PEMODELAN REGRESI LOGISTIK BINER TERHADAP PENERIMAAN PEGAWAI DI PT XYZ JAKARTA

Ali Ilham Sofiyat[1], Awaluddin Tjalla[2], Mahdiyah[3]
[1]Program Studi Matematika FST – Universitas Islam As-Syafi'iyah
[2,3] Program Doktor PEP Universitas Negeri Jakarta
Email : alisofiyat.fst@uia.ac.id

ABSTRAK

Penerimaan pegawai merupakan pola seleksi yang dilaksanakan PT XYZ untuk menerima pegawai yang ditempatkan di berbagai cabang perusahaan, melalui ujian tertulis. Dari banyaknya peminat, ada beberapa yang dinyatakan diterima dan sisanya dinyatakan tidak diterima. Oleh karena itu, ingin diketahui karakteristik serta pemodelan menggunakan metode regresi logistik biner dengan variabel respon adalah status penerimaan dimana bernilai satu yaitu diterima dan bernilai nol apabila gagal diterima. Berdasarkan hasil analisis diketahui bahwa jumlah pelamar yang dinyatakan diterima adalah sebesar 50% peserta, sedangkan sisanya sebesar 50% pelamar dinyatakan tidak diterima. Pada pemodelan regresi logistik biner diperoleh variabel yang berpengaruh yaitu *education* (lama pendidikan terakhir) dan *experience* (pengalaman kerja).

Kata Kunci—Regresi Logistik Biner, keputusan diterima

ABSTRACT

Recruitment is a selection pattern implemented by PT XYZ to accept employees placed in various company branches, through a written exam. From the many applicants, there were some who were accepted and the rest were not accepted. Therefore, we want to know the characteristics and modeling using the binary logistic regression method with the response variable is the status of acceptance which has a value of one is accepted and a value of zero if it fails to be accepted. Based on the results of the analysis it is known that the number of applicants who were declared accepted was 50% of the participants, while the remaining 50% of the applicants were declared not accepted. In binary logistic regression modeling, the influential variables are education (last length of education) and experience (work experience).

Keywords—*Binary Logistic Regression*, decision accepted

PENDAHULUAN

Dalam sebuah penelitian biasanya kita memodelkan hubungan antar 2 variabel, yaiitu variabel X (independent) dan Y (dependent). Metode yang biasa dipakai dalam penelitian seperti ini adalah regresi linier, baik sederhana maupun berganda. Namun, adakalanya regresi linier dengan metode OLS (*Ordinary Least Square*) yang dipakai tidak sesuai untuk digunakan.

Regresi linier yang sering digunakan kadang terjadi pelanggaran asumsi *Gauss-Markov*. Misalnya pada kasus dimana variabel dependent (Y) bertipe data nominal, sedangkan variabel bebas/prediktornya (X) bertipe data interval atau rasio. Misalkan ingin diketahui apakah siswa sudah ”Lulus” berdasarkan jenis kelamin, fakultas yang dipilih dan nilai UAN. Dalam kasus ini hanya ada 2 kemungkinan respon siswa, yaitu siswa “Lulus” dan “Tidak Lulus” di PTN.

Dari contoh kasus di atas, dapat diketahui bahwa tipe data variabel respon (Y) adalah ***nominal***, yaitu kategorisasi siswa Lulus atau Tidak Lulus (misal Lulus symbol angka 1 dan Tidak Lulus symbol angka 0), sedangkan tipe data untuk variabel bebas (X) setidak-tidaknya interval (skala likert). Bila metode regresi linier biasa diterapkan pada kasus semacam ini, menurut Kutner, dkk. (2004), akan terdapat 2 pelanggaran asumsi *Gauss-Markov*, yaitu pelanggaran terhadap batasan dari nilai dugaan (*fitted value*) dari variabel respon (Y), yaitu: 1). *Error* dari model regresi yang didapat tidak menyebar normal dan 2). Ragam (*variance*) dari error tidak homogen (terjadi heteroskedastisitas pada ragam error). Sedangkan, pelanggaran bagi batasan nilai dugaan Y (*fitted value*) adalah bahwa nilai dugaan yang dihasilkan dari model regresi linier biasa melebihi rentang antara 0 s.d. 1. Hal ini jelas tidak masuk akal, karena batasan nilai pada variabel Y (dalam kasus ini adalah Lulus =1 dan Tidak Lulus 0).

Untuk mengatasi masalah ini, diperkenalkan metode Regresi Logistik (*Logistic Regression)*. Regresi logistik (kadang disebut model logistik atau model logit), dalam statistika digunakan untuk prediksi probabilitas kejadian suatu peristiwa dengan mencocokkan data pada fungsi logit kurva logistic.

Regresi logistik adalah sebuah pendekatan untuk membuat model prediksi seperti halnya regresi linear atau yang biasa disebut dengan istilah *Ordinary Least Squares* (OLS) regression. Perbedaannya adalah pada regresi logistik, peneliti memprediksi variabel terikat yang berskala dikotomi. Skala dikotomi yang dimaksud adalah skala data nominal dengan dua kategori, misalnya: Ya dan Tidak, Baik dan Buruk atau Tinggi dan Rendah. Apabila pada OLS mewajibkan syarat atau asumsi bahwa error varians (residual) terdistribusi secara normal. Sebaliknya, pada regresi logistik tidak dibutuhkan asumsi tersebut sebab pada regresi logistik mengikuti distribusi logistik.

Asumsi yang harus dipenuhi dalam regresi logistik antara lain:

1) Regresi logistik tidak membutuhkan hubungan linier antara variabel independen dengan variabel dependen.
2) Variabel independen tidak memerlukan asumsi *multivariate normality*.
3) Asumsi homoskedastisitas tidak diperlukan
4) Variabel bebas tidak perlu diubah ke dalam bentuk metrik (interval atau skala ratio).
5) Variabel dependen harus bersifat dikotomi (2 kategori, misal: tinggi dan rendah atau baik dan buruk)
6) Variabel independen tidak harus memiliki keragaman yang sama antar kelompok variabel
7) Kategori dalam variabel independen harus terpisah satu sama lain atau bersifat eksklusif
8) Sampel yang diperlukan dalam jumlah relatif besar, minimum dibutuhkan hingga 50 sampel data untuk sebuah variabel prediktor (independen).
9) Regresi logistik dapat menyeleksi hubungan karena menggunakan pendekatan non linier log transformasi untuk memprediksi *odds ratio*. Odd dalam regresi logistik sering dinyatakan sebagai probabilitas.

Model persamaan aljabar layaknya OLS yang biasa kita gunakan adalah berikut:
$Y = B_0 + B_1X + e$. Dimana e adalah error varians atau residual.

Dengan regresi logistik, tidak menggunakan interpretasi yang sama seperti halnya persamaan regresi OLS. Model Persamaan yang terbentuk berbeda dengan persamaan OLS. Sebagaimana metode regresi biasa, Regresi Logistik dapat dibedakan menjadi 2, yaitu:

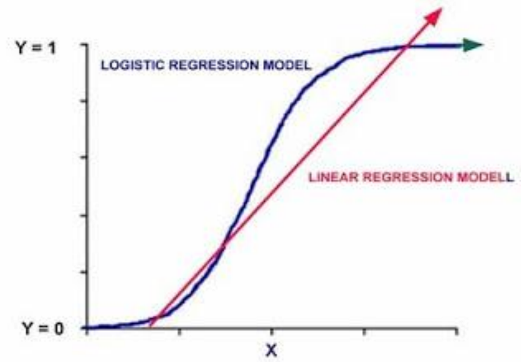
1) *Binary Logistic Regression* (Regresi Logistik Biner). Regresi Logistik biner digunakan ketika hanya ada 2 kemungkinan variabel respon (Y), misal membeli dan tidak membeli.
2) *Multinomial Logistic Regression* (Regresi Logistik Multinomial). Regresi Logistik Multinomial digunakan ketika pada variabel respon (Y) terdapat lebih dari 2 kategorisasi.

DASAR TEORI REGRESI LOGISTIK

Regresi Logistik Biner (*Binary Logistic Regression)* adalah metode statistika yang mempelajari tentang pola hubungan secara matematis antara satu variabel respon (y) yang bersifat nominal atau ordinal dengan satu atau lebih variabel prediktor (x). Perbedaan yang mendasar dengan model regresi linier yaitu pada variabel responnya. Variabel respon pada regresi logistik merupakan variabel biner atau dikotomus. Variabel prediktor dapat berupa variabel polikotomus (kategorik maupun interval). Sedangkan untuk regresi linier, variabel responnya minimal berskala interval. Perbedaan lainnya terlihat pada pemilihan model parametrik dan asumsi-asumsi yang mendasari kedua model. Walaupun demikian, prinsip-prinsip pendugaan parameter yang digunakan dalam analisis model regresi logistik sama dengan analisis model regresi linier (Hosmer and Lemeshow, 1989). Menurut jenis skala dan variabel respon yang digunakan regresi logistik dibagi menjadi 3 macam, yaitu regresi logistik biner, multinomial dan ordinal.

Pendugaan koefisien model regresi logistik tidak dapat dilakukan dengan metode OLS halnya regresi linear karena pelanggaran asumsi kehomogenan varians. Casella and Berger (2002) mengatakan bahwa metode estimasi yang biasanya dipakai adalah metode *Maximum Likelihood*, yang merupakan salah satu alternatif untuk memaksimalkan peluang pengklasifikasian obyek yang diamati menjadi kategori yang sesuai kemudian mengubahnya menjadi koefisien regresi yang sederhana. Metode ini mengasumsikan bahwa nilai ε mengikuti distribusi binomial.

Regresi logistik biner telah banyak digunakan secara luas sebagai salah satu alat analisis pemodelan ketika variabel responnya bersifat biner, yang merujuk pada penggunaan dua buah bilangan 0 dan 1 untuk menggantikan dua kategori pada variabel respon. Contoh variabel respon yang dimaksud adalah kesuksesan (sukses–gagal), kesetujuan (setuju–tidak setuju), keinginan membeli (ya–tidak), terpilih atau tidak terpilih, dan masih banyak lagi.



Gambar 1. Kurva Regresi Logistik dan Regresi Linier

Analisis regresi logistik biner adalah suatu regresi logistik antara variabel respon (y) dan variabel prediktor (x) dimana variabel y menghasilkan 2 kategori yaitu 0 dan 1 (Hosmer dan Lemeshow, 1989). Sehingga variabel y mengikuti distribusi Bernoulli dengan fungsi probabilitasnya sebagai berikut:

$$f(y) = \pi^{y}(1-\pi)^{1-y} \qquad y = 0, 1$$

Dimana y = 0 maka f(y) = $1 - \pi$ dan jika y = 1 maka f(y) = π

Selanjutnya fungsi regresi logistiknya dapat dituliskan:

$$f(z) = \frac{1}{1+e^{-z}} \text{ atau } f(z) = \frac{e^{z}}{1+e^{z}} \qquad (1)$$

Dimana $Z_i = \beta_0 + \beta_1.X_1 + \ldots + \beta_p.X_p$

Nilai e adalah bilangan dasar logaritma natural (ln) yang diperkirakan sebesar 2,71828128 ≈ 2,72.

Jika nilai z antara -□ dan □ maka *f(z)* terletak antara 0 dan 1 untuk setiap nilai z yang diberikan. Hal tersebut menunjukkan bahwa model logistik sebenarnya menggambarkan peluang/probabilitas atau resiko darisuatu obyek. Model regresi logistiknya adalah sebagai berikut:

$$\pi(x) = \frac{e^{\beta_0+\beta_1.X_1+\ldots+\beta_p X_p}}{1+e^{\beta_0+\beta_1.X_1+\ldots+\beta_p X_p}} \tag{2}$$

Atau

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1.x_1 + \ldots + \beta_p.x_p)}{1+\exp(\beta_0 + \beta_1.x_1 + \ldots + \beta_p.x_p)}$$

Dimana p = banyaknya variable predictor.

Bila model persamaan di atas ditransformasi dengan tranformasi logit, maka didapatkan bentuk logit seperti berikut:

$$g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1.x_1 + \ldots + \beta_p.x_p \tag{3}$$

1. Estimasi Parameter

Estimasi parameter dalam regresi logistic dapat dilakukan dengan metode *Maximum Likelihood*, namun apabila metode ini tiak menghasilkan solusi yang *close form*, maka dapat dilanjutkan dengan menggunakan iterasi *Newton Raphson*, hingga menghasilkan solusi yang konvergen.

Metode *Maximum Likelihood*

Pada suatu model dengan respon biner atau dikotomi (bernilai 0 dan 1) dimana antar pengamatan diasumsikan saling bebas, maka penduga parameter β dapat diperoleh dengan metode *Maximum Likelihood Estimation (MLE)* dimana dengan metode ini parameter diestimasi dengan memaksimumkan fungsi turunan pertama. Estimasi varian dan kovarian diperoleh dari turunan kedua fungsi log likelihood.

Jika x_i dan y_i meruapakan pasangan variable bebas dan terikat pada pengamatan ke-i dan diasumsikan bahwa setiap pasangan pengamatan saling independen dengan pasangan pengamatan lainnya, i = 1, 2, 3, …, n maka fungsi peluang (probabilitas) untuk setiap pasangan adalah sebagai berikut:

$$f(\beta, x_i) = \pi(x_i)^{y_i}(1-\pi(x_i))^{y_i} \qquad y_i = 0, 1 \tag{4}$$

Dengan $(x_i) = \dfrac{e^{\left(\sum_{j=0}^{p}\beta_j x_{ij}\right)}}{1+e^{\left(\sum_{j=0}^{p}\beta_j x_{ij}\right)}}$

Dimana ketika j = 0 maka nilai $x_{ij} = x_{i0} = 1$.

Setiap pasangan pengamatan diasumsikan saling bebas (independen) sehingga fungsi *Likelihood* merupakan gabungan dari fungsi distribusi masing-masing pasangan yaitu sebagai berikut:

$$l(\beta) = \prod_{i=1}^{n} f(\beta, x_i)$$

$$= \prod_{i=1}^{n} \pi(x_i)^{y_i} \left(1-\pi(x_i)\right)^{1-y_i}$$

$$= e\left[\sum_{j=0}^{p}\left(\sum_{i=1}^{n} y_i x_{ij}\right)\beta_j\right]\left[\prod_{i=1}^{n}\left(\frac{1}{1+e^{\sum_{j=0}^{p}\beta_j x_{ij}}}\right)\right]$$

Fungsi *likelihood* tersebut lebih mudah dimaksimumkan dalam bentuk log *l(β)* yang disebut juga dengan log *likelihood* (L(β)). Bentuk ini dapat didefinisikan sebagai berikut:

$$L(\beta) = \log l(\beta)$$

$$= \sum_{j=0}^{p}\left(\sum_{i=1}^{n} y_i x_{ij}\right)\beta_j - \sum_{i=1}^{n} \log\left(1 + e^{\sum_{j=0}^{p}\beta_j x_{ij}}\right)$$

Untuk mendapatkan nilai β dari (L(β)) yang maksimum maka dilakukan penurunan (diferensial) terhadap β dan hasilnya disamakan dengan 0.

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^{n} y_i x_{ij} - \sum_{i=1}^{n} x_{ij}\left[\frac{e^{\sum_{j=0}^{p}\beta_j x_{ij}}}{1+e^{\sum_{j=0}^{p}\beta_j x_{ij}}}\right] = 0$$

$$\sum_{i=1}^{n} y_i x_{ij} - \sum_{n=1}^{n} x_{ij}\hat{\pi}(x_i) = 0\ ;\ j = 0, 1, \ldots, p$$

Metode untuk mengestimasi varian dan kovarian dari estimasi koefisien parameter dikembangkan dengan mengikuti teori *Maximum Likelihood Estimation* (MLE) yang menyatakan bahwa estimasi varian dan kovarian diperoleh dari turunan kedua fungsi *likelihood* dapat diperoleh hasil sebagai berikut.

$$\frac{\partial L(\beta)}{\partial \beta_j \beta_u} = \sum_{i=1}^{n} x_{ij} x_{iu} \pi(x_i)(1-\pi(x_i))$$

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_u} = -\sum_{i=1}^{n} x_{ij} x_{iu} \pi(x_i)(1-\pi(x_i))$$

Apabila u = j, maka estimasi varians dapat ditulis

$$\frac{\partial^2 L(\beta)}{\partial (\beta_j)^2} = -\sum_{i=1}^{n} {x_{ij}}^2 \pi_i (1-\pi_i)\ ;\ j, u = 0, 1, \ldots, p$$

Metode *Newton Raphson*

Untuk memperoleh dugaan maksimum bagi parameter β karena pada persamaan likelihood didapatkan $\pi(x)$ yang non linier terhadap β maka digunakan metode *Newton Raphson* melalui iterasi $\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{(-1)} q^{(t)}$, dimana t = 0, 1, …sampai konvergen, dengan $q^T = \left[\frac{\partial L(\beta)}{\partial \beta_0}, \frac{\partial L(\beta)}{\partial \beta_1}, \ldots, \frac{\partial L(\beta)}{\partial \beta_k}\right]$ dan H merupakan matriks Hessian.

$$H = \begin{bmatrix} \frac{\partial^2 L(\beta)}{\partial {\beta_j}^2} & \frac{\partial^2 L(\beta)}{\partial \beta_j \beta_u} \\ \frac{\partial^2 L(\beta)}{\partial \beta_j \beta_u} & \frac{\partial^2 L(\beta)}{\partial {\beta_u}^2} \end{bmatrix}$$

Langkah-langkah iterasi *Newton Raphson* adalah sebagai berikut.

- Gunakan nilai dugaan awal $\beta^{(0)}$ dan dimasukan pada Persamaan (3) untuk mendapatkan $\pi^{(0)}$. Kemudian masukkan dalam Persamaan (10) dan (11).
- Selanjutnya untuk t > 0 digunakan nilai $\hat{\beta}^{(1)} = \beta^{(0)} - [H^{(0)}]^{-1} q^{(0)}$, nilai $\hat{\beta}$ digunakan untuk mencari $\pi^{(1)}$sehingga mendapatkan $q^{(1)}$ dan $H^{(1)}$ untuk memperoleh $\hat{\beta}^{(2)}$ sampai konvergen.

2. Pengujian Estimasi Parameter

Setelah parameter hasil estimasi diperoleh, maka dilakukan pengujian keberartian terhadap koefisien β secara univariat terhadap variabel respon yaitu dengan membandingkan parameter hasil maksimum *likelihood,* dugaan β dengan standar *error* parameter tersebut. Pengujian yang dilakukan adalah sebagai berikut.

a. Uji Serentak

Uji serentak dilakukan untuk mengetahui signifikansi parameter β secara keseluruhan atau serentak. Pengujian yang dilakukan sebagai berikut.

Hipotesis :

H_0 : $\beta_1 = \beta_2 = \ldots = \beta_j = 0$

H_1 : Paling tidak ada satu $\beta_j \neq 0$

Statistik uji (*Likelihood Ratio Test*) :

$$G = -2ln \frac{\left(\frac{n_1}{n}\right)^{n_1}\left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^{n} \hat{\pi}_i^{\,y_i}(1-\hat{\pi}_i)^{(1-y_i)}}$$

dimana : $n_1 = \sum_{i=1}^{n} y_i$; $n_0 = \sum_{i=1}^{n}(1-y_i)$; $n = n_1 + n_0$

Daerah Penolakan :

Tolak H_0 apabila nilai $G > \chi^2_{(v,\ \alpha)}$ atau *P-value* < α, dimana *v* adalah derajat bebas (banyak variabel prediktor yang ada di dalam model tanpa β_0).

b. Uji Individu

Uji individu ini dilakukan untuk mengetahui signifikansi parameter β secara individu (Hosmer dan Lemeshow, 1989).

Hipotesis :

H_0 : $\beta_j = 0$

H_1 : $\beta_j \neq 0$; $j = 1, 2, \ldots, p$

Statistik uji (Uji *Wald*) :

$$W^2 = \frac{\hat{\beta}_j^{\,2}}{SE(\hat{\beta}_j)^2}$$

Daerah Penolakan :

Tolak H_0 apabila $W_i^2 > \chi^2_{(v,\ \alpha)}$ atau *P-value* < α dengan *v* adalah derajat bebas banyaknya prediktor.

3. Pengujian Estimasi Parameter

Uji kesesuaian model berguna untuk mengetahui apakah model tanpa variabel-variabel yang tidak signifikan adalah model terbaik. Terdapat beberapa statistik uji yang dapat digunakan antara lain.

a. –2 log likelihood

$$G^2 = 2\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij} \log\left(\frac{x_{ij}}{m_{ij}}\right)$$

dimana : x_{ij} = nilai pengamatan

m_{ij} = frekuensi harapan

b. Goodness of fit

$$\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{(x_{ij}-m_{ij})^2}{m_{ij}}$$

METODE PENELITIAN

1. Sumber Data

Sumber data yang digunakan dalam penelitian ini merupakan data sekunder yaitu data biodata pelamar sebanyak 40 orang pada Seleksi Penerimaan Pegawai PT XYZ tahun 2017 yang diperoleh dari Kabag Kepegawaian PT XYZ Jakarta.

2. Variabel Penelitian

Variable yang diteliti adalah variable Y (keputusan penerimaan) (1 : diterimadan 0 : tidak diterima), X_1 = lama pendidikan terakhir (tahun), X_2 = lama pengalaman kerja (tahun) dan X_3 = Jenis kelamin pelamar (1 = Laki-laki dan 0 = perempuan).

HASIL ANALISIS DAN PEMBAHASAN

1. Statistika Deskriptif

Statistika deskriptif pelamar yang ikut seleksi Seleksi Penerimaan Pegawai PT XYZ tahun 2017 maka dianalisis dengan menggunakan statistika deskriptif dapat disajikan pada table 1 berikut:

Tabel 1. Lama Pendidikan Terakhir (Tahun) Pelamar

Variabel	Frekuensi	%
Lama Pendidikan Terakhir		
4 Tahun	16	40
6 Tahun	15	37,5
8 Tahun	9	22,5
Pengalaman Kerja		
0 – 3 Tahun	21	52,5
4 – 7 Tahun	11	27,5
Di atas 7 Tahun	8	20
Jenis Kelamin		
Perempuan	17	42,5
Laki-Laki	23	57,5
Penerimaan Kerja		
Diterima	20	50
Tidak Diterima	20	50

Dari hasil statistic deskriptif di atas, diketahui bahwa lama pendidikan terakhir pelamar selama 4 tahun sebanyak 40%, selama 6 tahun sebanyak 37,5% dan sisanya 22,5% selama 8 tahun. Pengalaman kerja pelamar 0 s/d 3 tahun sebanyak 52,5%, 4 s/d 7 tahun sebanyak 27,5% dan di atas 7 tahun sebanyak 20%. Selanjutnya pelamar 42,5%, sedangkan sebanyak 57,5%

adalah pelamar laki-laki. Terakhir jumlah pelamar yang diterima dan tidak diterima masing-masing sebanyak 50%.

2. Pemodelan Regresi Logistik Biner terhadap Pelamar Kerja di PT. XYZ

Pemodelan regresi logistik biner terhadap pelamar di PT XYZ tahun 2017 dilakukan dengan menggunakan variabel respon yaitu keputusan penerimaan, dimana bernilai satu apabila pelamar dinyatakan diterima dan bernilai 0 apabila pelamar dinyatakan tidak diterima.

a. Persamaan regresi logistic awal yang melibatkan Y dengan X_1, X_2 dan X_3

Tabel 2 *Omnibus Tests of Model Coefficients*

		Chi-square	df	Sig.
Step 1	Step	14.477	3	.002
	Block	14.477	3	.002
	Model	14.477	3	.002

Dari hasil di atas diperoleh nilai p = 0,002, dengan demikian dapat disimpulkan bahwa minimal ada satu variabel independen X yang signifikan mempengaruhi variabel Y.

Selain menggunakan nilai pada tabel 2 uji serentak juga dapat dilakukan dengan menggunakan nilai Hosmer-Lemeshow test seperti berikut:

Tabel 3. *Hosmer and Lemeshow Test*

Step	Chi-square	df	Sig.
1	4.004	8	.857

Dari hasil di atas diperoleh nilai p = 0,857, dengan demikian dapat disimpulkan bahwa model regresi logistik yang digunakan telah cukup mampu menjelaskan data (sesuai).

b. Persamaan regresi logistic secara parsial yang melibatkan Y dengan X_1, X_2 dan X_3

Tabel 4 *Variables in The Equation Block 1*

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Education	.673	.291	5.347	1	.021	1.960
	Experience	.312	.122	6.488	1	.011	1.365
	Sex(1)	-1.038	.847	1.502	1	.220	.354
	Constant	-4.662	1.797	6.727	1	.009	.009

a. Variable(s) entered on step 1: Education, Experience, Sex.

Dari tabel 4 di atas diperoleh hasil uji hipotesis sebagai berikut:

1) Variabel *education* (lama pendidikan terakhir) pelamar signifikan mempengaruhi Y (keputusan penerimaan pegawai) dengan nilai p = 0,021.
2) Variabel *experience* (pengalaman kerja) pelamar signifikan mempengaruhi Y (keputusan penerimaan pegawai) dengan nilai p = 0,011
3) Variabel *sex* (jenis kelamin) pelamar tidak signifikan mempengaruhi Y (keputusan penerimaan pegawai) dengan nilai p = 0,220.

Berdasarkan hasil di atas, variable jenis kelamin (X_3) dikeluarkan dari model untuk mendapatkan model regresi logistik yang baru.

c. Persamaan regresi logistic yang melibatkan Y dengan X_1 dan X_2

Uji serentak persamaan regresi logistic setelah variable s*ex* (jenis kelamin) dikeluarkan sebagai berikut:

Tabel 5. *Omnibus Tests of Model Coefficients* tanpa Variabel *Sex*

		Chi-square	df	Sig.
Step 1	Step	12.863	2	.002
	Block	12.863	2	.002
	Model	12.863	2	.002

Dari hasil di atas diperoleh nilai p = 0,002, dengan demikian dapat disimpulkan bahwa minimal ada satu variabel independen X yang signifikan mempengaruhi variabel Y. Selain menggunakan nilai pada tabel 5 uji serentak juga dapat dilakukan dengan menggunakan nilai Hosmer-Lemeshow test seperti berikut:

Tabel 6. *Hosmer and Lemeshow Test* tanpa Variabel *Sex*

Step	Chi-square	df	Sig.
1	9.400	7	.225

Dari hasil di atas diperoleh nilai p = 0,225, dengan demikian dapat disimpulkan bahwa model regresi logistik yang digunakan telah cukup mampu menjelaskan data (sesuai).

d. Persamaan regresi logistic secara parsial yang melibatkan Y dengan X_1 dan X_2

Tabel 7. *Variables in The Equation Block 1* tanpa Variabel *Sex*

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Education	.580	.265	4.780	1	.029	1.786
	Experience	.293	.121	5.908	1	.015	1.340
	Constant	-4.467	1.719	6.750	1	.009	.011

a. Variable(s) entered on step 1: Education, Experience.

Dari tabel 7 di atas diperoleh hasil uji hipotesis sebagai berikut:

1) Variabel *education* (lama pendidikan terakhir) pelamar signifikan mempengaruhi Y (keputusan penerimaan pegawai) dengan nilai p = 0,029.
2) Variabel *experience* (pengalaman kerja) pelamar signifikan mempengaruhi Y (keputusan penerimaan pegawai) dengan nilai p = 0,015

Setelah dilakukan uji serentak dan parsial didaptkan model regresi logistik (logit) sebagai berikut:

$$Logit(p) = \ln\left(\frac{p}{1-p}\right) = -4{,}467 + 0{,}58.Education + 0{,}293.Experience$$

Selanjutnya juga diperoleh nilai *Cox-Snell* R^2 dan *Nagelkerke R* memiliki analogi sama dengan nilai *R-square* pada regresi linier. Berdasarkan nilai *Cox-Snell* R^2 dapat dinyatakan bahwa sebanyak 36,7% keragaman dapat dijelaskan oleh model, sedangkan sisanya (63,3%) dijelaskan oleh faktor (variabel) lain diluar model penelitian.

Berdasarkan tabel 7 dapat diketahui nilai *odds ratio* untuk seluruh variabel prediktor. Pada variabel *education* (lama pendidikan terakhir) diperoleh nilai *odds ratio* sebesar 1,786. Hal ini berarti peluang diterima untuk pelamar yang lama pendidikannya lebih tinggi 1,786 kali lebih besar dibandingkan pelamar yang lama pendidikannya rendah. Nilai *odds ratio* untuk variabel nilai *experience* (pengalaman kerja) sebesar 1,340. Hal ini berarti peluang diterima untuk pelamar yang pengalamannya lebih lama 1,340 kali lebih besar dibandingkan pelamar yang lama pengalamannya baru.

PEMBAHASAN

Peluang pelamar mengalami keberhasilan diterima bekerja di PT XYZ Jaya setelah mencoba memasukkan lamaran pekerjaan adalah 0,7639 (76,39%) jika lama pendidikan terakhirnya lebih lama (lebih tinggi) dan pengalaman kerjanya lebih lama. Berdasarkan hasil penelitian yang telah dilakukan maka saran yang dapat diberikan oleh peneliti adalah perlu adanya persiapan yang matang untuk menghadapi penerimaan masuk kerja dikarenakan persaingan yang semakin ketat dan jumlah penerimaan yang sedikit. Metode yang digunakan dalam penelitian ini sudah relative sesuai dikarenakan proporsi antara diterima dan tidak diterima adalah seimbang.

DAFTAR PUSTAKA

Agresti, Alan (2007). *An Introduction to Categorical Data Analysis, 2nd edition*. New York:Wiley Interscience, John Wiley & Sons. INC.

Cucchiara, Andrew (1992) Applied Logistic Regression, *Technometrics, 34*:3, 358-359, DOI: 10.1080/00401706.1992.10485291

Hosmer D.W dan Stanley Lemeshow (2002), *Applied Logistic Regression*, *2nd edition*. New York: John Wiley & Sons. INC.

Gazpersz, V, (1995). *Teknik Analisis Dalam Penelitian Percobaan Jilid 2*. Bandung: Tarsito

Joanne Peng, Chao-Ying, Kuk Lida Lee dan Gary M. Ingersoll, "An Introduction to Logistic Regression Analysis and Reporting", *The Journal of Educational Research*, September 2002.

Myers, R.H (2000). *Classical and Modern Regression with Applications.2nd edition.* Singapore: Duxburry Thomson Learning.

Pagano, M, Kimberlee Gauvreau, Heather Mattie (2022), *Principles of Biostatistics, 3rd- edition*. Florida: CRC Press.