
Penerapan Metode *Stepwise* dan *Dominance Analysis* Pada Regresi Logistik Biner (Studi Kasus: Data Hipertensi Di Indonesia)

Muhammad Idman^{1*}, La Podje Talangko¹, Sitti Sahrinan¹
¹Departemen Statistika, Fakultas MIPA,
Universitas Hasanuddin, Makassar, 90245, Indonesia
* Corresponding author, email: muhammadidman@gmail.com

Abstract

Binary logistic regression is a method to describe the relationship between response variable that has two categories and one or more predictor variables. One of methods that can be used to obtain best model in logistic regression is stepwise. Stepwise method is a method that sets p_{entry} and p_{remove} as criteria to build model. Dominance analysis is used in this research to determine the importance rank of predictor variables by comparing the coefficient of determination (R^2) value before and after the predictor variable entered the model. Binary logistic regression can be used to find the relationship between hypertension and the factor risks. This study aims to obtain best model and to obtain the importance rank each predictor variable of binary logistic regression on data of hypertension in Indonesia. The result of this study shows that best model which is obtained is model with predictor variable of Heart Problems, High Cholesterol, Kidney Disease, Imperfect Vision, Breathlessness, and Nausea/ Vomitting. According to the value of R^2 McFadden, predictor variable of High Cholesterol infests first rank in the importance of predictor variable or gives the greatest contributions in explaining variety of Status of Hypertension than other predictor variables.

Keywords: Binary Logistic Regression, Dominance Analysis, Hypertension, p_{entry} , p_{remove} , Stepwise.

Abstrak

Regresi logistik biner merupakan metode untuk mendeskripsikan hubungan antara variabel respon yang memiliki dua kategori dengan satu atau lebih variabel prediktor. Salah satu metode yang digunakan untuk mendapatkan model terbaik pada regresi logistik adalah *stepwise*. Metode *stepwise* adalah metode yang menetapkan p_{entry} dan p_{remove} sebagai kriteria dalam membangun model. *Dominance analysis* digunakan untuk menentukan peringkat kepentingan variabel prediktor dengan membandingkan nilai koefisien determinasi (R^2) sebelum dan sesudah variabel prediktor dimasukkan ke dalam model. Peneliti ini bertujuan mendapatkan model terbaik dan mendapatkan peringkat kepentingan variabel prediktor regresi logistik biner pada data hipertensi di Indonesia. Hasil penelitian menunjukkan bahwa model terbaik yang didapatkan adalah model dengan variabel prediktor Masalah Jantung, Kolesterol Tinggi, Penyakit Ginjal, Pengelihatan Tidak Sempurna, Sesak Nafas, dan Mual/ Muntah. Berdasarkan nilai dari R^2 McFadden, variabel prediktor Kolesterol Tinggi menduduki peringkat pertama dalam kepentingan variabel prediktor atau memberikan kontribusi terbesar dalam menjelaskan keragaman Status Hipertensi dibandingkan variabel prediktor lain.

Kata Kunci: Regresi Logistik Biner, *Dominance Analysis*, Hipertensi, p_{entry} , p_{remove} , *Stepwise*.

1. Pendahuluan

Penyakit Tidak Menular (PTM) adalah jenis penyakit yang tidak dapat ditularkan dari orang ke orang melalui bentuk kontak apa pun. Salah satu penyakit tidak menular yang utama adalah hipertensi [1]. Hipertensi adalah penyakit yang didefinisikan sebagai

peningkatan tekanan darah secara menetap [2]. Seseorang dikatakan mengalami hipertensi jika tekanan darah berada di atas 140/90 mmHg [3].

Salah satu metode yang dapat digunakan untuk menganalisis hubungan antara penyakit hipertensi dan faktor-faktor risikonya adalah metode regresi logistik biner. Analisis regresi logistik biner digunakan untuk menjelaskan hubungan antara variabel respon yang berupa data dikotomik/biner dengan variabel prediktor yang berupa data berskala interval dan atau kategorik. Agar diperoleh hasil analisis yang optimal pada regresi logistik biner maka diperlukan model terbaik. Salah satu metode yang dapat digunakan untuk mendapatkan model terbaik pada regresi logistik adalah *Stepwise* [4]. Metode *Stepwise* adalah metode yang menetapkan p_{entry} dan p_{remove} sebagai kriteria dalam membangun model.

Setelah mendapatkan model terbaik, untuk mendapatkan informasi yang lebih akurat, dapat dilakukan *dominance analysis* pada model terbaik untuk menentukan peringkat kepentingan variabel prediktor. *Dominance analysis* sebagai pendekatan untuk mengetahui kontribusi tiap variabel prediktor sehingga bisa diketahui peringkat kepentingan masing-masing variabel prediktor [5]. Proses penentuan peringkat kepentingan variabel prediktor dilakukan dengan membandingkan nilai R^2 sebelum dan sesudah variabel prediktor dimasukkan ke dalam model [6].

Berdasarkan uraian tersebut, penelitian ini menerapkan metode *stepwise* untuk mendapatkan model terbaik regresi logistik biner dan *dominance analysis* untuk menentukan peringkat tingkat kepentingan tiap variabel prediktor regresi logistik biner pada data hipertensi di Indonesia.

2. Material dan Metode

Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu data hipertensi Indonesia yang diperoleh melalui *Indonesia Family Life Survey* 2014 dengan jumlah observasi 31.396. Variabel yang digunakan dalam penelitian ini terdiri dari satu variabel respon dan sembilan variabel prediktor. Status Hipertensi sebagai variabel respon (y), Masalah Jantung (x_1), Kolesterol Tinggi (x_2), Penyakit Ginjal (x_3), Masalah Psikis (x_4), Pengelihatian Tidak Sempurna (x_5), Merasa Tertekan (x_6), Merasa Lelah (x_7), Sesak Nafas (x_8), dan Mual/ Muntah (x_9).

2.1 Regresi Logistik Biner

Regresi logistik biner merupakan perkembangan dari sebuah interpretasi koefisien regresi logistik dengan situasi dimana variabel respon (y) adalah variabel kualitatif yang mempunyai skala nominal atau dikotomus. Variabel respon dikategorikan sama dengan nol atau satu. Dalam keadaan demikian, variabel respon mengikuti distribusi Bernoulli untuk setiap observasi tunggal [4]. Fungsi probabilitas untuk setiap observasi adalah sebagai berikut:

$$f(y_i) = \pi(x_i)^{y_i}(1 - \pi(x_i))^{1-y_i}; y = 0,1 \quad (1)$$

Menurut Hosmer & Lemeshow (2000), model probabilitas regresi logistik dengan melibatkan beberapa variabel prediktor (x_i) adalah sebagai berikut:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})} \quad (2)$$

2.2 Estimasi Parameter

Penaksir parameter pada regresi logistik dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) yaitu memaksimumkan fungsi *likelihood*. Fungsi *likelihood* diperoleh dengan mengalikan fungsi-fungsi kepadatan peluang dari y_i karena setiap pengamatan diasumsikan independen. Sehingga dapat dituliskan dalam bentuk:

$$L(\beta) = \prod_{i=1}^n (\pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}) \quad (3)$$

Untuk mendapatkan nilai β dengan memaksimumkan nilai fungsi *likelihood* digunakan bentuk logaritma natural dari fungsi *likelihood*, yang kemudian disebut fungsi log *likelihood*. Sehingga dapat dituliskan dalam bentuk:

$$l(\beta) = \sum_{i=1}^n [-\ln(1 + \exp(x_i^t \beta)) + y_i(x_i^t \beta)] \quad (4)$$

Untuk mendapatkan estimasi parameter terhadap β , maka persamaan $l(\beta)$ diturunkan terhadap β kemudian disamakan dengan nol agar diperoleh nilai β yang dapat memaksimumkan $l(\beta)$, sehingga didapatkan persamaan sebagai berikut:

$$\sum_{i=1}^n x_i^t \left(y_i - \frac{\exp(x_i^t \beta)}{1 + \exp(x_i^t \beta)} \right) = 0 \quad (5)$$

Selanjutnya karena hasil turunan pertama fungsi log *likelihood* terhadap masing-masing parameter tidak dapat diselesaikan secara analitik, maka pada penaksiran ini akan dilakukan pendekatan lain untuk mendapatkan nilai taksirannya. Nilai parameter β dari turunan pertama fungsi $l(\beta)$ didapatkan melalui suatu prosedur iteratif yang dilakukan dengan metode iterasi *Iteratively Reweighted Least Square* (IRLS). Adapun persamaan yang didapatkan dengan metode IRLS adalah sebagai berikut:

$$\hat{\beta}_{MLE} = (X^t \widehat{W} X)^{-1} X^t \widehat{W} \widehat{Z} \quad (6)$$

dengan:

$$\widehat{W} = \text{diag}[\hat{\pi}(x_i)(1 - \hat{\pi}(x_i))] \text{ dan } \widehat{Z} = X\hat{\beta}_t + \widehat{W}^{-1}(y - \hat{\pi}(x))$$

2.3 Metode Stepwise

Model regresi logistik terbaik dapat diperoleh dengan metode yang disebut metode *Stepwise*. Metode *Stepwise* merupakan gabungan dari metode *forward* dan *backward*. Metode ini bekerja dengan menetapkan nilai p_{entry} dan p_{remove} sebagai kriteria dalam membangun model. Kriteria inilah yang nantinya akan menentukan variabel prediktor

yang akan dimasukkan ke dalam model dan variabel prediktor yang akan dikeluarkan dari model yang dibandingkan dengan nilai $p - value$ dari uji G. Berikut adalah langkah-langkah metode *Stepwise*:

Tahap 0: Misalkan terdapat p variabel prediktor yang mungkin untuk masuk ke dalam model. tahap 0 dimulai dengan hanya memasukkan konstanta ke dalam model, kemudian mendapatkan nilai log-likelihood l_0 . Pada tahap ini juga, dilakukan pembentukan model regresi logistik univariabel pada setiap variabel prediktor yang ada dan membandingkan nilai log-likelihoodnya masing-masing.

Variabel prediktor yang dianggap penting adalah variabel prediktor yang mempunyai nilai $p - value$ terkecil. Jika variabel prediktor ini dinotasikan dengan x_{e_1} , maka $p_{e_1}^{(0)} = \min(p_j^{(0)})$. Subskrip e_1 menunjukkan variabel prediktor yang menjadi kandidat untuk dimasukkan ke dalam tahap 1. Dengan demikian akan dilanjutkan ke tahap 1 jika $p_{e_1}^{(0)} < p_E$; sebaliknya, akan dihentikan.

Tahap 1: Pada tahap ini dimulai dengan model regresi logistik yang berisi variabel prediktor x_{e_1} . Misalkan $l_{e_1}^{(1)}$ menunjukkan nilai log-likelihood pada model ini. Untuk menentukan apakah $p - 1$ variabel prediktor yang tersisa merupakan variabel prediktor yang penting setelah x_{e_1} berada dalam model. Misalkan variabel prediktor dengan $p - value$ terkecil pada tahap 1 adalah x_{e_2} , dimana $p_{e_2}^{(1)} = \min(p_j^{(1)})$. Jika nilai dari $p_{e_2}^{(1)} < p_E$ maka akan dilanjutkan ke tahap 2; sebaliknya akan dihentikan.

Tahap 2: Pada tahap ini dimulai dengan model regresi logistik yang berisi variabel prediktor x_{e_1} dan x_{e_2} . Ada kemungkinan bahwa setelah x_{e_2} ditambahkan ke model, x_{e_1} menjadi tidak penting lagi. Jadi, pada tahap 2 dilakukan pengecekan dengan metode *backward*.

Untuk memilih variabel prediktor mana yang akan menjadi kandidat untuk dihapus dari model dilakukan dengan memilih nilai $p - value$ terbesar. Misal x_{r_2} adalah variabel prediktor yang menjadi kandidat untuk dihapus dari model maka $p_{r_2}^{(2)} = \max(p_{e_1}^{(2)}, p_{e_2}^{(2)})$. Nilai p_R harus lebih besar dari nilai p_E . Jika $p_{r_2}^{(2)} > p_R$ maka x_{r_2} dihapus dari model, sebaliknya jika $p_{r_2}^{(2)} < p_R$ maka x_{r_2} tetap berada dalam model. Setelah melakukan pengecekan mana variabel yang harus dihapus dari model pada tahap 2 dilanjutkan metode *forward* seperti pada tahap 1.

Tahap 3: Pada tahap ini dilakukan sama seperti pada tahap 2. Membentuk model regresi logistik yang berisi variabel prediktor yang terpilih pada tahap sebelumnya, kemudian dilakukan pengecekan secara *backward* yang diikuti dengan pemilihan variabel prediktor yang tersisa secara *forward*. Proses berlanjut dengan cara seperti ini hingga tidak ada variabel yang memungkinkan dimasukkan atau dikeluarkan dari model [4].

2.4 Dominance Analysis

Dominance analysis merupakan suatu metode untuk dapat mengetahui variabel prediktor yang mendominasi variabel respon dengan membandingkan nilai kontribusi penambahan variabel prediktor yang masuk ke dalam model yang terdapat pada regresi logistik. Ukuran untuk dapat menentukan variabel prediktor yang mendominasi yaitu dengan menghitung besarnya kontribusi penambahan suatu variabel prediktor [6]. Kontribusi penambahan variabel prediktor pada model regresi logistik dapat menggunakan R^2 McFadden (R_M^2)

$$R_M^2 = 1 - \frac{l_p}{l_0} \quad (7)$$

Dominance analysis mendefinisikan nilai penambahan kontribusi variabel prediktor saat dimasukkan ke dalam model. Nilai kontribusi penambahan tersebut didefinisikan sebagai berikut:

$$C_{i(k)} = R_{Y.Xa}^2 - R_{Y.Xb}^2 \quad (8)$$

dengan :

$C_{i(k)}$ = Nilai kontribusi penambahan variabel prediktor ke-i ke dalam model dengan k variabel prediktor.

$R_{Y.Xa}^2$ = R^2 sesudah variabel prediktor ke-i masuk ke model.

$R_{Y.Xb}^2$ = R^2 sebelum variabel prediktor ke-i masuk ke model.

2.5 Hipertensi

Definisi hipertensi atau tekanan darah tinggi adalah peningkatan tekanan darah sistole lebih dari 140 mmHg dan tekanan darah diastole lebih dari 90 mmHg pada dua kali pengukuran dengan selang waktu lima menit dalam keadaan cukup istirahat/tenang. Peningkatan tekanan darah yang berlangsung dalam jangka waktu lama (persisten) dapat menimbulkan kerusakan pada ginjal (gagal ginjal), jantung (penyakit jantung koroner) dan otak (menyebabkan stroke) bila tidak dideteksi secara dini dan mendapat pengobatan yang memadai. Banyak pasien hipertensi dengan tekanan darah tidak terkontrol dan jumlahnya terus meningkat. Oleh karena itu, partisipasi semua pihak, baik dokter dari berbagai bidang peminatan hipertensi, pemerintah, swasta maupun masyarakat diperlukan agar hipertensi dapat dikendalikan [7].

Faktor resiko hipertensi adalah umur, jenis kelamin, riwayat keluarga, genetik (faktor resiko yang tidak dapat diubah/dikontrol), kebiasaan merokok, konsumsi garam, konsumsi lemak jenuh, kebiasaan konsumsi minum-minuman beralkohol, obesitas, kurang aktifitas fisik, stres, penggunaan estrogen [7].

Ada pun klasifikasi hipertensi terbagi menjadi:

1. Berdasarkan penyebab

a. Hipertensi Primer/Hipertensi Esensial

Hipertensi yang penyebabnya tidak diketahui, walaupun dikaitkan dengan kombinasi faktor gaya hidup seperti kurang bergerak dan pola makan. Terjadi pada sekitar 90% penderita hipertensi.

b. Hipertensi Sekunder/Hipertensi Non Esensial

Hipertensi yang diketahui penyebabnya. Pada sekitar 5-10% penderita hipertensi, penyebabnya adalah penyakit ginjal. Pada sekitar 1-2%, penyebabnya adalah kelainan hormonal atau pemakaian obat tertentu.

2. Berdasarkan bentuk Hipertensi

Hipertensi diastole, hipertensi campuran (diastole dan sistole yang meninggi), hipertensi sistole [7].

3. Hasil dan Diskusi

3.1 Pembentukan Model Dengan Metode *Stepwise*

Pembentukan model untuk regresi logistik biner menggunakan metode *Stepwise* (gabungan dari metode *forward* dan *backward*). Metode ini bekerja dengan menetapkan nilai p_{entry} dan p_{remove} sebagai kriteria dalam membangun model. Berikut adalah tahap dalam pembentukan model terbaik:

1) Tahap 0

Hasil dari $p - value$ tiap model pada tahap 0 menunjukkan bahwa model dengan variabel prediktor x_2 memiliki $p - value$ terkecil. Model dengan variabel prediktor x_2 memiliki $p - value = 1.94e-176 < p_{entry} = 0.05$ sehingga variabel prediktor x_2 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

2) Tahap 1

Hasil dari $p - value$ tiap model pada tahap 1 menunjukkan bahwa model dengan variabel prediktor x_2 dan x_1 memiliki $p - value$ terkecil. Model dengan variabel prediktor x_2 dan x_1 memiliki $p - value = 3.35e-38 < p_{entry} = 0.05$ sehingga variabel prediktor x_1 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

3) Tahap 2

Tahap sebelumnya telah didapatkan variabel prediktor yang memenuhi syarat untuk masuk pada model tahap 2 yaitu x_2 dan x_1 . Sebelum melakukan seleksi terhadap variabel prediktor yang akan digunakan pada model di tahap selanjutnya, terlebih dahulu dilakukan pengecekan secara *backward* untuk melihat apakah ada variabel yang layak untuk dikeluarkan dari model berdasarkan $p - value$. Hasil dari $p - value$ tiap model pada *backward elimination* tahap 2 menunjukkan bahwa model tahap 2 dengan variabel prediktor x_1 dihilangkan pada model memiliki $p - value$ terbesar. Model tersebut

memiliki $p - value = 3.35e-38 < p_{remove} = 0.05$ maka variabel prediktor x_1 tidak dikeluarkan dari model.

Selanjutnya pada tahap ini dilakukan pemilihan variabel prediktor yang akan digunakan pada model di tahap selanjutnya. Pemilihan variabel prediktor berdasarkan pada $p - value$ terkecil. Hasil dari $p - value$ tiap model pada tahap 2 model dengan variabel prediktor x_2, x_1 dan x_5 memiliki $p - value$ terkecil. Model dengan variabel prediktor x_2, x_1 dan x_5 memiliki $p - value = 1.01e-23 < p_{entry} = 0.05$ sehingga variabel prediktor x_5 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

4) Tahap 3

Tahap sebelumnya telah didapatkan variabel prediktor yang memenuhi syarat untuk masuk pada model tahap 3 yaitu x_2, x_1 dan x_5 . Sebelum melakukan seleksi terhadap variabel prediktor yang akan digunakan pada model di tahap selanjutnya, terlebih dahulu dilakukan pengecekan secara *backward* untuk melihat apakah ada variabel yang layak untuk dikeluarkan dari model berdasarkan $p - value$. Hasil dari $p - value$ tiap model pada *backward elimination* tahap 3 menunjukkan bahwa model tahap 2 dengan variabel prediktor x_5 dihilangkan pada model memiliki $p - value$ terbesar. Model tersebut memiliki $p - value = 1.01e-23 < p_{remove} = 0.05$ maka variabel prediktor x_5 tidak dikeluarkan dari model.

Selanjutnya pada tahap ini dilakukan pemilihan variabel prediktor yang akan digunakan pada model di tahap selanjutnya. Pemilihan variabel prediktor berdasarkan pada $p - value$ terkecil. Hasil dari $p - value$ tiap model pada tahap 3 menunjukkan bahwa model dengan variabel prediktor x_2, x_1, x_5 dan x_8 memiliki $p - value$ terkecil. Model dengan variabel prediktor x_2, x_1, x_5 dan x_8 memiliki $p - value = 9.81e-10 < p_{entry} = 0.05$ sehingga variabel prediktor x_8 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

5) Tahap 4

Tahap sebelumnya telah didapatkan variabel prediktor yang memenuhi syarat untuk masuk pada model tahap 4 yaitu x_2, x_1, x_5 dan x_8 . Sebelum melakukan seleksi terhadap variabel prediktor yang akan digunakan pada model di tahap selanjutnya, terlebih dahulu dilakukan pengecekan secara *backward* untuk melihat apakah ada variabel yang layak untuk dikeluarkan dari model berdasarkan $p - value$. Hasil dari $p - value$ tiap model pada *backward elimination* tahap 4 menunjukkan bahwa model tahap 4 dengan variabel prediktor x_8 dihilangkan pada model memiliki $p - value$ terbesar. Model tersebut memiliki $p - value = 9.81e-10 < p_{remove} = 0.05$ maka variabel prediktor x_8 tidak dikeluarkan dari model.

Selanjutnya pada tahap ini dilakukan pemilihan variabel prediktor yang akan digunakan pada model di tahap selanjutnya. Pemilihan variabel prediktor berdasarkan pada $p - value$ terkecil. Hasil dari $p - value$ tiap model pada tahap 4 menunjukkan bahwa model dengan variabel prediktor x_2, x_1, x_5, x_8 dan x_3 memiliki $p - value$ terkecil.

Model dengan variabel prediktor x_2, x_1, x_5, x_8 dan x_3 memiliki $p - value = 5.10e-07 < p_{entry} = 0.05$ sehingga variabel prediktor x_3 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

6) Tahap 5

Tahap sebelumnya telah didapatkan variabel prediktor yang memenuhi syarat untuk masuk pada model tahap 5 yaitu x_2, x_1, x_5, x_8 dan x_3 . Sebelum melakukan seleksi terhadap variabel prediktor yang akan digunakan pada model di tahap selanjutnya, terlebih dahulu dilakukan pengecekan secara *backward* untuk melihat apakah ada variabel yang layak untuk dikeluarkan dari model berdasarkan $p - value$. Hasil dari $p - value$ tiap model pada *backward elimination* tahap 5 menunjukkan bahwa model tahap 2 dengan variabel prediktor x_3 dihilangkan pada model memiliki $p - value$ terbesar. Model tersebut memiliki $p - value = 4.53e-05 < p_{remove} = 0.05$ maka variabel prediktor x_3 tidak dikeluarkan dari model.

Selanjutnya pada tahap ini dilakukan pemilihan variabel prediktor yang akan digunakan pada model di tahap selanjutnya. Pemilihan variabel prediktor berdasarkan pada $p - value$ terkecil. Hasil dari $p - value$ tiap model pada tahap 5 menunjukkan bahwa model dengan variabel prediktor x_2, x_1, x_5, x_8, x_3 dan x_9 memiliki $p - value$ terkecil. Model dengan variabel prediktor x_2, x_1, x_5, x_8, x_3 dan x_9 memiliki $p - value = 0.01 < p_{entry} = 0.05$ sehingga variabel prediktor x_9 memenuhi syarat untuk dimasukkan ke dalam model tahap selanjutnya.

7) Tahap 6

Tahap sebelumnya telah didapatkan variabel prediktor yang memenuhi syarat untuk masuk pada model tahap 6 yaitu x_2, x_1, x_5, x_8, x_3 dan x_9 . Sebelum melakukan seleksi terhadap variabel prediktor yang akan digunakan pada model di tahap selanjutnya, terlebih dahulu dilakukan pengecekan secara *backward* untuk melihat apakah ada variabel yang layak untuk dikeluarkan dari model berdasarkan $p - value$. Hasil dari $p - value$ tiap model pada *backward elimination* tahap 6 menunjukkan bahwa model tahap 6 dengan variabel prediktor x_9 dihilangkan pada model memiliki $p - value$ terbesar. Model tersebut memiliki $p - value = 1.29e-02 < p_{remove} = 0.05$ maka variabel prediktor x_9 tidak dikeluarkan dari model.

Selanjutnya pada tahap ini dilakukan pemilihan variabel prediktor yang akan digunakan pada model di tahap selanjutnya. Pemilihan variabel prediktor berdasarkan pada $p - value$ terkecil. Hasil dari $p - value$ tiap model pada tahap 6 bahwa model dengan variabel prediktor $x_2, x_1, x_5, x_8, x_3, x_9$ dan x_4 memiliki $p - value$ terkecil. Model dengan variabel prediktor $x_2, x_1, x_5, x_8, x_3, x_9$ dan x_4 memiliki $p - value = 0.24 > p_{entry} = 0.05$ sehingga variabel prediktor x_4 tidak layak untuk dimasukkan ke dalam model tahap selanjutnya. Karena tidak ada lagi variabel yang memungkinkan untuk dimasukkan ke dalam model, maka proses pembentukan model dengan metode *stepwise* berakhir pada tahap 6.

Model yang didapatkan pada tahap 6 yaitu model dengan variabel prediktor Masalah Jantung (x_1), Kolesterol Tinggi (x_2), Penyakit Ginjal (x_3), Pengelihan Tidak Sempurna (x_5), Sesak Nafas (x_8), dan Mual/ Muntah (x_9). Hasil pendugaan koefisien regresi logistik biner model dengan metode *stepwise* dapat dilihat pada Tabel 1 berikut ini:

Tabel 1. Pendugaan Estimasi Parameter Model Dengan Metode *Stepwise*

Parameter	Koefisien	Std. Error	$ W_j $	P
Konstanta (β_0)	-2.27	0.02	102.88	0.00*)
Masalah Jantung (x_1)				
Ya (β_1)	1.27	0.10	12.69	0.00*)
Kolesterol Tinggi (x_2)				
Ya (β_2)	1.65	0.06	26.86	0.00*)
Penyakit Ginjal (x_3)				
Ya (β_3)	0.65	0.12	5.27	0.00*)
Pengelihan Tidak Sempurna (x_5)				
Ya (β_5)	0.56	0.06	10.19	0.00*)
Sesak Nafas (x_8)				
Ya (β_{13})	0.35	0.06	5.83	0.00*)
Mual/Muntah (x_9)				
Ya (β_{14})	0.12	0.05	2.51	0.01*)

Keterangan : *) berpengaruh signifikan dengan $\alpha = 5\%$

3.2 Dominance Analysis

Penentuan peringkat kepentingan variabel prediktor yang mendominasi variabel Status Hipertensi dalam penelitian ini menggunakan nilai R^2 McFadden. Nilai R^2 McFadden digunakan untuk mengetahui besarnya masing-masing penambahan kontribusi di setiap variabel prediktor pada *dominance analysis*. Variabel prediktor yang digunakan adalah variabel yang terdapat pada model terbaik yaitu Masalah Jantung (x_1), Kolesterol Tinggi (x_2), Penyakit Ginjal (x_3), Pengelihan Tidak Sempurna (x_5), Sesak Nafas (x_8), dan Mual/ Muntah (x_9). Hasil *dominance analysis* terdapat pada Lampiran 20 dan dapat dilihat pada Tabel 2 berikut ini:

Tabel 2. Nilai Rata-Rata Kontribusi Setiap Variabel Prediktor

p	Rata-rata Kontribusi Variabel Prediktor					
	x_1	x_2	x_3	x_5	x_8	x_9
1	0.0100	0.0330	0.0020	0.0070	0.0020	0.0010
2	0.0090	0.0320	0.0002	0.0060	0.0020	0.0010
3	0.0080	0.0300	0.0020	0.0060	0.0020	0.0000
4	0.0070	0.0290	0.0010	0.0050	0.0020	0.0000
5	0.0060	0.0027	0.0010	0.0040	0.0010	0.0000
G_i	0.0080	0.0253	0.0012	0.0056	0.0018	0.0004

Berdasarkan uraian tersebut dapat disimpulkan bahwa x_2 memiliki kepentingan paling dominan dibandingkan dengan variabel lainnya. Dengan kata lain, variabel prediktor Kolesterol Tinggi (x_2) menduduki peringkat pertama dalam kepentingan variabel prediktor disusul Masalah Jantung (x_1), Pengelihan Tidak Sempurna (x_5), Penyakit Ginjal (x_3), Sesak Nafas (x_8), dan Mual/ Muntah (x_9).

4. Kesimpulan

Dari uraian yang diperoleh dari bab sebelumnya, dapat disimpulkan bahwa:

1. Hasil uji Hosmer dan Lemeshow menunjukkan bahwa data hipertensi di Indonesia cocok dimodelkan dengan model regresi logistik biner dengan model terbaik berdasarkan metode *stepwise* adalah sebagai berikut :

$$\pi(x) = \frac{\exp(-2.3 + 1.3x_1(ya) + 1.6x_2(ya) + 0.6x_3(ya) + 0.6x_5(ya) + 0.4x_8(ya) + 0.1x_9(ya))}{1 + \exp(-2.3 + 1.3x_1(ya) + 1.6x_2(ya) + 0.6x_3(ya) + 0.6x_5(ya) + 0.4x_8(ya) + 0.1x_9(ya))}$$

Variabel yang terdapat dalam model terbaik tersebut adalah Masalah Jantung (x_1), Kolesterol Tinggi (x_2), Penyakit Ginjal (x_3), Pengelihan Tidak Sempurna (x_5), Sesak Nafas (x_8), dan Mual/ Muntah (x_9)

2. Hasil *dominance analysis* menunjukkan bahwa Kolesterol Tinggi memberikan pengaruh paling besar terhadap Status Hipertensi berdasarkan nilai R^2 McFadden dibandingkan dengan variabel lainnya. Dengan kata lain, variabel prediktor Kolesterol Tinggi menduduki peringkat pertama dalam peringkat kepentingan variabel prediktor disusul Masalah Jantung, Pengelihan Tidak Sempurna, Penyakit Ginjal, Sesak Nafas, dan Mual/ Muntah. Adapun variabel yang tidak mempunyai pengaruh signifikan terhadap Status Hipertensi berdasarkan uji Wald yaitu Masalah Psikis, Merasa Tertekan dan Merasa Lelah.

Daftar Pustaka

- [1] Kementerian Kesehatan Republik Indonesia, *Rencana Strategis Kementerian Kesehatan Tahun 2015-2019*, Jakarta: Kementerian Kesehatan RI, 2015.
- [2] J. T. Dipiro, R. L. Talbert, G. C. Yee, G. R. Matzke, B. G. Wells and L. M. Posey, *Pharmacotherapy : A Pathophysiologic Approach*, United States: The McGraw-Hill Companies, Inc, 2011.
- [3] Joint National Committee, *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure*, *National Institutes of Health*, 2004.
- [4] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression Second Edition*, United States: John Wiley & Sons, Inc, 2000.
- [5] R. B. Darlington and A. F. Hayes, *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*, New York: Guilford Publications, 2017.
- [6] R. Azen and N. Traxel, Using Dominance Analysis to Determine Predictor Importance in Logistic Regression, *Journal of Educational and Behavioral Statistics*, pp. 319-347, 2009.
- [7] Pusat Data dan Informasi Kementerian Kesehatan RI, *Hipertensi*, Indonesia: Kementerian Kesehatan RI, 2014.