

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah teknik statistika yang berguna untuk memeriksa dan memodelkan hubungan antar variabel. Analisis regresi terbagi atas regresi linear dan regresi nonlinear. Model regresi nonlinear yang seringkali digunakan dalam analisis data kualitatif atau kategorik adalah regresi logistik (Insani, 2014). Regresi logistik merupakan salah satu metode klasifikasi yang berkaitan dengan respon data yang berupa kategorik, baik nominal maupun ordinal dengan variabel prediktor berupa kuantitatif atau kategorik (Agresti, 2007). Analisis regresi logistik biner digunakan untuk menjelaskan hubungan antara variabel respon yang berupa data dikotomik atau biner yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik dengan variabel prediktor yang berupa data berskala interval dan atau kategorik (Agresti, 2007).

Beberapa penelitian sebelumnya yaitu Nurhasanah, dkk (2021) melakukan penelitian mengetahui faktor-faktor yang mempengaruhi minat siswa SMA di Banda Aceh untuk melanjutkan studi ke Universitas Syiah Kuala, serta untuk membangun model minat siswa melalui pendekatan regresi logistik biner. Oshi dan Dhakal (2021) mengembangkan model prediksi untuk diabetes tipe 2 pada wanita dengan menggunakan regresi logistik, guna meningkatkan pemahaman tentang faktor risiko dan memungkinkan tindakan pencegahan terhadap diabetes. Selanjutnya, Kim dkk. (2023) menggunakan model regresi logistik biner untuk menganalisis faktor-faktor yang mempengaruhi tingkat depresi pada manula berpenghasilan rendah. Namun, penelitian-penelitian tersebut hanya menggunakan data *cross-sectional*. Sementara itu, dalam banyak penelitian di bidang ekonomi, epidemiologi, dan kesehatan, faktor waktu sering kali menjadi pertimbangan penting untuk mendiagnosis perubahan yang terjadi pada individu penelitian seiring berjalannya waktu. Oleh karena itu, penelitian ini mengembangkan model regresi logistik biner dengan menggunakan data longitudinal, yang memungkinkan analisis lebih mendalam terhadap perubahan dari waktu ke waktu.

Data longitudinal adalah gabungan dari data *cross-sectional* dan *time series*, yang diperoleh melalui pengamatan berulang terhadap beberapa individu dalam kurun waktu yang berbeda (Wu & Zhang, 2006). Menurut Wu dan Zhang (2006), terdapat beberapa keuntungan dalam menggunakan data longitudinal. Pertama, studi longitudinal lebih kuat secara statistik dibandingkan dengan studi *cross-sectional* pada sejumlah individu yang tetap. Dengan kata lain, untuk mencapai kekuatan uji statistik yang sama, studi longitudinal memerlukan jumlah individu yang lebih sedikit. Kedua, dengan jumlah individu yang sama,

pengukuran kesalahan (error) dalam data longitudinal menghasilkan estimasi efek perlakuan yang lebih efisien dibandingkan dengan data *cross-sectional*. Ketiga, data longitudinal mampu menyediakan informasi tentang perubahan yang terjadi pada individu dari waktu ke waktu, sementara data *cross-sectional* tidak dapat memberikan informasi tersebut.

Dalam analisis data longitudinal, model dapat dibagi menjadi dua kategori utama, yaitu model efek tetap (*fixed effect*) dan model efek acak (*random effect*). Model yang menggabungkan kedua komponen ini disebut model campuran (*mixed effect*). Dalam data longitudinal, variabel waktu dan individu merupakan komponen penting yang perlu diperhitungkan. Variabel waktu mencatat perubahan yang terjadi pada individu sepanjang periode pengamatan, sementara variabel individu mencerminkan perbedaan antar individu yang dapat mempengaruhi respon yang diamati. Kedua variabel ini berhubungan erat dengan efek acak, karena fluktuasi pada waktu dan perbedaan individu sering kali tidak dapat dijelaskan sepenuhnya oleh variabel prediktor yang tetap. Efek acak digunakan untuk mengakomodasi variasi yang tidak dapat diprediksi yang timbul dari perbedaan antar individu dan perubahan dari waktu ke waktu (Wu & Zhang, 2006). Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model regresi logistik dengan efek campuran, yang mempertimbangkan struktur data longitudinal serta memperhitungkan variabel prediktor sebagai efek acak, guna memperoleh estimasi yang lebih tepat dan akurat.

Kasus yang memuat variabel yang bersifat fixed dan random dapat ditemukan pada data pencapaian target produktivitas alat gali muat dalam kegiatan pemindahan lapisan tanah penutup di PT Kaltim Prima Coal (KPC) tahun 2022. PT KPC, salah satu perusahaan tambang batu bara dengan open-pit terbesar di dunia, berperan penting dalam memenuhi kebutuhan pasar domestik dan internasional. Untuk mencapai hasil produksi yang optimal, PT KPC perlu meningkatkan produktivitas, yang didukung oleh efisiensi kinerja alat-alat mekanis, atau produktivitas alat gali muat. Alat gali muat yang digunakan di PT KPC diukur berulang kali pada interval waktu tertentu, sehingga dapat dilakukan analisis longitudinal. Faktor yang mempengaruhi produktivitas alat gali muat pada aktivitas pemindahan lapisan tanah penutup di PT KPC dipengaruhi oleh variabel tetap yang dapat dikontrol yaitu faktor *spotting time* dan *loading time* yang kesemuanya dapat dikontrol dengan melakukan optimalisasi pada proses produksi. Sedangkan faktor tinggi *bench*, fragmentasi (bucket), dan metode penggalan & material pengaruhnya tidak dapat dikontrol karena data tersebut berdasarkan standar yang digunakan oleh perusahaan yang sulit dikontrol dan dipengaruhi oleh faktor alam sehingga banyak kerandoman yang bisa terjadi. Penelitian yang dilakukan oleh Gonçalves dan Silva (2020) serta Mu dan Wang (2019) mengonfirmasi bahwa faktor-faktor operasional yang dapat dikontrol, seperti waktu

kerja dan kecepatan alat, berperan besar dalam produktivitas alat gali muat, sedangkan faktor eksternal yang tidak dapat dikontrol, seperti kondisi geologi dan cuaca, dapat menambah kerandoman dalam hasil produksi. Berdasarkan uraian di atas, penulis berharap dengan model regresi logistik biner efek campuran pada data *longitudinal* diharapkan dapat mendeskripsikan faktor-faktor yang berpengaruh terhadap pencapaian target produktivitas alat gali muat pada aktivitas pemindahan lapisan tanah penutup di PT Kaltim *Prima Coal* Sangatta sehingga dapat dilakukan upaya untuk peningkatan produktivitas.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diperoleh, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana bentuk estimasi parameter model regresi logistik biner efek campuran pada data longitudinal?
2. Bagaimana bentuk model pencapaian target produktivitas alat gali muat pada aktivitas pemindahan lapisan tanah penutup di PT Kaltim *Prima Coal* menggunakan regresi logistik biner efek campuran pada data longitudinal?

1.3 Batasan Masalah

Agar penelitian ini tidak menimbulkan penafsiran yang lebih luas maka diberikan beberapa asumsi sebagai batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan terdiri dari satu variabel respon yang berskala biner.
2. Penelitian ini menggunakan estimasi parameter model regresi logistik efek campuran dengan *maksimum likelihood estimation*.
3. Menggunakan struktur data longitudinal dengan rentang waktu minggu dalam satu tahun 2022 dengan tiga individu pengamatan berdasarkan tipe alat gali muat atau excavator pada aktivitas pemindahan lapisan tanah penutup di PT Kaltim *Prima Coal*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah:

1. Untuk memperoleh bentuk estimasi parameter model regresi logistik biner efek campuran pada data longitudinal.
2. Untuk memperoleh bentuk model pencapaian target produktivitas alat gali muat pada aktivitas pemindahan lapisan tanah penutup di PT Kaltim *Prima Coal* menggunakan regresi logistik biner efek campuran pada data longitudinal?

1.5 Manfaat Penulisan

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah pemahaman teoritis dan praktis bagi peneliti dan pembaca tentang model regresi logistik biner data longitudinal.
2. Menjadi acuan bagi perusahaan untuk meningkatkan faktor-faktor yang berpengaruh agar dapat dilakukan upaya peningkatan tercapainya produktivitas alat gali muat di PT Kaltim *Prima Coal*.

BAB II TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan analisis untuk menjelaskan hubungan antara dua variabel atau lebih, yaitu variabel respon dan variabel prediktor. Secara umum model regresi linear adalah sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \quad (2.1)$$

dengan $i = 1, 2, \dots, n$; $j = 1, 2, \dots, p$; y_i adalah variabel respon, $x_{ij} = (x_{i1}, x_{i2}, \dots, x_{ip})$ adalah variabel prediktor, $\beta_j = (\beta_1, \beta_2, \dots, \beta_p)$ adalah parameter model regresi yang tidak diketahui, dan ε_i adalah galat yang bersifat acak, berdistribusi normal dengan mean nol dan variansi σ^2 (Montgomery dan Peck, 1992).

Persamaan (2.1) dapat dinyatakan dalam bentuk matriks pada Persamaan (2.2) sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.2)$$

dengan:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \text{ adalah vektor variabel respon berukuran } n \times 1,$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix} \text{ adalah vektor variabel prediktor berukuran } n \times (p + 1),$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \text{ adalah vektor parameter berukuran } (p + 1) \times 1,$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \text{ adalah vektor galat berukuran } (n \times 1).$$

2.2 Regresi Data Longitudinal

Regresi data longitudinal merupakan data yang diperoleh dari pengamatan berulang dalam kurun waktu berbeda pada beberapa subyek. Berbeda dengan data *cross section* yang diperoleh dari pengamatan dalam kurun waktu sekali pada beberapa individu, sehingga data longitudinal memiliki keunggulan dalam menganalisis perubahan individu yang diamati dari waktu ke waktu terhadap respon. Pada data longitudinal diasumsikan pengamatan dilakukan terhadap n individu-individu yang saling independen satu sama lain

yang diamati secara berulang dalam kurun waktu berbeda, akan tetapi antar pengamatan di dalam individu saling dependen sehingga terdapat korelasi (Wu dan Zhang, 2006).

Hubungan antara p variabel prediktor dengan variabel respon tunggal $(x_{1it}, x_{2it}, \dots, x_{pit}, y_{it})$ menyatakan pengamatan pada waktu ke- $t = 1, 2, \dots, T$ dari individu ke- $i = 1, 2, \dots, n$ yang mengikuti model regresi dapat ditulis dalam bentuk Persamaan (2.3) sebagai berikut:

$$y_{it} = \beta_0 + \beta_{i1}x_{1it} + \beta_{i2}x_{2it} + \dots + \beta_{ip}x_{pit} + \varepsilon_{it} \quad (2.3)$$

dengan:

y_{it} : variabel respon untuk unit individu ke- i dan unit waktu ke- t ,

β : parameter intersep atau konstanta,

β_{ip} : parameter koefisien prediktor ke- $j = 1, 2, \dots, p$ untuk unit individu ke- i ,

x_{pit} : variabel prediktor ke- $j = 1, 2, \dots, p$ untuk unit individu ke- i dan unit waktu ke- t ,

ε : galat untuk unit individu ke- i dan periode waktu ke- t ,

i : $1, 2, \dots, n$ adalah unit *cross section* (individu),

t : $1, 2, \dots, T$ adalah unit *time series* (waktu).

Adapun struktur dari data longitudinal secara umum dapat dilihat pada Tabel 1 sebagai berikut:

Tabel 1. Struktur Data Longitudinal Secara Umum

Individu ke- i	Pengamatan ke- t	Variabel Respon	Variabel Prediktor ke- j			
		Y_{it}	X_{1it}	X_{2it}	$\dots$	X_{pit}
1	1	y_{11}	x_{111}	x_{211}	$\dots$	x_{p11}
	2	y_{12}	x_{112}	x_{212}		x_{p12}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	T	y_{1T}	x_{1nT}	x_{2nT}		x_{pnT}
2	1	y_{21}	x_{121}	x_{221}	$\dots$	x_{p21}
	2	y_{22}	x_{121}	x_{222}		x_{p22}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	T	y_{2T}	x_{1nT}	x_{2nT}		x_{pnT}

Individu ke- i	Pengamatan ke- t	Varaiabel Respon	Variabel Prediktor ke- j			
		Y_{it}	X_{1it}	X_{2it}	...	X_{pit}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
n	1	y_{n1}	x_{1n1}	x_{2n1}	...	x_{pn1}
	2	y_{n2}	x_{n21}	x_{2n2}		x_{pn2}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	T	y_{nT}	x_{1nT}	x_{2nT}		x_{pnT}

2.3 Model Linear Efek Campuran

Model dengan dua komponen disebut dengan model campuran (*mixed effect*) karena sebagian variabel prediktor merupakan efek tetap dan variabel prediktor lainnya merupakan efek random (Michel Mouchart, 2009). Variabel respon y selain dipengaruhi variabel prediktor x atau dikenal dengan efek tetap (*fixed effect*) juga dipengaruhi oleh faktor keacakan dalam pemilihan sampel, waktu, area dan lain lain yang dikenal dengan efek acak (*random effect*). Model persamaan regresi efek campuran adalah:

$$y = x\beta + z\gamma + \varepsilon \quad (2.4)$$

dengan:

- y : vektor variabel respon,
- x : vektor variabel prediktor untuk efek tetap,
- β : vektor parameter koefisien efek tetap,
- z : vektor variabel prediktor untuk efek acak,
- γ : vektor parameter koefisien efek acak,
- ε : vektor galat acak.

Efek acak γ dan ε saling bebas dan γ diasumsikan berdistribusi normal multivariat dengan rata-rata nol dan variansi σ^2 atau $\gamma \sim N(\mu_\gamma, \sigma_\gamma^2)$ sedangkan *error* diasumsikan independen identik berdistribusi normal dengan rata-rata nol dan variansi σ^2 atau $\varepsilon \sim IID N(0, \sigma^2 I_n)$.

2.4 Regresi Logistik Biner

Regresi logistik adalah suatu metode analisis statistika untuk menganalisis hubungan antara variabel respon berupa data kualitatif dengan variabel prediktor berupa data kualitatif maupun data kuantitatif. Berbeda dengan regresi linier biasa, regresi logistik tidak mengasumsikan hubungan antara variabel respon dan prediktor secara linier. Regresi logistik termasuk dalam model linear umum (*Generalized Linear Models* atau GLM) dimana model linear umum merupakan pengembangan dari model linear klasik. Analisis regresi logistik dilakukan pada data yang memiliki peubah prediktor dengan peubah respon yang bersifat kategorik (berskala nominal atau ordinal) dengan p variabel prediktor berskala kategorik atau kontinu (berskala interval atau rasio). Menurut Hosmer dkk (2013), model probabilitas regresi logistik dengan variabel prediktor $x_1, x_2, \dots, x_p$ adalah:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})} \quad (2.5)$$

Fungsi $\pi(x_i)$ pada regresi logistik merupakan fungsi non linear sehingga perlu dilakukan transformasi ke dalam bentuk logit untuk memperoleh fungsi linear agar dapat dilihat hubungan antara variabel respon dan variabel prediktornya, sehingga dirumuskan sebagai berikut:

$$\begin{aligned} g(x_i) &= \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) \\ &= \ln\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}}{1 - \frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}}\right) \\ &= \ln\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}}{\frac{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}) - \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}}\right) \\ &= \ln\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}}{1}\right) \\ &= \ln\left(\frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})} (1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}))\right) \\ &= \ln(\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip})) \\ &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \end{aligned}$$

atau

$$g(x_i) = \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} \quad (2.6)$$

dengan:

$\pi(x_i)$: peluang kejadian sukses pada $y = 1$,

$g(x_i)$: nilai estimasi logit,

β_0 : konstanta (intersep),

β_p : parameter variabel prediktor ke- $j = (1, 2, \dots, p)$,

x_{ij} : variabel prediktor ke- $j = (1, 2, \dots, p)$ pada pengamatan ke- $i = (1, 2, \dots, n)$,

p : jumlah variabel prediktor.

Apabila variabel respon terdiri dari dua kategori maka metode regresi logistik yang dapat digunakan adalah regresi logistik biner (Hosmer & Lemeshow, 2013). Model regresi logistik biner termasuk dalam sebaran keluarga Eksponensial. Sebaran keluarga Eksponensial yang dimaksud adalah sebaran Bernoulli, yaitu sebaran dari peubah acak yang hanya mempunyai 2 kategori yaitu 0 dan 1 (Islamiyati, 2015). Variabel respon pada regresi logistik biner terdiri 2 kategori, misalkan $y = 1$ menyatakan hasil yang diperoleh 'sukses' dan $y = 0$ menyatakan hasil yang diperoleh 'gagal'. Karena variabel y hanya memiliki dua kategori, maka variabel y tersebut mengikuti distribusi Bernoulli (Hosmer & Lemeshow, 2000) dengan fungsi probabilitas:

$$f(y_i) = \pi_i(x_i)^{y_i}(1 - \pi_i(x_i))^{1-y_i}; y = 0, 1 \quad (2.7)$$

Jika gagal $y_i = 0$ maka $f(y_i) = 1 - \pi_i(x_i)$,

Jika sukses $y_i = 1$ maka $f(y_i) = \pi_i(x_i)$.

2.5 Metode Maximum Likelihood Estimation

Salah satu metode yang umumnya digunakan untuk mendapatkan estimator pada regresi logistik yaitu dengan metode estimasi likelihood maksimum (*Maximum Likelihood Estimation* (MLE)) yang diperkenalkan oleh R.A Fisher (Hosmer & Lemeshow, 2013). Prinsip dari metode MLE ini adalah untuk menemukan nilai estimasi β dengan memaksimumkan fungsi *likelihoodnya* dan mensyaratkan bahwa data harus mengikuti suatu distribusi tertentu. Jika regresi logistik biner mengasumsikan setiap pengamatan bersifat independen maka bentuk umum fungsi kepadatan mengikuti distribusi Bernoulli (Agresti, 2007), yaitu:

$$f(y_i) = \pi(x_i)^{y_i}(1 - \pi(x_i))^{1-y_i}$$

Maka fungsi *likelihood*

$$L(\beta) = \prod_{i=1}^n f(y_i)$$

$$= \prod_{i=1}^n (\pi(x_i)^{y_i} (1 - \pi(x_i))^{(1-y_i)}) \quad (2.8)$$

Untuk mendapatkan nilai β dengan memaksimumkan nilai fungsi *likelihood* digunakan bentuk logaritma natural dari fungsi *likelihood*, yang kemudian disebut fungsi *log-likelihood*, sehingga dapat ditulis Persamaan (2.9):

$$\begin{aligned} \ell &= \ln L(\beta) \\ &= \ln \left(\prod_{i=1}^n (\pi(x_i)^{y_i} (1 - \pi(x_i))^{(1-y_i)}) \right) \\ &= \sum_{i=1}^n (\pi(x_i)^{y_i} (1 - \pi(x_i))^{(1-y_i)}) \end{aligned} \quad (2.9)$$

Persamaan (2.9) diturunkan terhadap parameter β dan disamakan dengan nol, sehingga diperoleh Persamaan (2.10):

$$\sum_{i=1}^n x_i \left(y_i - \frac{\exp(x_i' \hat{\beta})}{1 + \exp(x_i' \hat{\beta})} \right) = 0 \quad (2.10)$$

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 Newton Raphson yaitu memaksimumkan fungsi *likelihood*.

Metode *Newton Raphson* merupakan metode iterasi untuk menyelesaikan persamaan nonlinear. Berdasarkan iterasi *Newton Raphson* akan diperoleh turunan kedua dari fungsi *In-likelihood* terhadap masing-masing parameter. Formulasi iterasi *Newton Raphson* dengan memulai iterasi $t = 0$ adalah Persamaan (2.11) berikut:

$$\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{-1} g^{(t)} \quad (2.11)$$

dengan:

- t : iterasi ke-0,1,2, ..., T
- $g^{(t)}$: matriks turunan pertama terhadap masing-masing parameter β
- $H^{(t)}$: matriks turunan kedua terhadap masing-masing parameter β .

Proses iterasi dengan menggunakan metode *Newton Raphson* akan berhenti jika terpenuhi kondisi konvergen, yaitu ketika $\|\beta^{(t+1)} - \beta^{(t)}\| \leq \varepsilon$, dimana ε adalah bilangan positif yang sangat kecil (Ehlers, 2002).

2.6 Pengujian Signifikansi Parameter

Uji yang digunakan untuk menguji parameter model regresi logistik ordinal adalah uji simultan untuk menentukan variabel yang signifikan kemudian dilanjutkan dengan uji parsial (Hosmer dan Lemeshow, 2013).

2.6.1 Uji Simultan

Pengujian ini dilakukan untuk mengetahui signifikansi model dan pengaruh variabel respon secara bersama-sama. Statistik uji yang digunakan untuk uji simultan adalah Uji G (*Likelihood Ratio Test*) (Hosmer dan Lemeshow, 2013). Hipotesis yang digunakan adalah sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (tidak ada pengaruh variabel prediktor secara simultan terhadap variabel respon).

H_1 : Paling tidak terdapat satu $\beta_j \neq 0$ (ada pengaruh paling sedikit satu variabel prediktor secara simultan terhadap variabel respon).

Statistik uji yang digunakan adalah:

$$G = -2 \ln \left[\frac{L_0(\omega)}{L_p(\omega)} \right] = 2 [\ln L_p(\omega) - \ln L_0(\omega)] \sim X_{\alpha, p}^2 \quad (2.12)$$

dengan

$L_0(\omega)$: Nilai fungsi *likelihood* tanpa variabel prediktor.

$L_p(\omega)$: Nilai fungsi *likelihood* dengan variabel prediktor.

p : Jumlah variabel prediktor

Kriteria keputusan:

H_0 ditolak jika nilai $G_{hit} > X_{\alpha, p}^2$ atau nilai $p - value < \alpha$

H_0 diterima jika nilai $G_{hit} < X_{\alpha, p}^2$ atau nilai $p - value > \alpha$

2.6.2 Uji Parsial

Uji parsial merupakan pengujian terhadap besarnya tingkat signifikansi dari variabel prediktor secara individu terhadap variabel respon. Menurut Hosmer dan Lemeshow (2013) untuk menguji signifikansi parameter digunakan Uji *Wald*. Hipotesis yang digunakan adalah sebagai berikut:

$H_0 : \beta_j = 0$, dengan $j = 1, 2, \dots, p$ (tidak ada pengaruh variabel prediktor ke- j terhadap variabel respon)

$H_1 : \beta_j \neq 0$, dengan $j = 1, 2, \dots, p$ (ada pengaruh variabel prediktor ke- j terhadap variabel respon)

Statistik Uji Wald yaitu:

$$W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \sim X_{\alpha,1}^2 \quad (2.13)$$

dengan:

$\hat{\beta}_j$: penduga estimator β_j

$SE(\hat{\beta}_j)$: penduga standar error dari estimator $\hat{\beta}_j$

Kriteria keputusan:

H_0 ditolak jika nilai $|W_j| > X_{\alpha,1}^2$ atau nilai $p - value < \alpha$

H_0 diterima jika nilai $|W_j| > X_{\alpha,1}^2$ atau nilai $p - value > \alpha$

2.7 Pemilihan Model Terbaik

Kriteria pemilihan model terbaik pada analisis regresi logistik digunakan untuk mendapatkan model yang terbaik yang mampu menjelaskan pola hubungan antara variabel prediktor dengan variabel respon, serta model yang terbaik yang mampu secara akurat melakukan prediksi. Untuk mendapatkan model terbaik yang mampu menjelaskan faktor-faktor yang berpengaruh terhadap variabel respon dapat digunakan kriteria AIC. Sedangkan untuk mendapatkan model terbaik yang mampu memprediksi secara akurat dapat digunakan nilai ketepatan klasifikasi.

2.7.1 Akaike's Information Criterion (AIC)

Pemilihan model terbaik berdasarkan kriteria AIC dilakukan dengan memilih model yang menghasilkan nilai AIC terkecil. Hal tersebut dikarenakan besarnya nilai AIC sejalan dengan nilai devians dari model. Semakin kecil nilai devians maka akan semakin kecil pula tingkat kesalahan yang dihasilkan sehingga model yang diperoleh akan semakin tepat. Adapun untuk mendapatkan nilai AIC dapat digunakan rumus sebagai berikut (Johnson & Wichern, 1992):

$$AIC = -2\log l(\beta_k) + 2p \quad (2.14)$$

Keterangan:

p : Jumlah parameter

$l(\beta_k)$: Maksimum *likelihood* dari model

2.7.2 Ketepatan Klasifikasi

Salah satu cara untuk melakukan evaluasi terhadap model regresi logistik biner adalah dengan menggunakan tabel klasifikasi (Hosmer & Lemeshow, 2013). Tabel klasifikasi merupakan tabel yang menyatakan jumlah data uji hasil prediksi yang diklasifikasikan benar atau salah berdasarkan data uji hasil sebenarnya.

Tabel 2. Tabel Klasifikasi

Hasil Observasi	Hasil Prediksi	
	π_0	π_1
π_0	n_{00}	n_{10}
π_1	n_{01}	n_{11}

Keterangan:

π_0 : Prediksi kategori 0 atau tidak tercapai,

π_1 : Prediksi kategori 1 atau tercapai,

n_{00} : jumlah y_i dari π_0 tepat diklasifikasikan sebagai π_0 ,

n_{01} : jumlah y_i dari π_0 salah diklasifikasikan sebagai π_1 ,

n_{10} : jumlah y_i dari π_1 salah diklasifikasikan sebagai π_0 ,

n_{11} : jumlah y_i dari π_1 tepat diklasifikasikan sebagai π_1 .

Akurasi menunjukkan tingkat performa model secara keseluruhan, dimana semakin tinggi akurasi hal ini berarti semakin baik tingkat performa model tersebut. Perhitungan akurasi menggunakan fungsi sebagai berikut (Han, Kamber, & Pei, 2012):

$$akurasi = \frac{n_{00} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}} \times 100\% \quad (2.15)$$

2.8 Odds Ratio

Pada pemodelan regresi logistik, pemahaman mengenai odds ratio (OR) sangat diperlukan. Rasio odds berperan dalam memudahkan interpretasi model regresi logistik yang diperoleh. Interpretasi parameter bertujuan untuk mengetahui arti dari nilai taksiran parameter pada variabel prediktor. Interpretasi koefisien dalam model regresi logistik dilakukan dalam bentuk *odds ratio* (Hosmer dan Lemeshow, 2013). Rasio odds berhubungan dengan transformasi logit, seperti yang diketahui agar menjadi bentuk yang linear fungsi logistik perlu ditransformasi sedemikian rupa. Tranformasi logit dapat dituliskan:

$$\ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

Sehingga model *odds ratio* dituliskan pada persamaan (2.16) sebagai berikut:

$$OR = \frac{\pi(x_i)}{1 - \pi(x_i)} = \exp\left(\beta_0 + \sum_{j=1}^k \beta_j x_{ij}\right) \quad (2.16)$$

2.9 Produktivitas OB Removal di PT KPC

Menurut UU No.4 Tahun 2009, pertambangan adalah sebagian atau seluruh tahapan kegiatan dalam rangka penelitian, pengelolaan dan pengusahaan mineral atau batubara yang meliputi penyelidikan umum, eksplorasi, studi kelayakan, konstruksi, penambangan, pengolahan dan pemurnian, pengangkutan dan penjualan, serta kegiatan pasca tambang. Salah satu produsen batubara terkemuka di Indonesia yang memenuhi kebutuhan dunia, memberikan nilai optimal bagi semua pemangku kepentingan adalah PT KPC. PT Kaltim Prima Coal (KPC) merupakan perusahaan berbadan hukum Indonesia yang bergerak di bidang pertambangan dan penjualan batubara baik untuk pelanggan domestik maupun internasional dari berbagai sektor industri. KPC merupakan salah satu anak perusahaan dari PT Bumi Resources (Tbk) yang berlokasi di Kabupaten Kutai Timur, Provinsi Kalimantan Timur dengan aktivitas penambangannya dilakukan dengan system tambang terbuka (*open-pit*). KPC mengelola area konsesi pertambangan dengan luas mencapai 84,938 hektar. Dengan didukung oleh lebih dari 4.499 orang karyawan dan 21.000 personel dari kontraktor dan perusahaan terkait, kapasitas produksi batubara KPC mencapai 70 juta ton per tahun (KPC, 2022).

Ada 14 divisi di PT KPC, salah satunya adalah *Mine Operation Division* (MOD) yang bertanggungjawab terhadap pelaksanaan kegiatan pertambangan yang terdiri dari kegiatan konstruksi infrastruktur dalam tambang, penambangan, dan pengangkutan menuju tempat pengolahan dan pemurnian. Kegiatan pertambangan yang dilakukan oleh PT KPC dimulai dari survey eksplorasi hingga pemuatan batubara ke kapal Gambar 3 (*Technical Handbook MOD_First Edition*, 2017).



Gambar 3 Bagan Alir Proses Penambangan PT KPC

Salah satu kegiatan yang sangat berpengaruh dalam kegiatan penambangan dan merupakan kegiatan awal sebelum proses penambangan dilakukan adalah kegiatan

pemindahan lapisan tanah (*overburden removal*). Tanah penutup (*overburden*) adalah semua lapisan tanah/batuan yang berada di atas dan langsung menutupi lapisan bahan galian berharga (batu batubara) sehingga perlu disingkirkan terlebih dahulu sebelum dapat menggali bahan galian berharga (batu batubara) tersebut. Dalam mengekspos atau membuka lapisan batubara diperlukan peralatan mekanis seperti alat gali muat digunakan untuk penggalian di *front* dan pengangkutan ke *disposal area* atau *dumping point*.

Kegiatan penambangan harus diselesaikan dengan target produksi yang ingin dicapai, PT KPC mempunyai target produksi 60 juta ton/tahun (KPC, 2022). Target produksi pemindahan overburden sangat berpengaruh terhadap hasil produksi perusahaan tersebut sehingga perlu adanya evaluasi produktivitas untuk dapat mengetahui apakah target produksi tercapai atau tidak. Tercapainya target produksi didukung oleh beberapa faktor penting seperti efisiensi kinerja alat-alat mekanis atau dikenal dengan istilah produktivitas alat gali muat. Produktivitas adalah suatu pendekatan interdisipliner untuk menentukannya tujuan yang efektif atau pembuat rencana yang dimana menggunakan sumber-sumber yang efisien namun kualitas yang tinggi tetap terjaga (Sinungan, 2003). Semua kegiatan pemindahan tanah penutup/ OB Removal dan penambangan batubara di PT KPC direkam oleh *dispatch system*. Produktivitas di PT KPC dihitung seberapa lama unit excavator memuat suatu truck/alat angkut material OB dalam satuan BCM per-konversi 1 jam kerja (workinghour) Bcm/hour. Rumus produktivitas alat muat (*Technical Handbook MOD_First Edition*, 2017):

$$productivity = \frac{OB\ Bcm}{spot\ time + load\ time} \quad (2.17)$$

Keterangan:

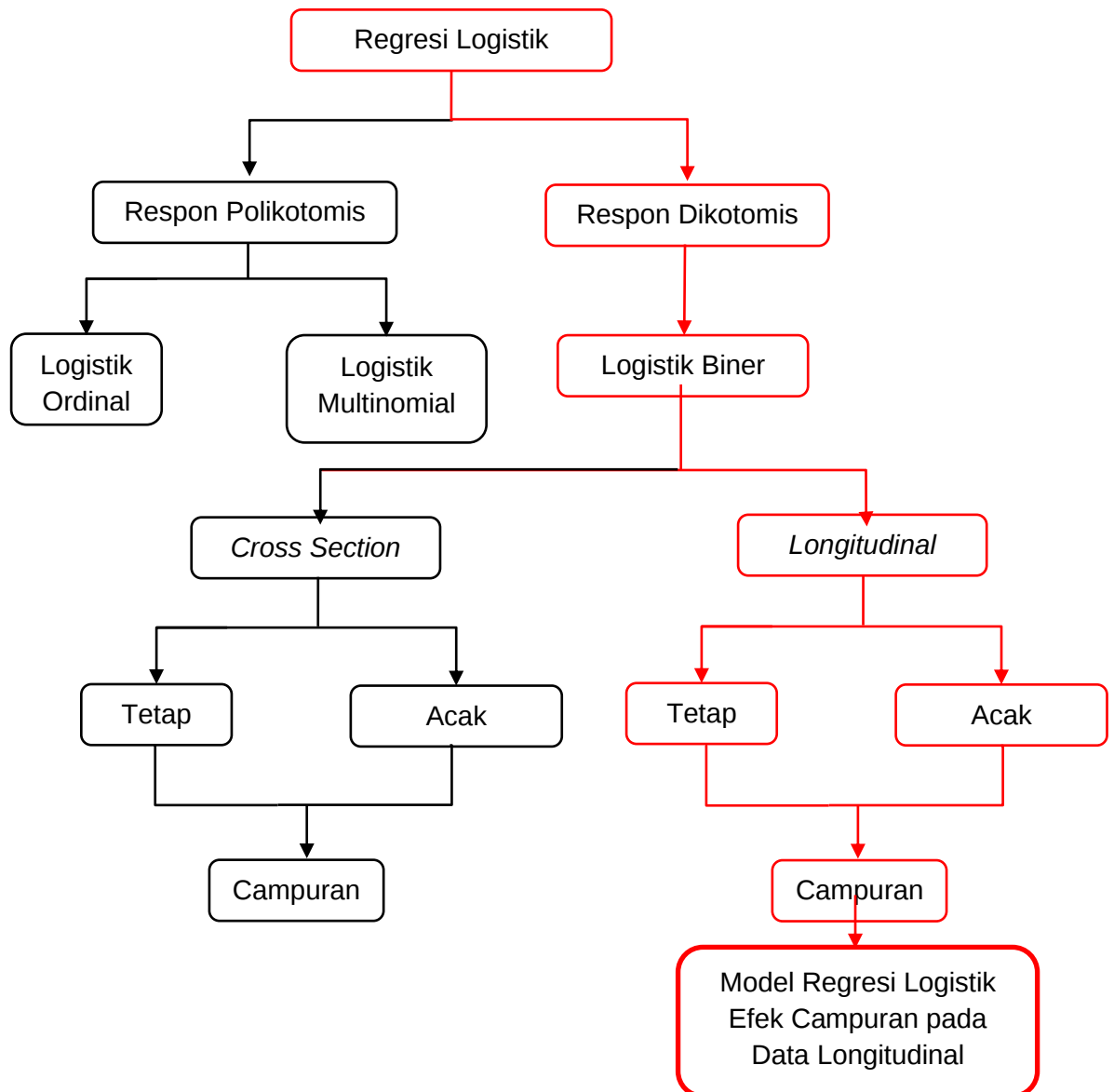
- productivity* = produktivitas alat gali muat dan angkut dalam satuan BCM per jam
- spot time* = waktu yang digunakan oleh *truck* untuk *maneuver* di *loadingpoint*, pada kegiatan pemuatan *digger* ke alat angkut.
- load time* = waktu/ durasi yang digunakan *digger* untuk memuat suatu OB *truck*.

PT KPC mengoperasikan alat gali dengan beberapa tipe untuk aktivitas pemidahan lapisan tanah penutup. Adapun target produktivitas alat muayang ditetapkan di PT KPC (*Technical Handbook MOD_First Edition*, 2017):

- 1) Excavator *Backhoe* Liebherr R9800 sebanyak tiga unit dengan produktivitas di PT KPC adalah 3069 Bcm/hr.
- 2) Excavator *Backhoe* Liebherr R996 sebanyak tiga unit dengan produktivitas di PT KPC adalah 2000 Bcm/hr.

- 3) Excavator *Backhoe* Hitachi R3600 sebanyak tiga unit dengan produktivitas di PT KPC adalah 1255 Bcm/hr.

2.10 Kerangka Konseptual



Gambar 4. Kerangka Konseptual