

Pemodelan Regresi Logistik Biner Bayesian (Studi Kasus: Pasien Kanker Leher Rahim RSUD Sayang Cianjur)

Tsania Dinni Islamiati^{*}, Abdul Kudus

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Indonesia.

*tsaniadinniislamiati@gmail.com, akudus69@unisba.ac.id

Abstract. In this study, we will discuss Bayesian binary logistic regression modeling in a case study of cervical cancer patients at Sayang Cianjur Hospital. Bayesian method is a simple mathematical formula used to simplify conditional probabilities, this method uses parameter estimation techniques by combining likelihood and prior distributions. The results of this study obtained seven independent variables that are significant to the incidence of cervical cancer. The variables are age, age of marriage, number of partners, weight, Pap smear test, history of contraception and occupation. Then the model obtained is:

$$\text{logit}(\hat{p}) = -0,164 + 0,051X_1 - 0,281X_2 + 1,949X_4 - 0,065X_5 + 5,915X_7 \\ + 0,928X_8 - 1,882X_{14}$$

Keywords: Cervical cancer, Bayesian method, logistic regression.

Abstrak. Le Pada penelitian ini akan dibahas tentang pemodelan regresi logistik biner Bayesian pada studi kasus pasien kanker leher rahim RSUD Sayang Cianjur. Metode Bayesian merupakan rumus matematika sederhana yang digunakan untuk menyederhanakan probabilitas bersyarat, metode ini menggunakan teknik estimasi parameter dengan menggabungkan *likelihood* dan distribusi prior. Hasil penelitian ini diperoleh tujuh variabel bebas yang signifikan terhadap kejadian kanker leher rahim. Variabel tersebut adalah usia, usia menikah, jumlah pasangan, berat badan, tes papsmear, riwayat kontrasepsi dan pekerjaan. Maka model yang diperoleh sebagai berikut:

$$\text{logit}(\hat{p}) = -0,164 + 0,051X_1 - 0,281X_2 + 1,949X_4 - 0,065X_5 + 5,915X_7 \\ + 0,928X_8 - 1,882X_{14}$$

Kata Kunci: Kanker leher rahim, metode Bayesian, regresi logistik.

1. Pendahuluan

Kanker Leher rahim merupakan kanker yang terjadi pada leher rahim yang disebabkan oleh infeksi *Human Papiloma Virus* (HPV). Berdasarkan data Dinas Kesehatan Jawa Barat tahun 2019, wilayah dengan persentase hasil pemeriksaan IVA tertinggi di Jawa Barat yakni Kabupaten Cianjur sebesar 5,4%. Banyak faktor-faktor yang mempengaruhi kanker leher rahim, pada penelitian ini akan dilakukan pemodelan terhadap faktor-faktor yang mempengaruhi kanker leher rahim, sehingga penelitian ini dapat digunakan untuk referensi dalam menurunkan resiko kanker leher rahim[1].

Data yang digunakan dalam penelitian ini adalah data pasien poli kandungan RSUD Sayang Cianjur bulan oktober-desember 2020. Pada penelitian ini variabel terikat yang digunakan bersifat biner yaitu terkena kanker leher rahim atau tidak. Untuk memodelkan variabel terikat yang bersifat biner salah satu metode yang dapat digunakan adalah regresi logistik biner. Untuk menduga parameter dalam regresi logistik biasanya menggunakan metode

maksimum *likelihood*, namun pada penelitian ini akan digunakan metode Bayesian. Metode Bayesian menggunakan fungsi *likelihood* dan distribusi *prior* untuk mengestimasi parameter model. Distribusi *prior* adalah distribusi awal yang digunakan untuk memberikan informasi mengenai parameter. Distribusi *prior* dari parameter terlebih dahulu ditentukan untuk selanjutnya mendapatkan distribusi *posterior*. Hal yang sulit dalam metode Bayesian adalah mengetahui bagaimana bentuk distribusi parameter tersebut. Maka dari itu untuk menduga parameter pada metode Bayesian melalui metode *Markov Chain Monte Carlo* (MCMC) dengan algoritma *Hamiltonian Monte Carlo* (HMC). Kelebihan mengestimasi parameter dengan metode Bayesian adalah tidak perlu diketahui distribusi *prior* dari populasi dan sampel yang digunakan dapat berukuran kecil [2].

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah “Bagaimana model regresi logistik biner Bayesian yang diterapkan pada data pasien poli kandungan RSUD Sayang Cianjur dan faktor-faktor apa saja yang mempengaruhi terjadinya kanker leher rahim pada pasien poli kandungan RSUD Sayang Cianjur. Adapun tujuan penelitian ini adalah memperoleh model regresi logistik biner Bayesian pada data pasien poli kandungan RSUD Sayang Cianjur sehingga didapat faktor-faktor resiko yang berpengaruh terhadap kanker leher rahim.

2. Metodologi

Distribusi Bernoulli

Suatu variabel acak X dikatakan berdistribusi Bernoulli jika terdapat dua kemungkinan misal sukses dan gagal. Jika peluang sukses dinotasikan dengan p dan peluang gagal dinotasikan q dimana $q = 1 - p$ maka fungsi kepadatan peluang distribusi Bernoulli dapat ditulis [3]:

$$f(x; p) = \begin{cases} p(1-p)^{1-x}, & x = 0, 1 \\ 0, & x \text{ lainnya} \end{cases} \quad (2.1)$$

Dimana x adalah nilai dari variabel acak dan p bernilai $0 \leq p \leq 1$. Variabel acak X berdistribusi Bernoulli dapat dinotasikan $X \sim \text{Bernoulli}(p)$.

Regresi Logistik Biner

Regresi logistik biner yaitu analisis regresi yang digunakan untuk melihat pengaruh variabel terikat (Y) dengan beberapa variabel bebas (X), dimana variabel terikat bersifat biner. Variabel biner yaitu variabel yang mempunyai dua kemungkinan [3]. Model umum regresi logistik biner adalah sebagai berikut:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (2.2)$$

Model regresi logistik biner diperoleh dengan melakukan transformasi logit terhadap persamaan (2.2), sehingga diperoleh sebagai berikut:

$$\text{logit}(\pi(x)) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2.3)$$

Fungsi *likelihood* Model Regresi Logistik Biner

Suatu $x_1, x_2, \dots, x_n$ tetap, fungsi *likelihood* yaitu fungsi dari β dan dinotasikan dengan $L(\beta)$. Jika $X_1, X_2, \dots, X_n$ adalah peubah acak dari $f(x_n; \beta)$ maka fungsi *likelihood* dapat ditulis sebagai berikut [9]:

$$f(y_i | \beta_0, \dots, \beta_p) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \rightarrow l(\beta) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad (2.4)$$

Untuk menentukan nilai $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ maka digunakan bentuk natural dari fungsi *likelihood*, kemudian tentukan turunan pertamanya terhadap setiap parameter, kemudian disyaratkan dengan nol [9].

Metode Bayesian

Pada metode Bayesian, parameter merupakan peubah acak yang mempunyai distribusi tertentu yang disebut dengan distribusi *prior*. Metode Bayesian menghubungkan fungsi *likelihood*

dengan distribusi *prior* untuk mendapatkan distribusi *posterior*. Penduga Bayesian diperoleh dari ekspektasi distribusi *posterior* [2]. Jika $y_1, y_2, \dots, y_n$ adalah sampel acak maka berdasarkan teorema Bayesian, rumus umum distribusi *posterior* dapat dinyatakan sebagai berikut:

$$p(\beta|y) = \frac{p(\beta, y)}{\int p(\beta) \cdot p(y|\beta) d\beta} = \frac{p(\beta) \cdot L(y|\beta)}{p(y)} \quad (2.5)$$

Dikarenakan $p(y)$ adalah konstanta densitas yang tidak mengandung parameter, maka rumus distribusi *posterior* dapat ditulis sebagai berikut:

$$p(\beta|y) \propto p(\beta) \cdot L(y|\beta) \quad (2.6)$$

Untuk mengestimasi parameter Bayesian maka harus menggunakan ekspektasi dari distribusi *posterior* maka rumusnya dapat ditulis sebagai berikut:

$$E(\beta) = \int \beta p(\beta|y) d\beta \quad (2.7)$$

Distribusi Prior

Distribusi prior yaitu distribusi awal yang digunakan untuk memberikan informasi mengenai parameter. Distribusi *prior* terlebih dahulu ditentukan untuk selanjutnya menentukan distribusi *posterior*. Dikarenakan dalam analisis regresi logistik menggunakan distribusi binomial maka distribusi *prior* yang digunakan adalah distribusi beta[4]. Distribusi *prior* beta dapat ditulis sebagai berikut:

$$p(\beta) \propto \beta^{a_1-1} (1 - \beta)^{a_2-1} \quad (2.8)$$

Markov Chain Monte Carlo

Metode *Markov Chain Monte Carlo* (MCMC) adalah metode yang efisien untuk menduga parameter Bayesian melalui pencarian distribusi *posterior*nya. MCMC merupakan proses pembangkitan sampel acak. Adapun algoritma yang digunakan pada penelitian ini adalah algoritma *Hamiltonian Monte Carlo* (HMC). Algoritma HMC mampu mendeteksi korelasi antar parameter[4].

Pengujian Signifikan Parameter

Pada metode Bayesian untuk menguji signifikan parameter dapat menggunakan *credible interval* atau selang kepercayaan [2].

Hipotesis

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0 \text{ dimana } j = 1, 2, \dots, p$$

Jika variabel acak Y memiliki peluang distribusi *posterior* $p(\beta|Y)$ dengan taraf nyata $(1 - \alpha) \times 100\%$ maka interval (β_l, β_u) dapat ditulis sebagai berikut:

$$P(\beta_l < \beta < \beta_u) = \int_{\beta_l}^{\beta_u} p(\beta|Y) d\beta = 1 - \alpha \text{ dimana } 0 \leq \alpha \leq 1 \quad (2.9)$$

Kriteria uji adalah tolak H_0 jika *credible interval* tidak mengandung 0. Maka dapat disimpulkan parameter yang diperoleh berpengaruh signifikan terhadap model.

Kestabilan Nilai Dugaan Parameter Model

Pemeriksaan ini digunakan untuk memeriksa kualitas sampel acak yang dibangkitkan oleh algoritma MCMC. Dapat dilihat dari *history trace plot* dan *correlation plot*. Jika pemeriksaan ini terpenuhi maka sampel yang dibangkitkan oleh algoritma MCMC telah *equilibrium* [2].

Interpretasi Model

Interpretasi setiap parameter dapat ditentukan dari nilai OR [3]. Dengan menggunakan rumus $\widehat{OR} = \exp(\hat{\beta})$.

Keakuratan Model

Keakurat model dapat diketahui dari kurva ROC-AUC digunakan untuk mengetahui seberapa layak model tersebut dapat menginterpretasikan data [2].

Kanker Leher Rahim

Kanker leher rahim yaitu keganasan yang terjadi pada leher rahim yang diakibatkan infeksi virus *Human Papiloma Virus* (HPV)[5]. Adapun faktor-faktor resiko pemicu kanker leher rahim:

1. Virus HPV: adalah faktor inisiator dari kanker leher rahim yang mengakibatkan gangguan pada sel lahir. Wanita dengan seksual aktif rentan terinfeksi oleh virus ini [5].

2. Usia: usia lebih dari 35 tahun sangat rentan terkena kanker, pada kelompok usia 55-64 tahun memiliki nilai prevalensi terkena kanker tinggi [5].
3. Usi pertama menikah: berhubungan seksual pada usia dini rentan terkena kanker leher rahim. Wanita yang menikah di usia < 20 tahun memiliki resiko lebih tinggi terkena leher rahim [6].
4. Merokok: wanita perokok memiliki resiko 2,5 kali lebih tinggi terkena kanker leher rahim sedangkan wanita perokok pasif memiliki 1,4 kali lebih tinggi terkena kanker daripada wanita yang jauh dari asap rokok [5].
5. Paritas: wanita yang sering melahirkan cenderung lebih rentan terkena kanker leher rahim [5].
6. Penggunaan alat kontrasepsi: penggunaan alat kontrasepsi > 50 tahun memiliki resiko tinggi terkena kanker leher rahim [5].
7. Riwayat keluarga: faktor keturunan membuat beberapa wanita lebih rentan terinfeksi virus HPV [5].
8. Pendidikan: prevalensi terkena kanker meningkat pada tingkat pendidikan Perguruan Tinggi [7].
9. Pekerjaan: wanita yang bekerja lebih beresiko terkena kanker disebabkan terpapar sesuatu di pekerjaan seperti bahan kimia, debu dll [6].
10. Tes papsmear: tes ini digunakan untuk mengetahui perubahan sel pada daerah leher rahim. Setiap wanita diharuskan melakukan tes ini 3 tahun sekali [6].
11. Hubungan seksual: wanita yang berhubungan seksual dengan > 1 pasangan memiliki resiko lebih tinggi terkena kanker leher rahim [5].
12. Siklus menstruasi: siklus menstruasi yang normal berkisar 5-7 hari, ketika wanita mengalami siklus menstruasi yang tidak normal maka harus dilakukan pemeriksaan tes papsmear [8].
13. Menopause: yaitu kondisi periode akhir menstruasi apabila selama 12 bulan tidak mengalami menstruasi [8].
14. Obesitas: angka kematian kanker leher rahim pada wanita diakibatkan oleh obesitas [8].

3. Pembahasan dan Diskusi

Data yang digunakan pada penelitian ini adalah data rekam medis pasien poli kandungan RSUD Sayang Cianjur bulan oktober-desember 2020. Diketahui pasien yang berobat sebanyak 100 orang. Terdapat 65 orang terkena kanker leher rahim dan 35 tidak terkena kanker leher rahim.

Pendugaan Parameter Model

Pengujian hipotesis terhadap parameter dilakukan dengan *credible interval*. Parameter yang signifikan menunjukkan bahwa variabel bebas tersebut berpengaruh terhadap model. Adapun nilai estimasi parameter adalah sebagai berikut:

Tabel 1. Nilai Credible Interval untuk Setiap Parameter Tahap I

Parameter	Credible Interval (CI)		Keterangan
	3%	97%	
Konstanta	-1,185	0,429	
Usia (X_1)	0,112	0,151	Berpengaruh
Usia menikah (X_2)	-0,439	-0,359	Berpengaruh
Paritas (X_3)	-0,423	0,116	Tidak berpengaruh
Jumlah pasangan (X_4)	0,693	2,36	Berpengaruh

Berat badan (X_5)	-0,102	-0,077	Berpengaruh
Menopause (X_6)	-2,815	1,005	Tidak berpengaruh
Tes papsmear (X_7)	5,4	7,23	Berpengaruh
Riwayat kontrasepsi (X_8)	0,226	2,054	Berpengaruh
Riwayat keluarga (X_9)	0,838	25,525	Berpengaruh
Status menikah (X_{10})	-0,208	1,564	Tidak berpengaruh
Siklus menstruasi (X_{11})	-0,426	1,732	Tidak berpengaruh
Merokok (X_{12})	-3,69	0,412	Tidak berpengaruh
Pendidikan SMP ($X_{13(1)}$)	-0,446	3,068	Tidak berpengaruh
Pendidikan SMA ($X_{13(2)}$)	-0,175	1,36	Tidak berpengaruh
Pendidikan Perguruan Tinggi ($X_{13(3)}$)	-0,873	0,852	Tidak berpengaruh
Pekerjaan (X_{14})	0,337	3,693	Berpengaruh

Berdasarkan tabel 1 diketahui bahwa variabel bebas usia, usia menikah, jumlah pasangan, berat badan, tes papsmear, riwayat kontrasepsi, riwayat keluarga dan pekerjaan signifikan pada iterasi 100.000. Dapat dilihat dari nilai *quantile* pada CI antara 3% dan 97% tidak mengandung nilai nol. Jika nilai antara 3% dan 97% memuat nilai nol maka parameter tidak signifikan. Langkah selanjutnya yaitu dilakukan analisis yang sama dengan memasukan variabel bebas yang berpengaruh signifikan terhadap model. Diperoleh sebagai berikut:

Tabel 1. Nilai Credible Interval untuk Setiap Parameter Tahap II

Parameter	Credible Interval (CI)		Keterangan
	3%	97%	
Koefisien	-0,801	0,715	
Usia (X_1)	0,036	0,07	Berpengaruh
Usia menikah (X_2)	-0,316	-0,241	Berpengaruh
Jumlah pasangan (X_3)	1,16	2,606	Berpengaruh
Berat badan (X_4)	-0,086	-0,062	Berpengaruh
Tes papsmear (X_7)	5,049	6,631	Berpengaruh
Riwayat kontrasepsi (X_8)	0,772	2,389	Berpengaruh
Riwayat keluarga (X_9)	-1,194	24,671	Tidak berpengaruh
Pekerjaan (X_{14})	0,299	3,409	Berpengaruh

Berdasarkan tabel 2 diketahui bahwa variabel bebas usia, usia menikah, jumlah pasangan, berat badan, tes papsmear, riwayat kontrasepsi dan pekerjaan signifikan pada iterasi

100.000. Dapat dilihat dari nilai *quantile* pada CI antara 3% dan 97% tidak mengandung nilai nol. Jika nilai antara 3% dan 97% memuat nilai nol maka parameter tidak signifikan. Langkah selanjutnya yaitu dilakukan analisis yang sama dengan memasukan variabel bebas yang berpengaruh signifikan terhadap model. Diperoleh sebagai berikut:

Tabel 3. Nilai Credible Interval untuk setiap Parameter tahap III

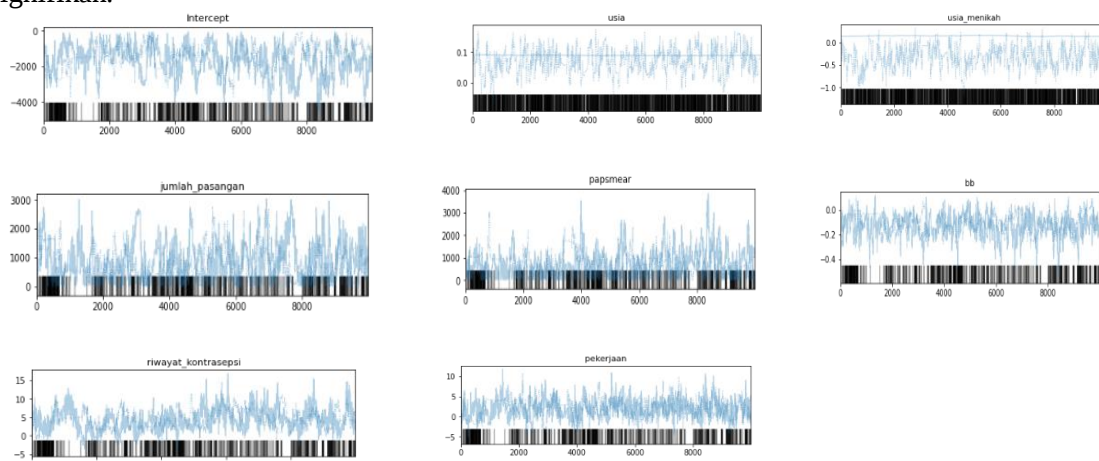
Parameter	Credible Interval (CI)		Keterangan
	3%	97%	
Koefisien	0,82	0,567	
Usia (X_1)	0,034	0,068	Berpengaruh
Usia menikah (X_2)	-0,317	-0,248	Berpengaruh
Jumlah pasangan (X_3)	1,27	2,608	Berpengaruh
Berat badan (X_4)	-0,076	-0,055	Berpengaruh
Tes papsmear (X_7)	5,187	6,625	Berpengaruh
Riwayat kontrasepsi (X_8)	0,013	1,648	Berpengaruh
Pekerjaan (X_{14})	0,407	3,31	Berpengaruh

Berdasarkan tabel 3 diketahui bahwa variabel bebas usia, usia menikah, jumlah pasangan, berat badan, tes papsmear, riwayat kontrasepsi dan pekerjaan signifikan pada iterasi 100.000. Dapat dilihat dari nilai *quantile* pada CI antara 3% dan 97% tidak mengandung nilai nol. Maka diperoleh model regresi logistik sebagai berikut:

$$\text{logit}(\hat{p}) = -0,164 + 0,051X_1 - 0,281X_2 + 1,949X_3 - 0,065X_4 + 5,915X_5 + 0,928X_7 - 1,882X_{14}$$

Kestabilan Nilai Dugaan Parameter

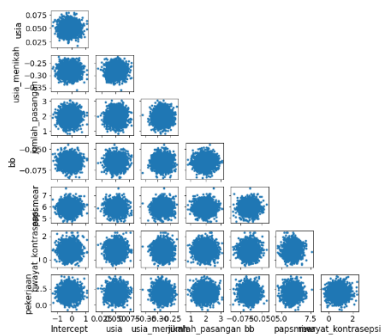
Dilakukan iterasi sebanyak 100000. Berikut *history trace plot* untuk setiap parameter yang signifikan.



Gambar 1. History trace plot

Berdasarkan *history trace plot* yang diperoleh, dapat disimpulkan bahwa setiap parameter telah stabil, selanjutnya akan dilihat kovarians dari pasangan dugaan parameter.

Berikut grafik kovaraians.



Gambar 2. Kovarians Setiap Pasangan Parameter

Berdasarkan gambar di atas, dapat disimpulkan bahwa untuk setiap pasangan parameter tidak ada hubungan antar pasangan parameter dikarenakan sebaran plot acak. Berdasarkan hasil *trace plot* dan grafik kovarians, sampel yang dibangkitkan oleh algoritma MCMC telah memenuhi sifat *equilibrium* yang bermakna parameter yang dihasilkan dalam keadaan setimbang sehingga tidak ada kecenderungan untuk berubah.

Interpretasi Koefisien

Berikut nilai odds ratio untuk setiap koefisien:

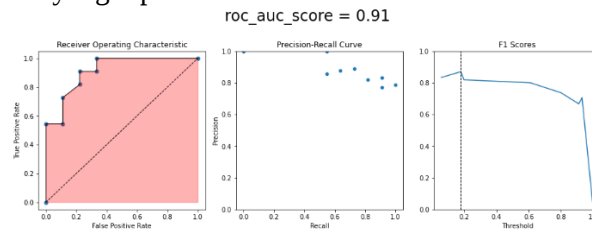
Tabel 4. Nilai Odds Ratio Setiap Parameter

Parameter	Koefisien	$\exp(\beta)$
Konstanta	-0,164	
Usia (X_1)	0,051	1,0523
Usia menikah (X_2)	-0,281	0,7550
Jumlah pasangan (X_3)	1,949	7,0217
Berat badan (X_4)	-0,065	0,9371
Tes papsmear (X_7)	5,915	370,5543
Riwayat kontrasepsi (X_8)	0,928	2,5294
Pekerjaan (X_{14})	1,882	6,5666

Nilai odds ratio untuk usia (X_1) yaitu 1,0523 artinya untuk usia pasien bertambah 1 tahun, maka peluang terkena kanker leher rahim lebih tinggi sebesar 1,0523. Nilai odds ratio usia menikah (X_2) yaitu 0,755 artinya untuk usia menikah pasien bertambah 1 tahun, maka peluang terkena kanker leher rahim lebih rendah sebesar 0,755 kali. Nilai odds ratio jumlah pasangan (X_3) yaitu 7,0217 artinya setiap jumlah pasangan pasien bertambah 1 orang maka peluang terkena kanker leher rahim akan lebih tinggi 7,0217 kali. Nilai odds ratio berat badan (X_4) yaitu 0,9371 artinya setiap berat badan bertambah 1 kg, maka peluang terkena kanker leher rahim akan lebih rendah sebesar 0,9371 kali. Nilai odds ratio tes papsmear (X_7) yaitu 370,5543, artinya wanita yang rutin melakukan tes papsmear akan lebih mudah terdeteksi kanker leher rahim daripada wanita yang tidak rutin melakukan tes papsmear. Nilai odds ratio riwayat kontrasepsi (X_8) yaitu 2,5294 artinya peluang wanita yang mempunyai riwayat penggunaan alat kontrasepsi akan lebih tinggi terkena kanker leher rahim. Nilai odds ratio pekerjaan (X_{14}) yaitu 6,556 artinya peluang wanita yang bekerja akan lebih tinggi terkena kanker leher rahim daripada wanita IRT.

Keakuratan Model

Berikut kurva ROC-AUC yang diperoleh:



Gambar 3. Kurva ROC-AUC

Berdasarkan nilai AUC model tersebut memberikan keakuratan 91% maka dapat disimpulkan bahwa model yang diperoleh sudah layak digunakan untuk mengetahui faktor-faktor yang mempengaruhi kanker leher rahim.

4. Kesimpulan

Dari 14 variabel bebas yang diduga faktor-faktor yang mempengaruhi kanker leher rahim, diperoleh tujuh variabel bebas yang berpengaruh nyata terhadap kanker leher rahim. Faktor-faktor tersebut adalah usia (X_1), usia menikah (X_2), jumlah pasangan (X_3), berat badan (X_4), tes papsmear (X_7), riwayat kontrasepsi (X_8) dan pekerjaan (X_{14}). Berdasarkan hasil uji kestabilan parameter model masing-masing plot menunjukkan model yang diperoleh telah stabil dan berdasarkan uji keakuratan model menunjukkan model yang diperoleh sudah layak digunakan untuk mengetahui faktor-faktor yang mempengaruhi kanker leher rahim. Model logit yang diperoleh adalah:

$$\text{logit}(\hat{p}) = -0,164 + 0,051X_1 - 0,281X_2 + 1,949X_3 - 0,065X_4 + 5,915X_5 + 0,928X_7 - 1,882X_{14}$$

Acknowledge

Penulis mengucapkan terimakasih kepada Bapa Abdul Kudus, S.Si., M.Si., Ph.D yang telah memberikan masukan dan saran sehingga penelitian ini dapat diselesaikan dengan baik.

Daftar Pustaka

- [1] Dinas Kesehatan Provinsi Jawa Barat, *Profil Kesehatan Jawa Barat Tahun 2019*, vol. 53, no. 9. 2019.
- [2] C. Robert and W. M. Boldstad, *Understanding Computational Bayesian Statistics*, vol. 25, no. 2. 2012.
- [3] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression Second Edition*. 2000.
- [4] D. B. R. Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, *Bayesian Data Analysis*, Third. USA: Taylor & Francis Group, 2014.
- [5] Kemenkes RI, *Pedoman Teknis Pengendalian Kanker Payudara dan Kanker Leher Rahim*, no. 1. Jakarta: Kementerian Kesehatan RI, 2013.
- [6] N. R. Dianti and M. A. Isfandiari, "Perbandingan Risiko Ca Serviks Berdasarkan Personal Hygiene Pada Wanita Usia Subur Di Yayasan Kanker Wisnuwardhana Surabaya," *J. PROMKES*, vol. 4, no. 1, p. 82, 2017, doi: 10.20473/jpk.v4.i1.2016.82-91.
- [7] Kementerian Kesehatan Republik Indonesia, *Beban Kanker di Indonesia*. Jakarta: Kementerian Kesehatan RI, 2019.
- [8] J. Spencer, *Cervical Cancer (Deadly Diseases and Epidemics)*. Switzerland: Chelsea House.
- [9] Wistara, R. A., Suliadi, & Kudus, A. (2015). Regresi Logistik pada Data Rare Event. *Prosiding Statistika*.
- [10] Shofwani Sheila Ghazia, Kudus Abdul. (2021). *Penentuan Kriteria Pengunjung dalam Pemilihan Green Hotel di Kota Bandung Menggunakan Metode Discrete Choice Experiment dengan Desain Choice Sets Kombinatorial*. *Jurnal Riset Statistika*, 1(1), 1-9.