

ANALISIS PENGARUH FAKTOR MEDIS TERHADAP PERKEMBANGAN PENYAKIT MENGGUNAKAN METODE REGRESI LINIER SECARA SIMULTAN DAN PARSIAL

Tresi Aprilia¹⁾, Supriyanto ²⁾, Muhammad Achsin Samas ¹⁾, Hafni Aulida ¹⁾

¹⁾ Dosen Program Studi Teknik Informatika

²⁾ Dosen Program Studi Keuangan dan Perbankan Digital

^{1.)} Universitas Selamat Sri

²⁾ Akademi Bisnis dan Keuangan Primaniyarta

Email : tresiaprilia98@gmail.com, supriyanto@abkprimaniyarta.ac.id,
samasachsins@gmail.com , hafniaulida@gmail.com

ABSTRAK

Perkembangan penyakit dipengaruhi oleh berbagai faktor medis yang dapat digunakan sebagai indikator dalam memprediksi kondisi kesehatan pasien. Penelitian ini bertujuan untuk menganalisis pengaruh faktor medis berupa usia (AGE), kadar low-density lipoprotein (LDL), dan high-density lipoprotein (HDL) terhadap perkembangan penyakit setelah satu tahun pengamatan (Y) menggunakan metode regresi linier secara simultan dan parsial. Data yang digunakan merupakan data medis pasien yang mencakup variabel perkembangan penyakit sebagai variabel dependen dan faktor-faktor medis sebagai variabel independen. Tahapan penelitian meliputi analisis deskriptif, visualisasi data menggunakan scatter plot, serta pembentukan model regresi linier sederhana dan regresi linier berganda. Hasil analisis menunjukkan bahwa variabel HDL memiliki hubungan negatif yang lebih kuat terhadap perkembangan penyakit dibandingkan AGE dan LDL. Sementara itu, variabel AGE menunjukkan hubungan positif namun relatif lemah, sedangkan LDL memiliki pengaruh yang sangat kecil terhadap perkembangan penyakit. Analisis simultan menunjukkan bahwa AGE, LDL, dan HDL secara bersama-sama berkontribusi dalam menjelaskan variasi perkembangan penyakit, meskipun kemampuan model masih terbatas. Hasil penelitian ini menunjukkan bahwa faktor medis tertentu memiliki tingkat pengaruh yang berbeda terhadap perkembangan penyakit. Oleh karena itu, model regresi linier dapat digunakan sebagai alat analisis untuk mengidentifikasi faktor-faktor yang berperan dalam perkembangan penyakit dan mendukung pengambilan keputusan di bidang kesehatan.

Kata Kunci: Penambangan Data, Regresi Linier, Prediksi Penyakit, Faktor Risiko Kesehatan, Analisis Medis.

ABSTRACT

Disease progression is influenced by various medical factors that can serve as indicators for predicting a patient's health condition. This study aims to analyze the effect of medical factors—namely age (AGE), low-density lipoprotein (LDL), and high-density lipoprotein (HDL)—on disease progression measured one year after baseline (Y), using simultaneous and partial linear regression. The dataset consists of 100 de-identified patient records drawn from a publicly available medical dataset, comprising disease progression as the dependent variable and the aforementioned medical factors as independent variables. The research stages include descriptive analysis, data visualization using scatter plots, and the construction of simple and multiple linear regression models. The results show that HDL has a stronger negative relationship with disease progression than AGE and LDL. AGE shows a weak positive relationship, while LDL has a very small effect on disease progression. The simultaneous analysis indicates that AGE, LDL, and HDL jointly contribute to explaining the variation in disease progression, although the model's explanatory power remains limited ($R^2 = 0.168$). These findings suggest that different medical factors exert different levels of influence on disease progression. Therefore, linear regression can serve as an analytical tool for identifying factors that play a role in disease progression and for supporting decision-making in the health sector.

Keywords: Data Mining, Linear Regression, Disease Prediction, Health Risk Factors, Medical Analysis.

1. Pendahuluan

Perkembangan penyakit dipengaruhi oleh berbagai faktor medis, seperti usia, kadar kolesterol, dan kadar gula darah [1]. Analisis terhadap faktor-faktor tersebut diperlukan untuk mengetahui variabel yang paling berpengaruh terhadap perkembangan penyakit sehingga dapat digunakan sebagai dasar pengambilan keputusan dalam bidang Kesehatan [2]. Metode regresi linier dapat digunakan untuk mengukur pengaruh masing-masing faktor secara parsial maupun secara simultan terhadap perkembangan penyakit [3].

Penggunaan regresi linier dalam penelitian ini tidak hanya ditujukan untuk memperoleh nilai prediksi, tetapi terutama untuk mengetahui arah dan besarnya hubungan antara faktor medis dengan perkembangan penyakit. Koefisien regresi yang bernilai positif menunjukkan kecenderungan peningkatan nilai perkembangan penyakit ketika nilai faktor medis meningkat, sedangkan koefisien negatif menunjukkan kecenderungan hubungan yang berlawanan. Dengan demikian, regresi linier memberikan dasar kuantitatif untuk membandingkan kontribusi beberapa faktor medis terhadap perkembangan penyakit. Pendekatan ini juga memungkinkan peneliti mengevaluasi seberapa besar variasi perkembangan penyakit yang dapat dijelaskan oleh faktor-faktor medis yang dimasukkan ke dalam model.

Salah satu teknik analisis dalam data mining yang sering digunakan untuk mengetahui hubungan antarvariabel numerik adalah regresi linier. Metode ini digunakan untuk memodelkan hubungan antara variabel dependen dengan satu atau lebih variabel independen [4]. Regresi linier tidak hanya berfungsi untuk memprediksi nilai suatu variabel, tetapi juga untuk mengukur besarnya pengaruh masing-masing variabel bebas terhadap variabel terikat [5]. Analisis regresi dapat dilakukan secara parsial untuk mengetahui pengaruh setiap variabel secara individu maupun secara simultan (berganda) untuk mengetahui pengaruh beberapa variabel secara bersama-sama terhadap variabel dependen [6].

Penelitian ini menggunakan dataset medis yang memuat informasi mengenai perkembangan penyakit setelah satu tahun pengamatan (Y) sebagai variabel dependen,

serta AGE, LDL, dan HDL sebagai variabel independen. Analisis dilakukan melalui visualisasi data menggunakan *scatter plot* untuk mengidentifikasi pola hubungan antarvariabel, kemudian dilanjutkan dengan pembentukan model regresi linier parsial dan regresi linier simultan. Pendekatan ini diharapkan mampu memberikan gambaran mengenai arah hubungan, besar pengaruh masing-masing faktor medis, serta kemampuan model dalam menjelaskan variasi perkembangan penyakit.

Rumusan Masalah Bagaimana pengaruh usia (AGE), LDL, dan HDL terhadap perkembangan penyakit (Y) secara parsial? Bagaimana pengaruh usia (AGE), LDL, dan HDL terhadap perkembangan penyakit (Y) secara simultan? Variabel manakah yang memiliki pengaruh paling dominan terhadap perkembangan penyakit?

Tujuan penelitian Menganalisis pengaruh usia (AGE), LDL, dan HDL terhadap perkembangan penyakit secara parsial. Menganalisis pengaruh usia (AGE), LDL, dan HDL terhadap perkembangan penyakit secara simultan. Menentukan faktor medis yang paling dominan memengaruhi perkembangan penyakit.

Hasil penelitian ini diharapkan dapat memberikan informasi mengenai faktor medis yang paling berpengaruh terhadap perkembangan penyakit sehingga dapat menjadi bahan pertimbangan dalam pengambilan keputusan di bidang kesehatan. Selain itu, penelitian ini diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang mengembangkan model prediksi penyakit menggunakan metode data mining maupun machine learning dengan variabel dan algoritma yang lebih kompleks.

2. Metode Penelitian

Berdasarkan penelitian yang dilakukan, dataset yang digunakan menggunakan data medis. Sumber dataset yang menggunakan data publik yang ada di *kaggle*. Berikut contoh dataset:

Tabel 1: Dataset Medis

NO	Y	AGE	LDL	HDL	TCH	LTG	GLU
1	128	53	143,8	58	4,0	5,0	91,0
2	102	53	115,2	67	3,0	4,3	100,0
3	302	23	131,4	65	3,0	4,6	91,0

4	198	65	160,6	40	5,0	4,5	84,0
5	95	41	104,4	56	3,0	4,0	76,0
...							
100	90	60	105,4	60,0	3,7	5,6	93,0

Berdasarkan data medis terlampir terdapat 100 data yang berhubungan dengan perkembangan penyakit setelah satu tahun amatan (Y), usia dalam tahun (AGE), *Low- Density Lipoproteins* (LDL), *High-Density Lipoproteins* (HDL), Total *Cholesterol* HDL (TCH), *Triglycerides* (LTG) dan *blood sugar* (GLU).

2.1 PERSAMAAN REGRESI LINIER PARSIAL

Regresi linier parsial adalah analisis regresi yang digunakan untuk mengetahui pengaruh satu variabel independen terhadap variabel dependen secara individual atau terpisah. Analisis ini bertujuan untuk melihat seberapa besar hubungan dan pengaruh masing-masing variabel bebas terhadap variabel terikat tanpa mempertimbangkan variabel bebas lainnya secara bersamaan [7]. Pengujian parsial dalam regresi dilakukan untuk mengetahui pengaruh masing-masing variabel independen terhadap variabel dependen melalui uji t dan interpretasi koefisien regresi [8].

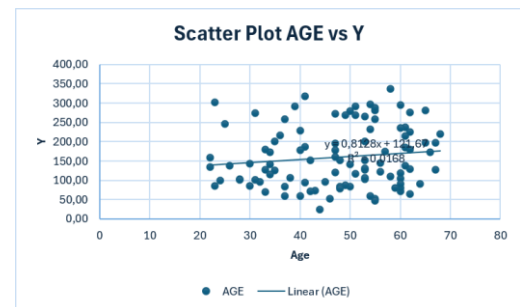
2.2 PERSAMAAN REGRESI LINIER SIMULTAN

Regresi linier simultan atau regresi linier berganda merupakan metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan dua atau lebih variabel independen secara bersama-sama. Analisis ini bertujuan untuk mengetahui seberapa besar pengaruh seluruh variabel independen terhadap variabel dependen serta mengidentifikasi variabel yang paling berkontribusi dalam model [9].

3. Hasil dan Pembahasan

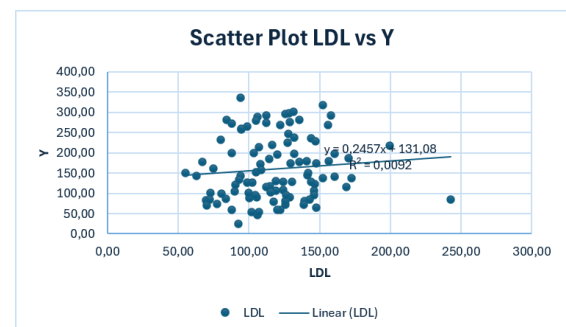
Pada bagian ini harus ditampilkan seluruh data yang diperoleh dengan informasi yang cukup bagaimana data tersebut diperoleh. Analisa statistik boleh disampaikan jika dianggap perlu, atau cukup ditampilkan dalam grafik disertai standar deviasi serta *error bars* dan penjelasan arti signifikansi secara statistik dari standar deviasi maupun *error bars* dalam grafik tersebut.

3.1 Visualisasi Data Dengan Scatter Plot



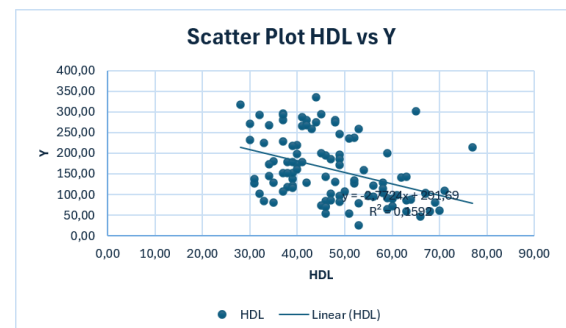
Gambar 1. Scatter Plot Variabel Age vs Y

Berdasarkan gambar 1, output yang diperoleh $Y=121,671+0,813AGE$ dan nilai $R^2 = 0,0168$. Setiap kenaikan usia 1 tahun diperkirakan meningkatkan nilai Y sebesar 0,813 satuan. Namun hubungan ini sangat lemah karena hanya menjelaskan 1,68% variasi Y.



Gambar 2. Scatter Plot Variabel LDL vs Y

Berdasarkan gambar 2, output yang diperoleh $Y=131,078+0,246LDL$ dan nilai $R^2 = 0,0092$. Setiap kenaikan LDL 1 unit diperkirakan meningkatkan Y sebesar 0,246 satuan. Hubungan sangat lemah, kurang dari 1%.



Gambar 3. Scatter Plot Variabel HDL vs Y

Berdasarkan gambar 3, output yang diperoleh $Y=291,687-2,772HDL$ dan nilai $R^2 = 0,1592$. Setiap kenaikan HDL 1 unit diperkirakan menurunkan Y sebesar 2,772 satuan. Hubungan negatif ini paling kuat dibanding AGE dan LDL.

3.2. Model Persamaan Regresi Linier Parsial

Tabel 2: Regression Statistics Variabel Age

REGRESSION STATISTICS	
Multiple R	0,129575961
R Square	0,01678993
Adjusted R Square	0,006757174
Standard Error	77,56710194
Observations	100

Berdasarkan Tabel 2, hasil regresi linier variabel AGE menunjukkan nilai Multiple R sebesar 0,1296 yang mengindikasikan hubungan linear yang sangat lemah antara usia dan perkembangan penyakit (Y). Nilai R Square sebesar 0,0168 menunjukkan bahwa AGE hanya mampu menjelaskan sekitar 1,68% variasi perkembangan penyakit, sedangkan 98,32% lainnya berkaitan dengan faktor lain di luar model. Nilai Adjusted R Square sebesar 0,0068 menunjukkan kemampuan penjelasan model setelah penyesuaian masih sangat rendah, sedangkan Standard Error sebesar 77,57 menunjukkan adanya penyimpangan yang cukup besar antara nilai aktual dan nilai yang diprediksi. Model ini menggunakan 100 observasi, sehingga secara keseluruhan AGE secara individual memiliki kemampuan yang terbatas dalam menjelaskan perkembangan penyakit.

Tabel 3: Anova Test Variabel Age

	df	ss	MS	F	Significance F
Regression	1	10068,9	10068,940	1,674	0,1988
Residual	98	589632,2	6016,655		
Total	99	599701,2	-		

Berdasarkan Tabel 3, hasil uji ANOVA variabel AGE menunjukkan nilai F sebesar 1,6735 dengan Significance F sebesar 0,1988. Nilai Significance F tersebut lebih besar dari tingkat signifikansi 0,05, sehingga model regresi AGE secara keseluruhan tidak menunjukkan pengaruh yang signifikan secara statistik terhadap perkembangan penyakit (Y). Nilai SS Regression sebesar 10.068,94 menunjukkan variasi Y yang dapat dijelaskan oleh AGE, sedangkan SS Residual sebesar 589.632,26 menunjukkan variasi Y yang tidak dapat dijelaskan oleh model. Dengan demikian, berdasarkan uji ANOVA, AGE secara individual belum cukup kuat untuk menjelaskan variasi perkembangan penyakit pada 100 observasi yang dianalisis.

Tabel 4: Hasil Summary Output Variabel Age

	Coefficients	Standard Error	t Stat	P-value
Intercept	121,67	30,8	3,95	0,00
X Variable 1	0,8127	0,6	1,29	0,20

Berdasarkan Tabel 4, hasil analisis regresi variabel AGE menunjukkan nilai intercept sebesar 121,67, sedangkan koefisien X Variable 1 (AGE) sebesar 0,8127 dengan t Stat sebesar 1,29 dan P-value sebesar 0,20. Koefisien AGE yang bernilai positif menunjukkan bahwa setiap kenaikan usia sebesar satu tahun diperkirakan meningkatkan nilai perkembangan penyakit (Y) sebesar 0,8127 satuan. Namun, karena nilai P-value $0,20 > 0,05$, pengaruh AGE terhadap perkembangan penyakit secara parsial tidak signifikan secara statistik. Dengan demikian, meskipun terdapat kecenderungan hubungan positif antara usia dan perkembangan penyakit, bukti statistik pada data penelitian ini belum cukup untuk menyatakan bahwa AGE berpengaruh signifikan terhadap Y. Hasil ini konsisten dengan nilai Significance F pada uji ANOVA sebelumnya sebesar 0,1988.

Tabel 5: Regression Statistics Variabel LDL

REGRESSION STATISTICS	
Multiple R	0,09584
R Square	0,00919
Adjusted R Square	-0,0009
Standard Error	77,8665
Observations	100

Berdasarkan Tabel 5, hasil Regression Statistics variabel LDL menunjukkan nilai Multiple R sebesar 0,0958, yang mengindikasikan bahwa hubungan linear antara LDL dan perkembangan penyakit (Y) tergolong sangat lemah. Nilai R Square sebesar 0,0092 menunjukkan bahwa LDL hanya mampu menjelaskan sekitar 0,92% variasi perkembangan penyakit, sedangkan sekitar 99,08% variasi lainnya berkaitan dengan faktor di luar model. Nilai Adjusted R Square sebesar -0,0009 menunjukkan bahwa setelah penyesuaian, kemampuan model dalam menjelaskan variasi Y sangat rendah. Sementara itu, Standard Error sebesar 77,87 menunjukkan adanya penyimpangan yang relatif besar antara nilai aktual dan nilai yang diprediksi model. Analisis ini menggunakan 100 observasi, sehingga secara keseluruhan LDL secara individual memiliki kemampuan yang sangat terbatas dalam menjelaskan perkembangan penyakit.

Tabel 6: Anova Test Variabel LDL

	df	ss	MS	F	Significance F
Regression	1	5508,8	5508,84	0,9086	0,3428
Residual	98	594192,3	6063,2		
Total	99	599701,2	-		

Berdasarkan Tabel 6, hasil uji ANOVA variabel LDL menunjukkan nilai F sebesar 0,9086 dengan Significance F sebesar 0,3428. Karena nilai Significance F $0,3428 > 0,05$, maka model regresi LDL secara keseluruhan tidak menunjukkan pengaruh yang signifikan secara statistik terhadap perkembangan penyakit (Y). Nilai SS Regression sebesar 5.508,84 menunjukkan variasi Y yang dapat dijelaskan oleh LDL, sedangkan SS Residual sebesar 594.192,36 menunjukkan variasi Y yang tidak dapat dijelaskan oleh model. Dengan demikian, berdasarkan uji ANOVA, LDL secara individual belum memberikan bukti yang cukup untuk menjelaskan variasi perkembangan penyakit secara signifikan pada 100 observasi yang dianalisis.

Tabel 7: Hasil Summary Output Variabel LDL

	Coefficients	Standard Error	t Stat	P-value
Intercept	131,077555	31,55	4,2	7E-05
X Variable 1	0,2457453	0,26	1,0	3E-01

Berdasarkan Tabel 7, hasil summary output variabel LDL menunjukkan nilai intercept sebesar 131,0776, sedangkan koefisien X Variable 1 (LDL) sebesar 0,2457 dengan t Stat sebesar 1,0 dan P-value sebesar 0,30. Koefisien LDL yang bernilai positif menunjukkan bahwa setiap kenaikan LDL sebesar satu unit diperkirakan meningkatkan nilai perkembangan penyakit (Y) sebesar 0,2457 satuan. Namun, karena nilai P-value $0,30 > 0,05$, maka pengaruh LDL terhadap perkembangan penyakit secara parsial tidak signifikan secara statistik. Dengan demikian, meskipun terdapat kecenderungan hubungan positif antara LDL dan perkembangan penyakit, hasil analisis belum memberikan bukti yang cukup bahwa LDL secara individual berpengaruh signifikan terhadap Y. Hasil ini konsisten dengan uji ANOVA pada Tabel 6 yang menghasilkan Significance F sebesar 0,3428

Tabel 8: Regression Statistics Variabel HGL

REGRESSION STATISTICS	
Multiple R	0,39904765
R Square	0,159239027
Adjusted R Square	0,150659833
Standard Error	71,72832325
Observations	100

Berdasarkan Tabel 8, hasil Regression Statistics variabel HDL menunjukkan nilai Multiple R sebesar 0,3990, yang menunjukkan adanya hubungan linear yang relatif lebih kuat dibandingkan AGE dan LDL, meskipun masih tergolong lemah. Nilai R Square sebesar 0,1592 menunjukkan bahwa HDL mampu menjelaskan sekitar 15,92% variasi perkembangan penyakit (Y), sedangkan sekitar 84,08% variasi lainnya berkaitan dengan faktor di luar model. Nilai Adjusted R Square sebesar 0,1507 menunjukkan bahwa setelah penyesuaian, kemampuan model dalam menjelaskan variasi Y masih sekitar 15,07%. Sementara itu, Standard Error sebesar 71,73 menunjukkan penyimpangan standar antara nilai aktual dan nilai yang diprediksi model. Dengan 100 observasi, hasil ini menunjukkan bahwa dibandingkan AGE dan LDL, HDL memiliki kemampuan yang lebih besar dalam menjelaskan variasi perkembangan penyakit secara individual.

Tabel 9: Anova Test Variabel HGL

	df	ss	MS	F	Significance F
Regression	1	95495,8	95495,829	18,56	3,914E-05
Residual	98	504205,3	5144,9524		
Total	99	599701,2			

Berdasarkan Tabel 9, hasil uji ANOVA variabel HDL menunjukkan nilai F sebesar 18,56 dengan Significance F sebesar $3,914 \times 10^{-5}$ (0,00003914). Karena nilai Significance F lebih kecil dari 0,05, maka model regresi HDL secara keseluruhan berpengaruh signifikan secara statistik terhadap perkembangan penyakit (Y). Nilai SS Regression sebesar 95.495,83 menunjukkan variasi Y yang dapat dijelaskan oleh HDL, sedangkan SS Residual sebesar 504.205,30 menunjukkan variasi yang belum dapat dijelaskan oleh model. Dengan demikian, HDL memiliki hubungan yang signifikan dengan perkembangan penyakit pada 100 observasi yang dianalisis dan, dibandingkan AGE dan LDL, memberikan kontribusi yang lebih kuat

dalam menjelaskan variasi Y. Hasil ini juga konsisten dengan nilai R Square sebesar 0,1592 pada Tabel 8.

Tabel 10: Hasil Summary Output Variabel HGL

	Coefficients	Standard Error	t Stat	P-value
Intercept	291,686	31,35	9,31	3,9E-15
X Variable 1	-2,772	0,64	-4,31	3,9E-05

Berdasarkan Tabel 10, hasil summary output variabel HDL menunjukkan nilai intercept sebesar 291,686, sedangkan koefisien X Variable 1 (HDL) sebesar -2,772 dengan t Stat sebesar -4,31 dan P-value sebesar $3,9 \times 10^{-5}$ (0,000039). Koefisien HDL yang bernilai negatif menunjukkan bahwa setiap kenaikan HDL sebesar satu unit diperkirakan menurunkan nilai perkembangan penyakit (Y) sebesar 2,772 satuan. Karena nilai P-value $0,000039 < 0,05$, maka HDL memiliki pengaruh yang signifikan secara statistik terhadap perkembangan penyakit secara parsial. Dengan demikian, HDL merupakan variabel medis yang memiliki hubungan negatif dan signifikan dengan Y, serta berdasarkan nilai koefisiennya memiliki kontribusi yang lebih kuat dibandingkan AGE dan LDL dalam model regresi parsial. Hasil ini konsisten dengan uji ANOVA pada Tabel 9 yang menghasilkan Significance F sebesar $3,914 \times 10^{-5}$ dan nilai R Square sebesar 0,1592.

3.3. Model Persamaan Regresi Linier Simultan

Tabel 11: Regression Statistics Variabel Age, LDL dan HDL

REGRESSION STATISTICS	
Multiple R	0,409817371
R Square	0,167950277
Adjusted R Square	0,141948724
Standard Error	72,09521943
Observations	100

Berdasarkan Tabel 11, hasil Regression Statistics secara simultan untuk variabel AGE, LDL, dan HDL menunjukkan nilai Multiple R sebesar 0,4098, yang menunjukkan adanya hubungan linear positif dengan kekuatan relatif lemah antara ketiga faktor medis secara bersama-sama dan perkembangan penyakit (Y). Nilai R Square sebesar 0,1680 menunjukkan bahwa AGE, LDL, dan HDL secara bersama-sama mampu menjelaskan sekitar 16,80% variasi perkembangan penyakit, sedangkan sekitar 83,20% variasi lainnya berkaitan dengan faktor di luar model. Nilai Adjusted R Square

sebesar 0,1419 menunjukkan bahwa setelah penyesuaian terhadap jumlah variabel dan observasi, kemampuan model dalam menjelaskan variasi Y sebesar sekitar 14,19%. Sementara itu, Standard Error sebesar 72,0952 menunjukkan besarnya penyimpangan standar antara nilai aktual dan nilai yang diprediksi model.

Tabel 12: Anova Test Variabel Age, LDL dan HDL

	df	ss	MS	F	Significance F
Regression	3	100720,32539	33573,32539	6,46	0,0005
Residual	96	498981,720665	5197,720665		
Total	99	599701,32539	33573,32539		

Berdasarkan Tabel 12, hasil uji ANOVA secara simultan untuk variabel AGE, LDL, dan HDL menunjukkan nilai F sebesar 6,46 dengan Significance F sebesar 0,0005. Karena nilai Significance F $0,0005 < 0,05$, maka model regresi secara simultan signifikan secara statistik, sehingga AGE, LDL, dan HDL secara bersama-sama memiliki hubungan yang signifikan dengan perkembangan penyakit (Y). Nilai SS Regression sebesar 100.720 menunjukkan variasi Y yang dapat dijelaskan oleh ketiga variabel independen, sedangkan SS Residual sebesar 498.981 menunjukkan variasi Y yang belum dapat dijelaskan oleh model. Dengan demikian, meskipun masing-masing variabel tidak semuanya signifikan secara parsial, kombinasi AGE, LDL, dan HDL secara bersama-sama memberikan kontribusi yang signifikan terhadap perkembangan penyakit pada 100 observasi yang dianalisis. Hasil ini mendukung nilai R Square sebesar 0,16795, yang menunjukkan bahwa model mampu menjelaskan sekitar 16,80% variasi perkembangan penyakit.

Tabel 13: Hasil Summary Output Variabel Age, LDL dan HDL

	Coefficients	Standard Error	t Stat	P-value
Intercept	266,63	54,01	4,94	3E-06
X Variable 1	0,60	0,60	1,00	3E-01
X Variable 2	-0,04	0,25	-0,17	9E-01
X Variable 3	-2,74	0,67	-4,11	8E-05

Berdasarkan persamaan model regresi linier simultan didapatkan hasil setiap kenaikan usia 1 tahun diperkirakan meningkatkan perkembangan penyakit sebesar 0,813 satuan. Setiap kenaikan LDL 1 diperkirakan

meningkatkan perkembangan penyakit sebesar 0,246 satuan. Setiap kenaikan HDL 1 unit diperkirakan menurunkan perkembangan penyakit sebesar 2,772 satuan. Jika LDL dan HDL dianggap tetap, kenaikan usia 1 tahun meningkatkan Y sebesar 0,599 satuan. Jika AGE dan HDL tetap, kenaikan LDL 1 unit akan menurunkan Y sebesar 0,043 satuan. Jika AGE dan LDL tetap, kenaikan HDL 1 unit akan menurunkan Y sebesar 2,736 satuan.

4. Kesimpulan

Berdasarkan hasil analisis regresi linear berganda, diperoleh model persamaan:

$$Y = 266,627 + 0,599(\text{AGE}) - 0,043(\text{LDL}) - 2,736(\text{HDL})$$

Model ini menunjukkan hubungan antara variabel usia (AGE), kadar LDL, dan kadar HDL terhadap perkembangan penyakit (Y) setelah satu tahun pengamatan. Koefisien AGE bernilai positif, yang mengindikasikan bahwa semakin bertambah usia seseorang maka nilai perkembangan penyakit cenderung meningkat. Sebaliknya, koefisien HDL bernilai negatif, yang menunjukkan bahwa semakin tinggi kadar HDL maka perkembangan penyakit cenderung menurun. Di antara ketiga variabel yang dianalisis, HDL memiliki pengaruh paling besar terhadap perubahan nilai Y karena memiliki nilai koefisien terbesar secara absolut, yaitu -2,736. Artinya, setiap kenaikan satu satuan HDL diperkirakan akan menurunkan nilai perkembangan penyakit sebesar 2,736 satuan dengan asumsi variabel lain tetap. Sementara itu, pengaruh LDL dalam model relatif kecil dengan koefisien sebesar -0,043, sehingga kontribusinya terhadap perubahan Y tidak sebesar AGE maupun HDL. Nilai koefisien determinasi (R^2) sebesar 0,168 itu mewakili bahwa model mampu menjelaskan sekitar 16,8% variasi perkembangan penyakit berdasarkan variabel AGE, LDL, dan HDL. Dengan kata lain, masih terdapat sekitar 83,2% variasi perkembangan penyakit yang dipengaruhi oleh faktor-faktor lain yang tidak dimasukkan ke dalam model. Oleh karena itu, model regresi ini cukup berguna untuk menggambarkan hubungan awal antarvariabel, tetapi belum dapat digunakan sebagai model prediksi yang sangat kuat karena kemampuan penjelasannya masih relatif rendah.

Berdasarkan hasil analisis, penelitian lanjutan disarankan untuk memasukkan variabel lain yang berkaitan dengan kondisi kesehatan, seperti indeks massa tubuh (BMI), tekanan darah, kadar glukosa, maupun faktor-faktor lainnya. Penambahan variabel tersebut diharapkan dapat meningkatkan kemampuan model dalam menjelaskan variasi perkembangan penyakit dan menghasilkan model prediksi yang lebih akurat.

5. Daftar Pustaka

- [1] J. Amadeus, M. Sangtutu, A. A. Lee, A. Kusuma, and S. Elfi, "Pemeriksaan Kolesterol, Gula Darah, dan Asam Urat Sebagai Upaya Deteksi Dini Penyakit Tidak Menular," vol. 3, no. 8, pp. 4408–4418, 2025.
<https://doi.org/10.59837/jpmmba.v3i8.3281>
<https://jurnalpengabdianmasyarakatbangsa.com/index.php/jpmmba/article/view/3281>
- [2] L. Amdan and M. Rafi, "Analisis Faktor-Faktor Yang Mempengaruhi Pertumbuhan Ekonomi di Indonesia," vol. 3, no. 1, pp. 108–119, 2023. DOI: <https://doi.org/10.56799/ekoma.v3i1.2089>
<https://al-haramjournal.co.id/EKOMA/article/view/2089>
- [3] H. A. Budiman, A. K. Mutaqin, P. Statistika, F. Matematika, and P. Alam, "Analisis Regresi Linier Berganda untuk Menentukan Faktor-Faktor Penyebab Demam Berdarah Dengue di Provinsi Jawa Barat Tahun 2023," 2023.
<https://doi.org/10.29313/bcss.v5i2.19971>
<https://proceedings.unisba.ac.id/index.php/BCSS/article/view/19971>
- [4] U. Selamat and S. Kendal, "Analisis Perbandingan Metode Deep Learning Dalam," vol. 22, pp. 16–26, 2025. DOI: <https://doi.org/10.58839/jd.v22i1.Juli.1544>
<https://ojs.ustj.ac.id/dinamis/article/view/1544>
- [5] B. Sihombing, A. P. Mangande, and P. S. Akuntansi, "Honda Beat Street Pada Pt Astra Honda Area Tanah Hitam," vol. 17, no. 2, pp. 17–24, 2019.
<https://ojs.ustj.ac.id/dinamis/article/view/760>
- [6] F. S. Irma Susanti 1, "Gamma-Pi: Jurnal Matematika dan Terapan Gamma-Pi: Jurnal Matematika dan Terapan," vol. 4, pp. 10–17, 2022. DOI: <https://doi.org/10.33059/jgp.v4i2.5591>
<https://ejurnalunsam.id/index.php/jgp/article/view/5591>

- [7] A. I. Fairuzsyifa, Y. S. Nugroho, J. T. Informatika, U. M. Surakarta, and U. M. Surakarta, "Analisis regresi linier berganda pengaruh minat calon mahasiswa di universitas muhammadiyah surakarta menggunakan python," pp. 265–272, 2022. DOI: <https://doi.org/10.33795/jip.v10i2.5014> <https://jurnal.polinema.ac.id/index.php/jip/article/view/5014>
- [8] F. R. Aflah and M. F. Hamdani, "Penerapan Regresi Linier Berganda dalam Menilai Hubungan Antar Variabel dalam Penelitian Kuantitatif," vol. 5, pp. 4195–4211, 2025. DOI: <https://doi.org/10.31004/innovative.v5i3.19319> <https://j-innovative.org/index.php/Innovative/article/view/19319>
- [9] J. Informatika, D. Rekayasa, K. Jakakom, C. A. Rahmat, Y. Novianto, and C. A. Rahmat, "Penerapan Metode Regresi Linier Berganda Untuk Mengestimasi Laju Pertumbuhan Penduduk Kabupaten Musi Banyuasin Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," vol. 3, no. April, pp. 359–369, 2023. DOI: <https://doi.org/10.33998/jakakom.2023.3.1.732> <https://ejournal.unama.ac.id/index.php/jakakom/article/view/732>