembiayaan kesehatan adalah masalah utama dalam pelayanan kesehatan, jadi diperlukan asuransi yang dapat menjamin kesehatan individu tersebut[2]. Selain itu, hasil penelitian menunjukkan bahwa guncangan kesehatan secara signifikan mengurangi pendapatan individu dan rumah tangga. Ini menunjukkan bahwa asuransi kesehatan dan layanan keuangan dapat memainkan peran penting dalam mengurangi dampak negatif guncangan kesehatan pada rumah tangga, karena dampaknya akan lebih parah pada rumah tangga miskin atau berpenghasilan rendah[3].

Menurut Financial Fitness Index(2021), tingkat kesehatan keuangan masyarakat Indonesia masih rendah, dengan skor 37,72 dari 100, menunjukkan bahwa orang-orang di negara ini tidak dapat menabung secara teratur, menyimpan dana untuk situasi darurat, dan memenuhi kebutuhan mereka. Meskipun akses ke layanan keuangan semakin luas, kondisi tersebut menunjukkan betapa kurangnya pengetahuan masyarakat tentang keuangan. Dengan meningkatnya kesadaran finansial, orang mulai menyadari betapa pentingnya asuransi kesehatan untuk melindungi mereka dari biaya kesehatan yang tidak terduga. Menurut survei Dana Reksa, kepemilikan asuransi kesehatan swasta dan pemerintah terus meningkat. Karena itu,

perusahaan asuransi kesehatan memiliki peran penting dalam memengaruhi keputusan pembelian masyarakat. Oleh karena itu, penelitian ini bertujuan untuk menganalisis beberapa faktor yang mempengaruhi keputusan pembelian asuransi kesehatan[4].

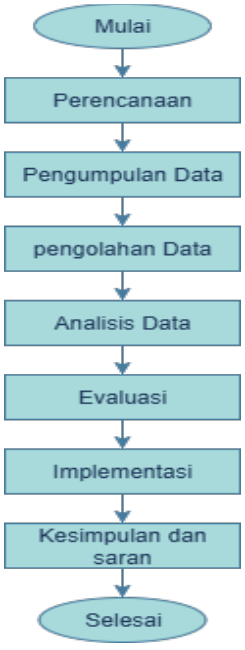
Tinjauan Literatur

Berbagai penelitian telah dilakukan untuk mengkaji prediksi penggunaan asuransi dan pemanfaatan regresi linear berganda dalam berbagai konteks. Menyoroti pentingnya pengelolaan risiko dalam sistem asuransi kesehatan di Indonesia, termasuk peran managed care dan teknologi digital [5]. Selanjutnya, menggunakan algoritma Random Forest untuk memprediksi premi asuransi, dan menemukan bahwa status merokok menjadi salah satu penentu utama kenaikan premi[6].

Sejalan dengan itu, peneliti lain membuktikan bahwa machine learning efektif dalam memprediksi perpanjangan polis asuransi, dengan algoritma Random Forest yang menghasilkan akurasi terbaik[7]. Dalam sektor lain, menerapkan regresi linear berganda untuk memprediksi tonase kelapa sawit dan memperoleh akurasi hingga 99,99%, menunjukkan kekuatan metode ini dalam prediksi data multivariat [8]. Pada kasus yang lain, mengaplikasikan regresi linear berganda untuk menganalisis pengaruh pertumbuhan ekonomi dan IPM terhadap kemiskinan di Indonesia, dan menemukan kedua variabel tersebut berpengaruh signifikan [9]. Terakhir, ada penelitian menunjukkan bahwa regresi linear sederhana dapat digunakan untuk memprediksi harga mobil bekas dengan akurasi tinggi melalui aplikasi Streamlit [10]. Tinjauan ini menunjukkan bahwa regresi linear berganda merupakan metode yang kuat dan fleksibel, baik dalam prediksi premi asuransi maupun dalam konteks sosial ekonomi lainnya, sehingga relevan untuk digunakan dalam penelitian prediksi penggunaan asuransi berdasarkan karakteristik pelanggan.

METODE PENELITIAN

Analisis penelitian menggunakan bahasa pemrograman Python dalam metode machine learning regresi linear berganda. Adapun alur penelitiannya adalah sebagai berikut:



Gambar 1. Diagram alur penelitian

Berikut adalah alur yang dilakukan dalam penelitian

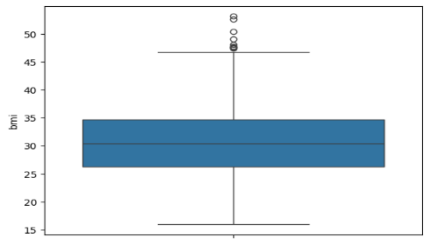
1. Perencanaan: Menentukan tujuan penelitian, memilih metode regresi linear berganda, serta menentukan sumber data.
2. Pengumpulan Data: Dataset terdiri dari 1338 baris data dengan variabel age, sex, bmi, children, smoker, charges.
3. Pengolahan data: melakukan pembersihan variabel yang tidak diperlukan dan melakukan encoding terhadap variabel dengan tipe data string.

4. Analisis data: Melakukan uji asumsi klasik (normalitas, multikolinearitas, heteroskedastisitas) serta uji F dan uji T.
5. Evaluasi Model: Menghitung R^2, MAE, RMSE untuk menilai performa model.
6. Implementasi: penerapan python untuk memprediksi premi asuransi kesehatan.

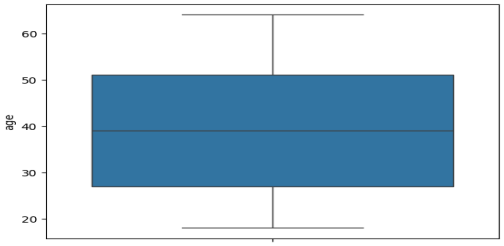
HASIL PENELITIAN

Berdasarkan analisis yang telah dilakukan kita mendapatkan hasil dari setiap pengujian analisis dan hasil prediksi untuk premi asuransi.

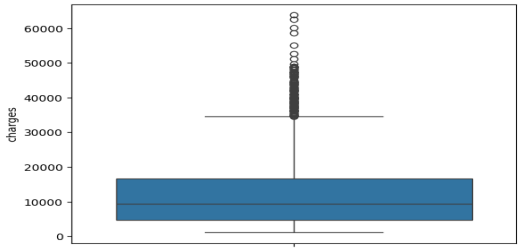
1. Uji outlier



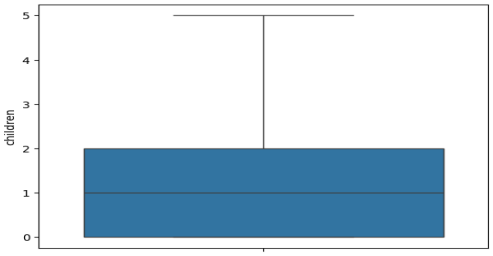
Gambar 2(a). Uji Outlier bmi



Gambar 2(b). Uji Outlier age



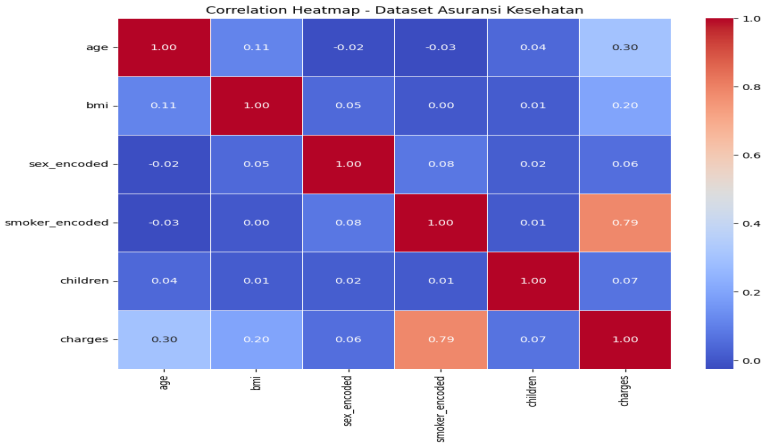
Gambar 3(a). Uji Outlier charges



Gambar 3(b). Uji Outlier children

Berdasarkan visualisasi boxplot diatas dapat kita lihat bahwa terdapat dua variabel x yang memiliki outlier yaitu bmi dan smoker. Namun, disini outlier tidak perlu dihapus karena nilai tersebut menjadi salah satu acuan untuk melakukan prediksi terhadap pembayaran premi dan tidak dianggap sebagai anomali (salah input).

2. korelasi setiap variabel



Gambar 4. Korelasi Variabel X dan Y

Heatmap diatas menunjukan hubungan dari setiap variabel, dengan nilai korelasi dari -1 sampai 1. semakin nilai korelasi mendekati 1, semakin kuat hubungan relasi antara dua variabel nya. Dari korelasi heatmap diatas menunjukan bahwa smoker_encode memiliki hubungan korelasi yang kuat terhadap variabel Y dengan korelasi sebesar 0.79, yang berpengaruh terhadap perhitungan premi asuransi.

3. implementasi metode regresi linear berganda

   Hasil persamaan regresi linear berganda yang terbentuk yaitu:

$$Y = -12.052 + 257X_1 - 128X_2 + 322X_3 + 474X_4 + 23.823X_5$$

   a. Nilai konstanta sebesar -12.052 yang artinya nilai konstanta ini merupakan nilai premi ketika semua variabel independen bernilai 0.
   b. Nilai koefisien variabel X1 (age) sebesar 257, yang berarti akan meningkatkan premi sebesar 257 setiap penambahan usia.
   b. Nilai koefisien variabel X2 (sex) sebesar -128 , Jenis kelamin menurunkan premi sebesar 128
   c. Nilai koefisien variabel X3 (bmi) sebesar 322, yang artinya setiap kenaikan berat badan 1 kg akan meningkatkan premi sebesar 322
   d. Nilai koefisien variabel X4 (children) sebesar 474, yang artinya setiap penambahan 1 anak akan meningkatkan premi sebesar 474.
   e. Nilai koefisien variabel X5 (smoker) sebesar 23.823, yang artinya setiap setiap perokok akan menaikan premi sebesar 23.823. pengaruh rokok sangat besar terhadap pembayaran premi.

   Seseorang yang merokok memiliki pengaruh yang paling besar terhadap premi, diikuti oleh children, bmi, usia dan sex. Menurut model, usia, status merokok, jumlah anak, dan indeks massa tubuh (BMI) mempengaruhi besaran premi asuransi kesehatan.

PEMBAHASAN

Tabel 1. Data Asli

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520
...	...	...	...	...	...	...	...
1337	61	female	29.070	0	yes	northwest	29141.36030

Tabel 2. Data yang sudah di Cleaning

	age	sex	bmi	children	smoker	charges
0	19	0	27.900	0	1	16884.92400
1	18	1	33.770	1	0	1725.55230
2	28	1	33.000	3	0	4449.46200
3	33	1	22.705	0	0	21984.47061
4	32	1	28.880	0	0	3866.85520
...	...	...	...	...	...	...
1337	61	0	29.070	0	1	29141.36030

Berikut adalah tahapan yang dilakukan

1. Perencanaan

   Tentukan tujuan penelitian, pilih dataset personal biaya medis dari Kaggle, dan mulai menggunakan metode analisis regresi linear berganda. Model persamaan regresi linear berganda adalah bagian dari analisis regresi linear yang melihat bagaimana variabel terikat (dependen) y dan beberapa variabel bebas (independen) x berinteraksi satu sama lain [11]. Persamaan regresi linear berganda yang umum adalah:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots\ldots + b_nX_n$$

2. Pengumpulan data

   Dataset terdiri dari 1338 baris data dengan variabel: usia (age), jenis kelamin (sex), indeks massa tubuh (BMI), jumlah anak (children), status merokok (smoker), dan biaya asuransi (charges).

3. Pengolahan data

   Pembersihan data dengan menghapus kolom data yang tidak diperlukan dan melakukan encoding pada kolom sex dan smoker. Pada encoding menggunakan bilang biner. Kolom sex di encoding (female : 0, male:1) dan encoding smoker (yes: 1, no: 0).

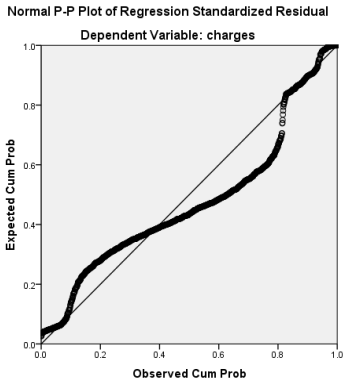
4. Analisis Data

   Age, sex, body massa index, children, and smoker adalah variabel independen yang digunakan untuk membangun model regresi linear berganda. Persamaan model penelitian adalah sebagai berikut:

$$Y = -12.052 + 257X_1 - 128X_2 + 322X_3 + 474X_4 + 23.823X_5$$

   Uji asumsi klasik dilakukan meliputi:

   a. Uji Normalitas Residual (Kolmogorov-Smirnov)



Gambar 5. Uji Normalitas Residual

Tabel 3. Uji normalitas residual

		Unstandardized Residual
N		1338
Normal Parameters[a,b]	Mean	0000000
	Std. Deviation	6058.365095
Most Extreme Differences	Absolute	.163
	Positive	.163
	Negative	-.092
Test Statistic		.163
Asymp. Sig. (2-talled)		.000[c]

Pengujian dilakukan untuk menentukan apakah nilai residual terdistribusi secara normal. Nilai residual harus terdistribusi secara normal dalam model regresi yang baik. Cara untuk mengetahuinya adalah dengan melihat penyebaran data pada sumber diagonal pada plot regression standardized normal P-P. Ini akan berfungsi sebagai dasar untuk membuat keputusan. Model regresi dapat digunakan untuk memprediksi variabel bebas jika tersebar di sekitar garis dan mengikuti garis diagonal, tetapi sebaliknya[12].

$$D = max|F0(x_i) - Sn(x_i)|$$

Keterangan:

D : Nilai Statistika Kolmogorov-Smirnov

$F_0(X_i)$: Probabilitas kumulatif normal (FT)

$S_n(X_i)$: Probabilitas kumulatif empiris (FS)

X_i : Angka pada data

kriteria pengujiannya adalah:

1. Jika nilai signifikansi > 0,05, maka data berdistribusi normal
2. Jika nilai signifikansi < 0,05, maka data tidak berdistribusi normal

b. Uji Multikolinearitas (VIF)

Uji multikolinearitas digunakan untuk menentukan apakah ada hubungan linear yang sempurna atau mendekati sempurna di antara beberapa atau semua variabel bebas dalam model regresi [13].

Rumus umum nya yaitu:

$$VIF = \frac{1}{(1 - R^2)}$$

Tabel 4. Uji Multikolinearitas

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std.Error	Beta			Tolerance	VIF
1	(Constant)	-12052.462	951.260		-12.670	.000		
	age	257.735	11.904	.229	21.651	.000	.985	1.015
	sex	-128.640	333.361	-.005	-.386	.700	.991	1.009
	bmi	322.364	27.419	.162	11.757	.000	.986	1.015
	children	474.411	137.856	.047	3.441	.001	.998	1.002
	smoker	23823.393	412.523	.794	57.750	.000	.994	1.006

Dependent Variable: charges

1. Jika nilai VIF < 10 maka terbebas dari multikolinearitas
2. Jika nilai VIF >10 maka akan terjadi multikolinearita

c. Uji Heteroskedastisitas

Tabel 5. Uji Heteroskedasitas

		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std.Error	Beta	t	Sig.
1	(Constant)	-85.876	583.118		-.147	.883
	age	-4.716	7.297	-.015	-.646	.518
	sex_encoded	-48.998	204.348	-.006	-.240	.811
	bmi	105.725	16.807	.147	6.290	.000
	children	89.361	84.505	.025	1.057	.290
	smoker_encoded	5539.889	252.875	.510	21.908	.000

Dependent Variable: ABS_RES. Uji heteroskedasitas atau uji identik dimaksudkan untuk melihat apakah residual memiliki sifat identik (homoskedastisitas) atau tidak identik (heteroskedasitas)[14]. residual dikatakan homoskedastisitas jika signifikansi > 0,05.

5. Evaluasi Model

```
X_train shape: (936, 5)
X_test shape: (402, 5)
y_train shape: (936,)
y_test shape: (402,)
```

Gambar 6. Train Test Split

Evaluasi dilakukan dengan menghitung:

a. Koefisien Determinasi (R^2)

Tabel 6. Uji Determinasi

Model	R	R Square	Adjusted R Square	Std.Error of the Estimate
1	.866[a]	.750	.749	6069.725250

Predictors: (Constant), smoker_encoded, bmi, children, sex_encoded, age. Dependent Variable: charges

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(yi - \hat{y})^2}{\sum_{i=1}^{n}(yi - \bar{y})^2}$$

Uji Koefisien Determinasi merupakan bagian dari analisis regresi yang bertujuan untuk mengetahui seberapa besar kemampuan model regresi dalam menjelaskan variabilitas variabel dependen (Y) yang disebabkan oleh variabel independen (X). R-squared berkisar antara 0 sampai 1[15].

b. Uji F

Tabel 7. Uji F

Model		Sum of Squares	df	Mean Square	F	Sig
1.	Regression	1.470E+11	5	2.940E+10	798.019	.000[b]
	Residual	4.907E+10	1332	36841564.60		
	Total	1.961E+11	1337			

Dependent Variabel: charges. Predictors: (Constant), smoker_encoded, bmi, children, sex_encoded, age

$$F = \frac{R2^2/K}{(1 - R^2)/(n - k - 1)}$$

Proses pemeriksaan menentukan apakah variabel independen dan variabel dependen memiliki pengaruh yang signifikan satu sama lain. Uji F akan menguji model secara keseluruhan[16]. Dasar pengambilan uji f dapat dilihat melalui 2 acuan, pertama dengan melihat nilai signifikan yaitu :

1. Jika nilai Sig < 0,05, maka hipotesis diterima, artinya motivasi (x1) dan minat (x2) secara simultan berpengaruh terhadap prestasi (Y).
2. Jika nilai Sig > 0,05, maka hipotesis ditolak, artinya motivasi (x1) dan minat (x2) secara simultan tidak berpengaruh terhadap prestasi (Y).

Kedua, dengan melihat perbandingan nilai t hitung dan t tabel yaitu :

1. H0 diterima apabila Fhitung > Ftabel untuk tingkat signifikan α = 5%
2. Ha diterima apabila Fhitung > Ftabel untuk tingkat signifikan α = 5%

c. Uji T

Tabel 8. Uji T

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std.Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-12052.462	951.260		-12.670	.000		
	age	257.735	11.904	.229	21.651	.000	.985	1.015
	sex	-128.640	333.361	-.005	-.386	.700	.991	1.009
	bmi	322.364	27.419	.162	11.757	.000	.986	1.015
	children	474.411	137.856	.047	3.441	.001	.998	1.002
	smoker	23823.393	412.523	.794	57.750	.000	.994	1.006

Dependent Variable: charges

$$t = \frac{bi}{SE(bi)}$$

Uji t adalah salah satu uji statistik yang digunakan untuk mengetahui apakah masing-masing variabel independen (X) memiliki pengaruh yang signifikan secara individual terhadap variabel

dependen (Y). Pengaruh masing-masing variabel bebas terhadap variabel terikat diukur melalui uji parsial (uji t)[17].

Nilai thitung akan dibandingkan dengan nilai tabel.

1. H0 diterima apabila -ttabel ≤ thitung ≤ ttabel (dengan tingkat signifikan α = 5%).
2. Ha diterima apabila thitung < - ttabel atau thitung > ttabel (dengan tingkat signifikan α = 5%).

Acuan lainnya melihat dari nilai sig nya.

1. Jika nilai Sig uji t < 0,05 maka H0 ditolak dan H1 diterima. Artinya terdapat pengaruh antara variabel independen terhadap variabel dependen.
2. 2. Jika nilai Sig uji t > 0,05 maka H0 diterima dan H1 ditolak. Artinya tidak ada pengaruh anta variabel independen terhadap variabel dependen

d. Mean Absolute Error (MAE)

Nilai rata-rata dan nilai prediksi yang bernilai mutlak positif dikenal sebagai Mean Absolute Error (MAE)[15].

$$MAE = \left(\frac{1}{n}\right)\sum_{t=1}^{n}|A_t - F_t|$$

Berdasarkan hasil perhitungan algoritma menggunakan python diperoleh nilai MAE sebesar 4171.01.

e. Root Mean Square Error (RMSE)

Untuk mengetahui seberapa besar kesalahan, gunakan root mean square error (RMSE). RMSE merupakan akar dari MSE dan metode berbasis gradien. Hasil prediksi ditampilkan dengan jelas: semakin baik prediksi yang dibuat semakin rendah nilai RMSE [18].

$$RMSE = \sqrt{\left(\left(\frac{1}{n}\right)\sum_{t=1}^{n}(A_t - F_t)^2\right)}$$

Berdasarkan hasil perhitungan algoritma menggunakan python diperoleh nilai MAE sebesar 5831.29

6. Implementasi

```
Prediksi: 3652.90351750234
Interval 95%: (np.float64(-96877.08816302365), np.float64(104182.89519802835))
```

Gambar 7. Interval prediksi pembayran asuransi

Hasil perhitungan interval prediksi menunjukkan bahwa nilai prediksi adalah 3.652,90, dan interval kepercayaan 95 persen berada di antara −96.877,09 dan 104.182,90. Rentang interval yang sangat lebar ini menunjukkan tingkat ketidakpastian prediksi yang tinggi. Ini mungkin disebabkan oleh heteroskedastisitas model dan varians residual yang besar. Akibatnya, meskipun model mampu memberikan nilai prediksi, presisi prediksinya masih rendah dan diperlukan perbaikan lebih lanjut.

KESIMPULAN

Hasil uji determinasi menunjukkan nilai Adjusted R Square sebesar 0,749, yang menunjukkan bahwa usia, sex_encoded, bmi, anak-anak, dan smoker_encoded dapat menyumbang 74,9% dari variasi variabel charges. Variabel lain di luar model penelitian dijelaskan 25,1% dari variasi tersebut. Hasil uji parsial (uji t) menunjukkan bahwa usia, rasio massa tubuh (BMI), jumlah anak, status perokok, dan jenis kelamin memiliki pengaruh yang signifikan terhadap besarnya premi asuransi kesehatan. Sebaliknya, variabel usia dan jenis kelamin tidak memiliki pengaruh yang signifikan terhadap pembayaran premi. Dan dilihat dari uji F secara keseluruhan variabel dependen dan independen memiliki pengaruh yang signifikan satu sama lain. Beberapa variabel menunjukkan pelanggaran pada asumsi normalitas residual dan heteroskedastisitas, menurut model regresi. Evaluasi kinerja model menggunakan MAE dan RMSE menunjukkan tingkat kesalahan prediksi yang signifikan, sehingga hasil prediksi harus ditafsirkan dengan hati-hati. Dibandingkan dengan prediksi individual yang sangat presisi, model ini lebih tepat digunakan untuk estimasi umum karena hasil interval prediksi yang relatif lebar menunjukkan bahwa prediksi individual memiliki tingkat ketidakpastian yang tinggi.

Untuk menjelaskan perbedaan dalam pembayaran premi asuransi kesehatan swasta di Indonesia, model regresi linear berganda sangat cocok. Nilai R Square sebesar 0,749 menunjukkan bahwa usia, jenis kelamin, indeks massa tubuh (BMI), jumlah anak, dan status perokok dapat bertanggung jawab atas sebagian besar variasi pada variabel charges. Menurut hasil uji parsial (uji t), usia, rasio massa tubuh (BMI), jumlah anak, dan status perokok memiliki dampak signifikan terhadap besarnya pembayaran premi asuransi kesehatan. Sebaliknya, variabel jenis kelamin tidak memiliki dampak signifikan. Hasil pengujian asumsi klasik menunjukkan bahwa model tidak mengalami multikolinearitas. Namun, asumsi normalitas residual dan heteroskedastisitas telah dibatalkan. Ada tingkat kesalahan prediksi yang signifikan, jadi hasil prediksi harus diinterpretasikan dengan hati-hati. Ini ditunjukkan oleh evaluasi kinerja model menggunakan MAE dan RMSE.

REFERENCES

[1] B. A. B. Iv, P. Asuransi, D. A. N. Sistem, and P. Kapitasi, "Bab iv. perjalanan asuransi kesehatan dan sistem pembayaran kapitasi di indonesia," pp. 86–133.

[2] A. S. Dewi, "Tinjauan Terhadap Peranan Asuransi Kesehatan dalam Menanggulangi Risik Tertanggung Pada PT. Prudential Life Assurance," *Juripol (Jurnal Institusi Politeknik Ganesha Medan)*, vol. 7, no. 2, pp. 22–33, 2024.

[3] R. Umaroh and L. Listiono, "Dampak Guncangan Kesehatan Terhadap Konsumsi Rumah Tangga Di Indonesia: Adakah Peran Asuransi Kesehatan Dan Kredit Mikro?," *Jurnal Ekonomi Kesehatan Indonesia*, vol. 8, no. 1, p. 11, 2023, doi: 10.7454/eki.v8i1.6227.

[4] J. E. Trisakti *et al.*, "FAKTOR- FAKTOR YANG MEMPENGARUHI KEPUTUSAN PEMBELIAN untuk mengelola keuangan agar dapat mencapai tujuan & kesejahteraan di masa depan . Adapun yang tingkat akses terhadap layanan keuangan cukup tinggi , namun tidak dibarengi dengan tingkat literasi," vol. 3, no. 1, pp. 291–302, 2023.

[5] A. Heryana, "Asuransi Kesehatan & Managed Care".

[6] D. I. Sumantiawan and U. N. Karangturi, "Metode Analasis Menggunakan Algoritma Random Forest," vol. 1, no. 1, pp. 1–8, 2024, [Online]. Available: https://journal.unuha.ac.id/index.php/JICode/article/view/3294/859

[7] R. Pratama, M. I. Herdiansyah, D. Syamsuar, and A. Syazili, "Prediksi Customer Retention Perusahaan Asuransi Menggunakan Machine Learning," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, pp. 96–104, 2023, doi: 10.32736/sisfokom.v12i1.1507.

[8] A. R. and I. F., "Implementasi Metode Regresi Linear Berganda Pada Sistem Prediksi Jumlah Tonase Kelapa Sawit di PT . Paluta Inti Sawit," *Jurnal Pendidikan Tambusai*, vol. 7, no. 1, pp. 2926–2934, 2023.

[9] A. Herdiana, "Studi Kasus Kemiskinan Di Indonesia Level Provinsi Dan Faktor-Faktor Yang Mempengaruhinya Menggunakan Regresi Linear Berganda," *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, vol. 10, no. 1, pp. 89–93, 2022, doi: 10.24252/msa.v10i1.23361.

[10] P. Aulia Azhar, M. Arya Pratama, and R. Fitriani, "Prediksi Harga Mobil Audi Bekas Menggunakan Model Regresi Linear dengan Framework Streamlit," *Journal of Technology and Informatics (JoTI)*, vol. 6, no. 1, pp. 22–28, 2024, doi: 10.37802/joti.v6i1.763.

[11] C. Tirta, S. Gunawan, C. Jauhari, and F. Nadya, "Model Penentuan Harga Premi Asuransi Kesehatan dengan Regresi Linear Berganda," no. April, 2022.

[12] RERE and ardy, "122-126-1-Pb," *Manajemen*, vol. 6, pp. 1–10, 2020.

[13] I. Susanti and F. Saumi, "Penerapan Metode Analisis Regresi Linear Berganda Untuk Mengatasi Masalah Multikolinearitas Pada Kasus Indeks Pembangunan Manusia (Ipm) Di Kabupaten Aceh Tamiang," *Gamma-Pi: Jurnal Matematika dan Terapan*, vol. 4, no. 2, pp. 38–42, 2022.

[14] M. Janah and A. Y. Kartini, "Penerapan Metode Regresi Linier Berganda Pada Kasus Balita Gizi Buruk Di Kabupaten Bojonegoro," *Jurnal Statistika dan Komputasi*, vol. 1, no. 2, pp. 74–82, 2022, doi: 10.32665/statkom.v1i2.1170.

[15] P. S. Matematika, F. Sains, U. Kristen, and S. Wacana, "JAMBURA JOURNAL OF PROBABILITY AND STATISTICS Volume 2 Nomor 1, Mei 2021," vol. 2, no. January 2020, pp. 17–29, 2021.

[16] G. Mardiatmoko, "Pentingnya Uji Asumsi Klasik Pada Analisis Regresi Linier Berganda (Studi Kasus Penyusunan Persamaan Allometrik Kenari Muda [Canarium Indicum L.])," *BAREKENG: Jurnal Ilmu*

Matematika dan Terapan, vol. 14, no. 3, pp. 333–342, 2020, doi: 10.30598/barekengvol14iss3pp333-342.

[17] Z. Iba and A. Wardhana, *Regresi Linier Sederhana dan Berganda*, no. July. 2024.

[18] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," *Journal of Applied Informatics and Computing*, vol. 5, no. 1, pp. 95–102, 2021, doi: 10.30871/jaic.v5i1.3184.