

**PENERAPAN METODE ANALISIS REGRESI LOGISTIK BINER PADA  
DATA INDEKS PEMBANGUNAN MANUSIA DI KALIMANTAN TAHUN  
2023**

**(Skripsi)**

**Oleh  
MARISA ELFAIRA DAFITRI  
1817031087**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2025**

## **ABSTRACT**

### **IMPLEMENTATION OF BINARY LOGISTIC REGRESSION METHOD FOR HUMAN DEVELOPMENT INDEX DATA IN KALIMANTAN IN 2023**

**By**

**MARISA ELFAIRA DAFITRI**

Binary logistic regression is a method used to find the relationship between dependent variable (Y) that is binary (dichotomus) and independent variables (X). This study aims to get the best binary logistic regression model and to know the accuracy of the classification of factors that affect human development index in Kalimantan in 2023. This study uses data on human development index in Kalimantan in 2023. The results of this study is factor that significantly affect human developments index is the expected length of schooling factor and average length of schooling factor. The binary logictic regression model obtained is human development index =  $-27,9688 + 2,4074(\text{the expected length of schooling}) X_1 - 0,3472(\text{percentage of poor people}) X_2$  with the accuracy of the classification obtained is 80,36%.

**Keywords** : Binary logistic regression analysis, human development index

## **ABSTRAK**

### **PENERAPAN METODE ANALISIS REGRESI LOGISTIK BINER PADA DATA INDEKS PEMBANGUNAN MANUSIA DI KALIMANTAN TAHUN 2023**

**Oleh**

**MARISA ELFAIRA DAFITRI**

Regresi logistik biner ialah metode yang berfungsi untuk mencari hubungan antara variabel dependen (Y) bersifat biner (*dichotomus*) dengan variabel independen (X). Penelitian ini bertujuan untuk memperoleh model regresi logistik biner terbaik serta mengetahui ketepatan klasifikasi terhadap faktor yang mempengaruhi indeks pembangunan manusia di Kalimantan pada tahun 2023. Data yang digunakan pada penelitian ini yaitu data tentang indeks pembangunan manusia di Kalimantan pada tahun 2023. Hasil yang diperoleh dalam penelitian ini yaitu faktor yang mempengaruhi indeks pembangunan manusia secara signifikan ialah variabel harapan lama sekolah dan variabel rata-rata lama sekolah. Model regresi logistik biner yang diperoleh yaitu indeks pembangunan manusia =  $-27,9688 + 2,4074(\text{harapan lama sekolah}) X_1 - 0,3472(\text{persentase penduduk miskin}) X_2$  dengan ketepatan klasifikasi yang diperoleh sebesar 80,36%.

**Kata Kunci :** Analisis regresi logistik biner, indeks pembangunan manusia

**PENERAPAN METODE ANALISIS REGRESI LOGISTIK BINER PADA  
DATA INDEKS PEMBANGUNAN MANUSIA DI KALIMANTAN TAHUN  
2023**

Oleh

*Marisa Elfaira Dafitri*

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar  
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2025**



Judul Skripsi

**: PENERAPAN METODE ANALISIS  
REGRESI LOGISTIK BINER PADA DATA  
INDEKS PEMBANGUNAN MANUSIA DI  
KALIMANTAN TAHUN 2023**

Nama Mahasiswa

**: Marisa Elfaira Dafitri**

Nomor Pokok Mahasiswa

**: 1817031087**

Jurusan

**: Matematika**

Fakultas

**: Matematika dan Ilmu Pengetahuan Alam**



**MENYETUJUI**

**1. Komisi Pembimbing**

*Mustafa*

**Prof. Drs. Mustofa Usman, M.A., Ph.D.**  
NIP. 195701011984041001

*AA*

**Dr. Notiragayu, S.Si., M.Si.**  
NIP. 197311092000122001

**2. Ketua Jurusan Matematika**

*Mary*

**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP. 19740316 200501 1 001



**MENGESAHKAN**

1. **Tim Penguji**

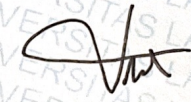
**Ketua : Prof. Drs. Mustofa Usman, M.A., Ph.D.** .....



**Sekretaris : Dr. Notiragayu, S.Si., M.Si.** .....



**Penguji Bukan Pembimbing : Drs. Nusyirwan, M.Si.** .....



2. **Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Dr. Eng. Heri Satria, S.Si., M.Si.**  
NIP. 19711001 200501 1 002

**Tanggal Lulus Ujian Skripsi : 30 Januari 2025**



## PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : **Marisa Elfaira Dafitri**

Nomor Pokok Mahasiswa : **1817031087**

Jurusan : **Matematika**

Judul Skripsi : **PENERAPAN METODE ANALISIS REGRESI  
LOGISTIK BINER PADA DATA INDEKS  
PEMBANGUNAN MANUSIA DI  
KALIMANTAN TAHUN 2023**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 12 Februari 2025

Penulis



Marisa Elfaira Dafitri

## **RIWAYAT HIDUP**

Penulis bernama lengkap Marisa Elfaira Dafitri dilahirkan di Bandar Lampung pada 29 Maret 2000. Penulis merupakan anak kedua dari dua bersaudara dari pasangan Bapak Sugianto, S.E. dan Ibu Dra. Endah Sulistiani.

Penulis pertama kali menempuh pendidikan awal di TK KARTIKA II-6 Bandar Lampung pada tahun 2005-2006. Penulis melanjutkan pendidikannya ke sekolah dasar di SD KARTIKA II-5 Bandar Lampung pada tahun 2006-2012. Selanjutnya, penulis melanjutkan jenjang pendidikannya di SMP Negeri 1 Bandar Lampung pada tahun 2012-2015, dan sekolah menengah atas di SMA Negeri 9 Bandar Lampung pada tahun 2015-2018. Pada tahun 2018 penulis terdaftar sebagai mahasiswa S1 Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur SBMPTN.

Selama menjadi mahasiswa penulis aktif dalam organisasi Himpunan Mahasiswa Matematika (HIMATIKA) FMIPA UNILA pada Bidang Eksternal sebagai Anggota Pengurus periode 2019. Pada tahun 2021, Penulis telah melaksanakan Kuliah Kerja Nyata (KKN) di Kecamatan Langkapura, Kota Bandar Lampung dan Kerja Praktik (KP) di Badan Koordinasi Keluarga Berencana Nasional (BKKBN) Provinsi Lampung. Selama menjadi mahasiswa, Penulis juga mengikuti program Kampus Merdeka yaitu Studi Independen Bersertifikasi (SIB) di PT. Microsoft Indonesia dan Mari Belajar.



## KATA INSPIRASI

*Dan apabila hamba-hamba-Ku bertanya kepadamu (Muhammad) tentang Aku, maka sesungguhnya Aku dekat. Aku Kabulkan permohonan orang yang berdoa apabila dia berdoa kepada-Ku. Hendaklah mereka itu memenuhi (perintah)-Ku dan beriman kepada-Ku, agar mereka memperoleh kebenaran*  
(Q.S. Al-Baqarah: 186)

*Sungguh atas kehendak Allah semua ini terwujud. Tiada kekuatan kecuali dengan pertolongan Allah.*  
(Q.S. Al-Kahfi: 39)

*Takdir Tuhan Ada Di Ujung Usaha Manusia*  
(Habiburrahman El-Shirazy, Novel "Bumi Cinta")

*"Education never ends, Watson. It is a series of lessons with the greatest for the last."*  
(Sir Arthur Conan Doyle on behalf of Sherlock Holmes)

*Believe in yourself. If you don't take chances, then you'll never know what it's going to be like.*  
(Niall James Horan)

## **PERSEMBAHAN**

Alhamdulillahirabbil'alamin.

Segala puji bagi Allah SWT atas rahmat dan karunia-Nya sehingga skripsi ini dapat terselesaikan dengan baik. Salawat serta salam kepada junjungan kita Nabi Muhammad SAW. Dengan penuh rasa syukur penulis persembahkan karya tulis ini kepada:

### **Bapak Sugianto, S.E. dan Ibu Dra. Endah Sulistiani.**

Sebagai tanda bakti, hormat dan rasa terima kasihku kepada Bapak dan Ibu yang telah memberikan dukungan, kasih sayang, doa, serta semangat selama ini. Terimakasih telah mempercayai saya bahwa dapat menyelesaikan pendidikan ini.

### **Kakakku Ryan Adi Saputra dan Keluarga**

Terima kasih kepada kakakku dan keluarga yang selalu memberikan doa, dukungan dan memberikan semangat selama langkah ini.

### **Dosen Pembimbing dan Pembahas**

Terima kasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu dalam membimbing, memberikan motivasi, memberikan arahan serta ilmu yang berharga.

### **Rekan dan Sahabat**

Terima kasih kepada para sahabat yang telah memberikan banyak dukungan dan kebersamaannya. Serta para rekan yang telah kebersamai setiap langkah ini.

**Almamater Tercinta, Universitas Lampung**



## SANWACANA

Puji dan syukur penulis ucapkan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Metode Analisis Regresi Logistik Biner Pada Data Indeks Pembangunan Manusia Di Kalimantan Tahun 2023” yang merupakan salah satu syarat untuk memperoleh gelar Sarjana Matematika di Universitas Lampung.

Penulis menyadari bahwa dalam penulisan skripsi ini tidak terlepas dari bimbingan, dukungan, bantuan dan doa dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku Dosen Pembimbing I yang senantiasa membimbing, memberi masukan, serta arahan kepada penulis dalam menyelesaikan skripsi ini.
2. Ibu Dr. Notiragayu, S.Si., M.Si., selaku Pembimbing II yang telah memberikan motivasi, bimbingan arahan serta masukan kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Bapak Drs. Nusyirwan, M.Si., selaku dosen penguji yang telah memberikan kritik dan saran selama proses penyusunan skripsi ini.
4. Ibu Dr. Fitriani, S.Si., M.Sc., selaku Pembimbing Akademik yang telah memberikan arahan, bimbingan, masukan, dan nasihat sehingga penulis dapat menyelesaikan proses perkuliahan dengan baik
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

7. Seluruh dosen, staff, serta karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Bapak, Ibu, Mas Adi, Mba Widy, Mikhayla, Tavisha dan seluruh keluarga besar yang selalu memberikan doa, dukungan, dan semangat kepada penulis.
9. Sahabat-sahabat seperjuangan, Reajeng, Putsal, Elsa, Refi, Amel, Eja, dan Hanifah yang selalu memberikan bantuan, dukungan, motivasi dan berbagi keceriaan bersama penulis selama masa perkuliahan.
10. Sahabat-sahabat terbaikku, Akira, Yayas, dan Erika yang selalu memberikan dukungan, bantuan, motivasi, serta menjadi tempat penulis berkeluh kesah.
11. Zeta, Riris, dan Rika yang telah menjadi tempat bertukar pikiran, memberikan saran, dan saling memberikan semangat selama proses penyelesaian skripsi.
12. Teman-teman Matematika 2018 atas kebersamaan selama masa perkuliahan hingga akhir.
13. Semua pihak yang telah membantu yang tidak bisa penulis sebutkan satu persatu.

Penulis menyadari masih banyak kekurangan dalam penulisan skripsi ini. Oleh karena itu, kritik dan saran sangat diharapkan agar dapat menjadi pelajaran dan perbaikan untuk kedepannya. Semoga skripsi ini dapat memberikan manfaat baik bagi penulis maupun bagi pihak yang membutuhkan.

Bandar Lampung, 12 Februari 2025  
Penulis,

Marisa Elfaira Dafitri



## DAFTAR ISI

Halaman

<b>DAFTAR TABEL .....</b>	<b>xv</b>
<b>DAFTAR GAMBAR.....</b>	<b>xvi</b>
<b>I. PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang dan Masalah .....	1
1.2. Tujuan Penelitian.....	3
1.3. Manfaat Penelitian.....	3
<b>II. TINJAUAN PUSTAKA.....</b>	<b>4</b>
2.1. Analisis Regresi Logistik Biner .....	4
2.2. Multikolinearitas .....	6
2.3. Maximum Likelihood Estimation (MLE) .....	7
2.4. Uji Simultan .....	11
2.5. Uji Parsial .....	12
2.6. Uji Kesesuaian Model .....	13
2.7. Interpretasi Koefisien Parameter .....	14
2.8. Ketepatan Klasifikasi .....	15
2.9. Indeks Pembangunan Manusia.....	16
<b>III. METODOLOGI PENELITIAN.....</b>	<b>18</b>
3.1. Waktu dan Tempat Penelitian .....	18
3.2. Data Penelitian .....	18
3.3. Metode Penelitian.....	21

<b>IV. PEMBAHASAN.....</b>	<b>23</b>
4.1. Statistika Deskriptif .....	23
4.2. Menguji Asumsi Multikolinearitas.....	25
4.3. Pembentukan Model Awal Regresi Logistik Biner.....	26
4.4. Pengujian Estimasi Parameter Secara Simultan Model Awal .....	27
4.5. Pengujian Estimasi Parameter Secara Parsial Model Awal .....	28
4.6. Model Akhir Regresi Logistik Biner.....	29
4.7. Uji Kesesuaian Model .....	30
4.8. Interpretasi Koefisien Parameter .....	31
4.9. Ketepatan Klasifikasi .....	43
<b>V. KESIMPULAN .....</b>	<b>44</b>
<b>DAFTAR PUSTAKA .....</b>	<b>45</b>
<b>LAMPIRAN.....</b>	<b>47</b>



## DAFTAR TABEL

Tabel	Halaman
1. Koefisien Parameter .....	14
2. Perhitungan Ketepatan Klasifikasi .....	15
3. Data Penelitian .....	19
4. Analisis Statistika Deskriptif Variabel Independen .....	24
5. Nilai VIF Pada Data Indeks Pembangunan Manusia .....	26
6. Nilai Estimasi Parameter Model Awal .....	26
7. Hasil Uji G ( <i>Likelihood Ratio</i> ) .....	27
8. Hasil Uji Wald .....	28
9. Nilai Estimasi Parameter Model Akhir .....	30
10. Hasil Uji Kessesuaian Model .....	30
11. Hasil <i>Odds Ratio</i> .....	31
12. Nilai <i>Odds</i> Dan <i>Probabilitas</i> .....	41
13. Ketepatan Klasifikasi .....	43

## DAFTAR GAMBAR

Gambar	Halaman
1. Diagram Alir Analisis Regresi Logistik Biner.....	22
2. Persentase Indeks Pembangunan Manusia (IPM) di Kalimantan Tahun 2023.....	23



## I. PENDAHULUAN

### 1.1. Latar Belakang dan Masalah

Analisis regresi adalah salah satu alat analisis yang bertujuan untuk mengetahui pengaruh suatu variabel terhadap variabel lain (Prasetyo, R.A., dan Helma, 2022). Analisis regresi diterapkan hampir pada seluruh aspek kehidupan, yakni sektor pertanian, sektor ekonomi dan keuangan, sektor industri dan tenaga kerja, sejarah, administrasi publik, serta ilmu lingkungan. Variabel dalam analisis regresi terbagi menjadi 2, yaitu variabel dependen dan variabel independen. Variabel dependen dalam analisis regresi ada yang bersifat dikotomis (biner) dengan nilai 0 atau 1. Analisis regresi yang cocok digunakan untuk variabel dependen yang bersifat biner adalah analisis regresi logistik. Menurut Peng, Lee, dan Ingersoll (2002), analisis regresi logistik lebih tepat digunakan ketika variabel dependen bersifat dikotomis (biner). Hal ini dikarenakan pada regresi linear mengasumsikan bahwa variabel dependen terdistribusi normal, sedangkan pada regresi logistik variabel dependen mengikuti distribusi *Bernoulli*. Selain itu, pada regresi linear probabilitas dari prediksi yang dihasilkan bisa berada diluar rentang 0 dan 1, yang tidak bermakna pada hasil variabel biner. Oleh karena itu, analisis regresi logistik adalah model yang cocok untuk data dengan variabel bersifat biner.

Menurut Hosmer dan Lemeshow (2000), regresi logistik adalah metode regresi yang bisa menjelaskan keterkaitan antara variabel dependen kategorik dengan satu atau lebih variabel independen. Regresi logistik juga bisa ditafsirkan sebagai pendekatan untuk membuat model prediksi. Regresi logistik pada statistik sering juga disebut sebagai model logistik atau model logit. Berdasarkan variabel

dependennya, terdapat beberapa jenis regresi logistik yaitu regresi logistik biner, multinomial, ordinal, serta multilevel. Regresi logistik biner merupakan metode dari regresi logistik yang sangat sederhana.

Regresi logistik biner yakni metode yang memiliki fungsi mencari keterkaitan antara variabel dependen (Y) yang memiliki karakteristik biner (*dichotomus*) dengan variabel independen (X). Regresi logistik biner banyak dipakai sebagai metode/alat analisis saat variabel dependen adalah biner, yang mengacu pada pemakaian dua angka, yaitu: 0 dan 1 sebagai pengganti kategori pada dua variabel dependen. Beberapa contoh variabel dependen tersebut adalah ya-tidak, benar-salah, pria-wanita, tinggi-rendah, hadir-absen, sakit-sehat, hidup-mati, dan lain sebagainya.

Indeks pembangunan manusia merupakan salah satu variabel dependen yang bersifat biner karena memiliki 2 kategori yaitu tinggi dan rendah. Indeks Pembangunan Manusia (IPM) adalah indikator komposit yang berguna untuk menghitung perolehan pembangunan pada aspek taraf hidup manusia. Pada tahun 1990, *United Nations Development Programme* (UNDP) memperkenalkan konsep pembangunan manusia yang menjelaskan bahwa manusia adalah kekayaan bangsa yang sesungguhnya (Badan Pusat Statistik, 2022). Indeks Pembangunan Manusia (IPM) berguna untuk memetakan apakah suatu negara masuk dalam kelas negara maju, negara berkembang, atau negara yang tertinggal serta juga untuk menilai dampak dari kebijakan ekonomi pada taraf hidup. Nilai IPM juga dapat menentukan peringkat atau level pembangunan suatu wilayah.

Berdasarkan penelitian yang telah dilakukan oleh Azagi, dkk. (2022), mengenai analisis faktor-faktor yang berpengaruh terhadap IPM di Pulau Jawa dengan menggunakan regresi logistik biner diperoleh nilai ketepatan klasifikasi sebesar 96,43%. Menurut penelitian Wulandari, dkk. (2021), tentang pemodelan literasi membaca siswa dengan regresi logistik biner diperoleh nilai ketepatan klasifikasi sebesar 81,8%.

Berdasarkan latar belakang yang telah dituliskan, maka analisis regresi logistik biner digunakan oleh penulis pada penelitian ini. Penelitian ini mengaplikasikan data indeks pembangunan manusia di Kalimantan dengan beberapa variabel independen yakni harapan lama sekolah (HLS), persentase penduduk miskin (PPM), laju pertumbuhan penduduk (LPP) dan persentase penduduk (PP).

## **1.2. Tujuan Penelitian**

Tujuan penelitian ini ialah sebagai berikut:

1. Memperoleh model regresi logistik biner terbaik terhadap faktor yang mempengaruhi indeks pembangunan manusia di Kalimantan pada tahun 2023.
2. Mengetahui ketepatan klasifikasi terhadap faktor yang mempengaruhi indeks pembangunan manusia di Kalimantan pada tahun 2023.

## **1.3. Manfaat Penelitian**

Manfaat penelitian ini ialah sebagai berikut:

1. Mengetahui model regresi logistik biner terbaik terhadap faktor yang mempengaruhi indeks pembangunan manusia di Kalimantan pada tahun 2023.
2. Mengetahui besarnya ketepatan klasifikasi terhadap faktor yang mempengaruhi indeks pembangunan manusia di Kalimantan pada tahun 2023.

## II. TINJAUAN PUSTAKA

### 2.1. Analisis Regresi Logistik Biner

Menurut Hosmer dan Lemeshow (2000), regresi logistik adalah alat regresi yang bisa menjelaskan keterkaitan antara variabel dependen kategorik dan satu atau lebih variabel independen. Terdapat beberapa model regresi logistik yakni regresi logistik ordinal, regresi logistik biner, serta regresi logistik multinomial. Regresi logistik biner ialah metode yang berfungsi untuk mencari keterkaitan antara variabel dependen (Y) yang memiliki sifat biner (*dichotomus*) dengan variabel independen (X) yang memiliki sifat kategorik atau kontinu. Variabel dependen (Y) pada regresi logistik biner mempunyai dua kategori yaitu kejadian “sukses” atau “gagal” misalnya: benar-salah, tinggi-rendah, pria-wanita, ya-tidak, sakit-sehat, hidup-mati, dan seterusnya. Variabel dependen (Y) disimbolkan sebagai  $Y=1$  (sukses) serta  $Y=0$  (gagal), sehingga mengikuti distribusi *Bernoulli* dengan fungsi probabilitas :

$$f(y) = \pi^y(1 - \pi)^{1-y} \quad y = 0,1 \quad (2.1)$$

dimana :

$\pi$  = peluang kejadian

$y$  = peubah acak yang bernilai 0 atau 1

Untuk  $y = 0$ , maka  $f(0) = \pi(x_i)^0(1 - \pi(x_i))^{1-0} = 1 - \pi(x_i)$

Untuk  $y = 1$ , maka  $f(1) = \pi(x_i)^1(1 - \pi(x_i))^{1-1} = \pi(x_i)$

Bentuk model regresi logistik dapat ditulis sebagai berikut :



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

dimana :

$\pi(x)$  = probabilitas sukses dengan nilai probabilitas  $0 \leq \pi(x) \leq 1$

$p$  = banyaknya variabel independen

$\beta_0$  = intersep model regresi logistik

$\beta_1, \dots, \beta_p$  = koefisien regresi yang mengukur pengaruh setiap variabel independen

Untuk memudahkan dalam memperkirakan parameter dari regresi logistik biner, maka  $\pi(x)$  pada persamaan diatas ditransformasikan dengan transformasi logit dari  $\pi(x)$  (Menezes, Liska, Cirillo, dan Vivanco, 2017). Transformasi logit dari  $\pi(x)$  dirumuskan sebagai berikut :

$$\begin{aligned} \pi(x) &= \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \\ (\pi(x))(1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}) &= e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ (\pi(x)) + (\pi(x)e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}) &= e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ (\pi(x)) &= e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} - \pi(x)e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ (\pi(x)) &= (1 - \pi(x))e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ \frac{\pi(x)}{1 - \pi(x)} &= e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) &= \ln e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \\ \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) &= \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \end{aligned}$$

Sehingga, persamaan yang didapatkan adalah :

$$g(\pi(x)) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2.3)$$

Fungsi  $g(\pi(x))$  adalah fungsi logit model regresi logistik biner dengan  $p$  variabel independen.

Menurut Abdulqader (2017), analisis regresi logistik biner memiliki beberapa asumsi, yaitu:

1. Hubungan linear antara variabel dependen dan variabel independen tidak dibutuhkan.
2. Pada regresi logistik biner variabel dependen harus memiliki 2 kategori atau bersifat dikotomi.
3. Antar variabel independen tidak terdapat multikolinearitas.
4. Variabel independen pada regresi logistik biner tidak membutuhkan pengujian normal multivariat.
5. Variabel independen tidak perlu berbentuk interval atau skala rasio (bentuk metrik).
6. Diperlukan sampel yang besar untuk variabel independen dengan jumlah minimum yaitu 50 sampel.
7. Tidak diperlukan keragaman varians antar kelompok variabel independen.
8. Tidak diperlukan asumsi homokedastisitas.

## 2.2. Multikolinearitas

Salah satu asumsi regresi logistik biner adalah antar variabel independen tidak boleh terdapat multikolinearitas. Multikolinieritas adalah suatu kondisi antar variabel independen terdapat hubungan atau korelasi yang linier. Model regresi logistik biner disebut baik apabila tidak terdapat multikolinearitas karena bisa mengakibatkan estimasi parameter pada model regresi yang diperoleh mempunyai *error* yang besar. Suatu model dikatakan tidak memiliki multikolinearitas apabila nilai pada *Variance Inflation Factor* (VIF) < 10. Jika nilai pada VIF > 10 menandakan adanya multikolinearitas antar variabel independen (Myers, 1990).

Untuk memeriksa asumsi multikolinearitas pada model regresi dapat dilihat dari nilai *Variance Inflation Factor* (VIF) untuk tiap-tiap variabel independen terhadap variabel dependen. Nilai VIF dapat dihitung menggunakan rumus berikut :

$$VIF = \frac{1}{(1-R_j^2)} \quad ,j = 1, 2, \dots, p \quad (2.4)$$

dengan  $R_j^2$  merupakan koefisien determinasi antar variabel independen  $X_j$ .

### 2.3. Maximum Likelihood Estimation (MLE)

*Maximum Likelihood Estimation* (MLE) ialah suatu pendugaan yang berfungsi untuk memperkirakan nilai parameter yang tidak diketahui. Estimasi parameter model regresi logistik biner dikerjakan dengan menggunakan *Maximum Likelihood Estimation* (MLE). Metode MLE berfungsi untuk memperoleh solusi maksimal dari fungsi *likelihood*, tetapi hanya mendapat sistem persamaan nonlinier (Kleinbaum & Klein, 2010).

Pada regresi logistik pengamatan berdistribusi Bernoulli dengan variabel  $x_i$  dan  $y_i$  merupakan pasangan variabel dependen dan variabel independen pada pengamatan ke- $i$  dengan  $i = 1, 2, \dots, n$ , sehingga fungsi probabilitasnya yaitu :

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

dengan

$$\pi(x_i) = \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})}$$

Saat  $j = 0$  maka  $x_{ij} = x_{i0} = 1$ . Tiap-tiap pasangan pengamatan dianggap independen. Oleh karena itu, fungsi *likelihood*-nya adalah kumpulan dari fungsi distribusi setiap pasangan. Menurut Hosmer & Lemeshow (2000), jika  $i$  adalah sampel pada sebuah populasi sehingga model fungsi *likelihood* untuk  $n$  sampel independen adalah :

$$\begin{aligned} l(\beta) &= \prod_{i=1}^n \pi(x_i)^{y_i}(1 - \pi(x_i))^{1-y_i} \quad (2.5) \\ &= \left\{ \prod_{i=1}^n (1 - \pi(x_i)) \right\} \left\{ \prod_{i=1}^n \exp \left( \ln \left( \frac{\pi(x_i)}{(1 - \pi(x_i))} \right)^{y_i} \right) \right\} \end{aligned}$$

$$\begin{aligned}
&= \left\{ \prod_{i=1}^n (1 - \pi(x_i)) \right\} \left\{ \exp \left( \sum_{i=1}^n y_i \ln \left( \frac{\pi(x_i)}{1 - \pi(x_i)} \right) \right) \right\} \\
&= \left\{ \prod_{i=1}^n \left( 1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right) \right)^{-1} \exp \left\{ \sum_{i=1}^n \left( \sum_{j=0}^p \beta_j x_{ij} y_i \right) \right\} \right\}
\end{aligned}$$

dimana :

$y_i$  = pengamatan pada variabel ke-i

$\pi(x_i)$  = peluang untuk variabel independen ke-i

$i = 1, 2, \dots, n$ .

Fungsi *likelihood* diatas kemudian dioptimalkan dalam bentuk  $\ln l(\beta)$  serta dirumuskan dengan  $L(\beta)$ , sehingga:

$$\begin{aligned}
l(\beta) &= \ln l(\beta) \\
&= \ln \left[ \left\{ \prod_{i=1}^n \left( 1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right) \right)^{-1} \right\} \exp \left\{ \sum_{i=1}^n \left( \sum_{j=0}^p \beta_j x_{ij} y_i \right) \right\} \right] \\
&= \left[ \ln \left( \exp \left\{ \sum_{i=1}^n \left( \sum_{j=0}^p \beta_j x_{ij} y_i \right) \right\} \right) \right] \left[ \ln \left( \prod_{i=1}^n \left( 1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right) \right)^{-1} \right) \right] \\
&= \left\{ \sum_{j=0}^p \beta_j \left( \sum_{i=1}^n x_{ij} y_i \right) \right\} - \sum_{i=1}^n \ln \left\{ 1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right) \right\}
\end{aligned}$$

Akan dicari nilai  $\beta$  dan  $l(\beta)$  dengan dilakukan turunan pertama  $l(\beta)$  terhadap  $\beta$  dan hasilnya sama dengan nol.

$$\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} \left[ \sum_{j=0}^p \beta_j \left( \sum_{i=1}^n x_{ij} y_i \right) \right] - \left[ \sum_{i=1}^n \ln \left\{ 1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right) \right\} \right] \\
&= \sum_{i=1}^n x_{ij} y_i - \left[ \sum_{i=1}^n x_{ij} \left\{ \frac{\exp \left( \sum_{j=0}^p \beta_j x_{ij} \right)}{1 + \exp \left( \sum_{j=0}^p \beta_j x_{ij} \right)} \right\} \right]
\end{aligned}$$

sehingga

$$\sum_{i=1}^n x_{ij} y_i - \sum_{i=1}^n x_{ij} \pi(x_i) = 0 \tag{2.6}$$

dengan  $j = 0, 1, \dots, p$ .



Karena persamaan (2.6) berbentuk implisit, maka persamaan tersebut memerlukan metode numerik untuk mengerjakannya. Digunakan metode numetik *Newton Raphson* dengan persamaan berikut (Agresti, 2007) :

$$\beta^{(r+1)} = \beta^{(r)} + (H^{(r)})^{-1} g^{(r)} \quad (2.7)$$

dimana:

$\beta^{(r+1)}$  = Nilai parameter  $\beta$  pada iterasi ke-  $r + 1$

$\beta^{(r)}$  = Nilai parameter  $\beta$  pada iterasi ke-  $r$

$(H^{(r)})^{-1}$  = Invers dari matriks *Hessian*

$g^{(r)}$  = Vektor kemiringan (*slope*)

Hasil dari turunan parsial yang pertama akan menghasilkan  $g^{(r)}$  yang disebut sebagai vektor kemiringan (*slope*).

$$g^{(r)} = \begin{bmatrix} \frac{\partial l(\beta)}{\partial \beta_0} \\ \frac{\partial l(\beta)}{\partial \beta_1} \\ \vdots \\ \frac{\partial l(\beta)}{\partial \beta_j} \end{bmatrix} = \begin{pmatrix} \sum_{i=1}^n x_{i0} y_i - \sum_{i=1}^n x_{i0} \pi(x_i) \\ \sum_{i=1}^n x_{i1} y_i - \sum_{i=1}^n x_{i1} \pi(x_i) \\ \vdots \\ \sum_{i=1}^n x_{ij} y_i - \sum_{i=1}^n x_{ij} \pi(x_i) \end{pmatrix}$$

Dalam bentuk matriks ditulis sebagai berikut:

$$\begin{aligned} g_p^{(r)} &= X^T Y - X^T \pi(x)^{(r)} \\ &= X^T (Y - \pi(x)^{(r)}) \end{aligned}$$

Selanjutnya, menghitung turunan parsial kedua dari persamaan *ln likelihood* terhadap parameter  $\beta$ .

$$\begin{aligned} \frac{\partial^2 l(\beta)}{\partial \beta_p^2} &= \frac{\partial}{\partial \beta_p} \left( - \sum_{i=1}^n x_{ij} \pi(x_i) \right) \\ &= \frac{\partial}{\partial \beta_p} \left( - \sum_{i=1}^n x_{ij} \left( \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})} \right) \right) \\ &= - \sum_{i=1}^n x_{ip}^2 \pi(x_i) (1 - \pi(x_i)) \end{aligned} \quad (2.8)$$

Kemudian, akan dihasilkan matriks *Hessian* dimana setiap elemennya merupakan hasil dari turunan parsial kedua dari fungsi  $l(\beta)$ .

$$H = \begin{bmatrix} \frac{\partial^2 l(\beta)}{\partial \beta_0^2} & \frac{\partial^2 l(\beta)}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 l(\beta)}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 l(\beta)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l(\beta)}{\partial \beta_1^2} & \cdots & \frac{\partial^2 l(\beta)}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\beta)}{\partial \beta_p \partial \beta_0} & \frac{\partial^2 l(\beta)}{\partial \beta_p \partial \beta_1} & \cdots & \frac{\partial^2 l(\beta)}{\partial \beta_p^2} \end{bmatrix}$$

Dalam bentuk matriks menjadi

$$H = X^T V^{(r)} X$$

Jadi, persamaan dari iterasi *Newton Raphson* dapat ditulis sebagai berikut:

$$\beta^{(r+1)} = \beta^{(r)} + (X^T V^{(r)} X)^{-1} X^T (Y - \pi(x)^{(r)})$$

Pada setiap iterasi ke-  $r$  berlaku:

$$H(\beta_p^{(r)}) = \frac{\partial^2 l(\beta)}{\partial \beta_p^2} = - \sum_{i=1}^n x_{ip} x_{ip} \pi(x_i)^{(r)} (1 - \pi(x_i)^{(r)})$$

$$g_p^{(r)} = \frac{\partial l(\beta)}{\partial \beta_j} = \sum_{i=1}^n x_{ij} y_i - \sum_{i=1}^n x_{ij} \pi(x_i)$$

$$\pi(x_i) = \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})}$$

Iterasi *Newton Raphson* dicari dengan langkah-langkah sebagai berikut:

- Menetapkan nilai dugaan awal  $\beta_{(0)}$
- Membuat vektor  $g^{(r)}$  dan matriks Hessian ( $H^{(r)}$ )
- Masukkan nilai  $\beta_0$  ke vektor  $g^{(r)}$  dan matriks Hessian ( $H^{(r)}$ ), sehingga diperoleh vektor  $g^{(r)}(\beta_0)$  dan matriks Hessian  $H^{(r)}(\beta_0)$
- Mulai dari iterasi  $r = 0$  dilakukan iterasi pada persamaan

$$\beta^{(r+1)} = \beta^{(r)} + (H^{(r)}(\beta_r))^{-1} g^{(r)}(\beta_r)$$

- e. Ulangi iterasi sampai iterasi ke-  $r = r + 1$  konvergen. Iterasi berhenti jika  $\beta^{(r+1)} \approx \beta^{(r)}$  atau  $\|\beta^{(r+1)} - \beta^{(r)}\| \leq \varepsilon$  dengan  $\varepsilon$  adalah bilangan yang sangat kecil

## 2.4. Uji Simultan

Uji simultan bertujuan melihat signifikansi koefisien  $\beta$  secara bersamaan atau keseluruhan terhadap variabel dependen (Hosmer dan Lemeshow, 2000). Uji simultan dilakukan dengan uji *Likelihood Ratio* dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_j \neq 0 \quad ; j = 1, 2, 3, \dots, p$$

Statistik Uji :

$$G = -2 \ln \left( \frac{\text{likelihood tanpa variabel independen}}{\text{likelihood dengan variabel independen}} \right)$$

$$G = -2 \ln \left[ \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1-\hat{\pi}_i)^{(1-y_i)}} \right] \quad (2.9)$$

Atau

$$G = 2 \{ \sum_{i=1}^n [y_i \ln(\hat{\pi}_i) + (1 - y_i) \ln(1 - \hat{\pi}_i)] - [n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)] \}$$

dimana :

$n_0$  = Jumlah pengamatan dengan kategori  $y = 0$

$n_1$  = Jumlah pengamatan dengan kategori  $y = 1$

$n$  = Jumlah pengamatan ( $n_0 + n_1$ )

$p$  = Banyaknya parameter

Daerah keputusan :

Tolak  $H_0$  jika  $G > X^2_{(p,\alpha)}$

Terima  $H_0$  jika  $G < X^2_{(p,\alpha)}$

Statistik uji G berdistribusi *Chi-Square* dengan taraf signifikansi sebesar  $\alpha$  serta p adalah derajat bebas yakni jumlah variabel independen dalam model.

Kesimpulan:

Tolak  $H_0$  artinya model memiliki variabel independen yang signifikan secara serentak atau bersamaan terhadap model.

## 2.5. Uji Parsial

Uji parsial bertujuan melihat signifikansi masing-masing parameter terhadap variabel dependen. Uji parsial dilakukan dengan uji Wald (Hosmer dan Lemeshow, 2000) dengan hipotesis sebagai berikut :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 \quad ; j = 1, 2, 3, \dots, p$$

Statistik Uji :

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.10)$$

Daerah keputusan :

Tolak  $H_0$  jika  $W > X^2_{(\alpha, p)}$

Terima  $H_0$  jika  $W < X^2_{(\alpha, p)}$

Statistik uji W adalah statistika uji Wald dengan  $SE(\hat{\beta}_j)$  merupakan estimasi standar *error* parameter yang berdistribusi *Chi-Square* dengan taraf signifikansi sebesar  $\alpha$  dan p merupakan derajat bebas.

Kesimpulan:

Tolak  $H_0$  maknanya variabel independen secara signifikan mempengaruhi variabel dependen.



## 2.6. Uji Kesesuaian Model

Uji kesesuaian model merupakan pengujian dengan tujuan melihat apakah model yang telah dibuat sesuai atau tidak. Model dikatakan sudah sesuai jika pada hasil pengamatan dan hasil model yang diprediksi tidak menunjukkan perbedaan (Hosmer dan Lemeshow, 2000). Hipotesis yang digunakan :

$H_0$  = Model sudah sesuai

$H_1$  = Model tidak sesuai

Uji kesesuaian model menggunakan statistik uji *Hosmer-Lemeshow Goodness of Fit* (Peeters, Dewil, dan Smets, 2012) :

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.11)$$

dengan  $\bar{\pi}_k$  dan  $O_k$  sebagai berikut :

$$\bar{\pi}_k = \frac{1}{n'_k} \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n'_k}$$

$m_j$  merupakan jumlah subjek pada  $c_k$  kombinasi variabel independen.

$$O_k = \sum_{j=1}^{c_k} y_j$$

dimana :

$O_k$  = Observasi pada grup ke-k

$\bar{\pi}_k$  = rata-rata taksiran peluang

$n'_k$  = banyaknya observasi pada grup ke-k

$g$  = banyaknya kategori semua variabel independen

Daerah keputusan :

Tolak  $H_0$  jika  $\hat{C} > X^2_{(g-2, \alpha)}$  atau jika P-value  $< \alpha$

Terima  $H_0$  jika  $\hat{C} < X^2_{(g-2, \alpha)}$  atau jika P-value  $> \alpha$

Statistik uji  $\hat{C}$  berdistribusi *Chi-Square* dengan taraf signifikansi sebesar  $\alpha$  dan  $g-2$  merupakan derajat bebas.

Kesimpulan :

Tolak  $H_0$  artinya model yang terbentuk sudah sesuai.

## 2.7. Interpretasi Koefisien Parameter

Interpretasi koefisien parameter digunakan untuk mengetahui apakah terdapat relasi fungsional pada variabel dependen dan variabel independen serta memaparkan bagaimana variabel independen mempengaruhi setiap perubahan pada variabel dependen (Hosmer dan Lemeshow, 2000). Model dalam regresi logistik terdapat dua nilai  $\pi(x)$  dan dua nilai  $1 - \pi(x)$ . Berikut merupakan tabel dari koefisien parameter.

Tabel 1. Koefisien Parameter

Variabel Dependen (Y)	Variabel Independen (X)	
	X = 1	X = 0
Y = 1	$\pi(1) = \frac{\exp(\beta_0 + \beta_j)}{1 + \exp(\beta_0 + \beta_j)}$	$\pi(0) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$
Y = 0	$1 - \pi(1) = \frac{1}{1 + \exp(\beta_0 + \beta_j)}$	$1 - \pi(0) = \frac{1}{1 + \exp(\beta_0)}$

Rasio kecenderungan (*odds ratio*) merupakan suatu bilangan yang menggambarkan perbandingan antara individu yang mengalami kejadian tertentu dengan dengan individu yang tidak, baik dalam sampel maupun pada tingkat populasi. *Odds ratio* digunakan dalam interpretasi koefisien parameter regresi dengan persamaan sebagai berikut :

$$\begin{aligned}
 OR &= \left( \frac{\frac{\pi(1)}{(1 - \pi(1))}}{\frac{\pi(0)}{(1 - \pi(0))}} \right) \\
 &= \left( \frac{\frac{\exp(\beta_0 + \beta_j)}{1 + \exp(\beta_0 + \beta_j)}}{\frac{1}{1 + \exp(\beta_0 + \beta_j)}} \right) \left( \frac{\frac{1}{1 + \exp(\beta_0)}}{\frac{\exp(\beta_0)}{1 + \exp(\beta_0)}} \right)
 \end{aligned}$$

$$\begin{aligned}
&= \left( \frac{\exp(\beta_0 + \beta_j)}{\exp(\beta_0)} \right) \\
&= \exp(\beta_j)
\end{aligned}
\tag{2.12}$$

Interpretasi dari OR pada model regresi logistik berdasarkan nilai dari  $\exp(\beta_j)$ . Artinya, apabila variabel independen X naik 1 unit perubahan, maka nilai variabel dependen Y akan naik sebesar  $\exp(\beta_j)$  kali, sedangkan variabel lainnya tetap (Srimaneekarn, Hayter, Liu, dan Tantipoj, 2022).

## 2.8. Ketepatan Klasifikasi

Ketepatan klasifikasi model berfungsi melihat apakah pengklasifikasian data yang dihasilkan sudah tepat atau belum (Agresti, 2007). Menurut Sari, Susilawati, dan Srinadi (2016), untuk menentukan ketepatan klasifikasi nilai yang digunakan adalah nilai *Apparent Error Rate* (APER). APER menjelaskan nilai yang dikategorikan dengan salah oleh fungsi klasifikasi. Nilai APER dapat dilihat pada tabel ketepatan klasifikasi sebagai berikut :

Tabel 2. Perhitungan Ketepatan Klasifikasi

Aktual		Prediksi		Total
		Y		
		0	1	
Y	0	$n_{11}$	$n_{12}$	$n_h$
	1	$n_{21}$	$n_{22}$	$n_i$
Total		$n_j$	$n_k$	$n$

dimana :

$n_{11}$  = Jumlah pengamatan dari Y(0) tepat diklasifikasikan sebagai Y(0),

$n_{12}$  = Jumlah pengamatan dari Y(0) tepat diklasifikasikan sebagai Y(1),

$n_{21}$  = Jumlah pengamatan dari Y(1) tepat diklasifikasikan sebagai Y(0),

$n_{22}$  = Jumlah pengamatan dari Y(1) tepat diklasifikasikan sebagai Y(1),

$n_h$  = Jumlah pengamatan aktual dari Y(0),

$n_i$  = Jumlah pengamatan aktual dari Y(1),

$n_j$  = Jumlah pengamatan prediksi dari Y(0),

$n_k$  = Jumlah pengamatan prediksi dari Y(1),

$n$  = Total pengamatan keseluruhan.

Menurut Ramandhani, Sudarno, dan Safitri (2017), perhitungan nilai APER menunjukkan pengamatan yang diprediksi dengan tidak tepat dengan fungsi klasifikasi menggunakan rumus sebagai berikut :

$$APER = \frac{n_{12} + n_{21}}{n} \times 100\% \quad (2.13)$$

Ketepatan klasifikasi = 100% - APER (%)

## 2.9. Indeks Pembangunan Manusia

*United Nations Development Programme* (UNDP) mengemukakan gagasan tentang pembangunan manusia pada tahun 1990 dalam laporan berjudul *Human Development Report* (HDR) yang menjelaskan bahwa manusia adalah kekayaan bangsa yang sesungguhnya. Bukan hanya input dari pembangunan, pembangunan manusia menempatkan manusia sebagai tujuan akhir pembangunan (Badan Pusat Statistik, 2022).

Indeks Pembangunan Manusia (IPM) berguna untuk memetakan apakah suatu negara masuk dalam kelas negara maju, negara berkembang, atau negara yang

tertinggal serta untuk menilai dampak dari kebijakan ekonomi pada kualitas hidup. Informasi data IPM dipaparkan secara rutin tiap tahun di tingkat nasional, provinsi, serta kabupaten/kota sejak tahun 2014. Pemaparan IPM secara berkala berdasarkan wilayah memberikan kesempatan bagi tiap provinsi serta kabupaten/kota memahami gambaran perkembangan pembangunan manusia di wilayahnya, baik capaian, laju, posisi, maupun kesenjangan atau ketimpangan antar wilayah.



### **III. METODOLOGI PENELITIAN**

#### **3.1. Waktu dan Tempat Penelitian**

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2024/2025 di jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### **3.2. Data Penelitian**

Pada penelitian ini, digunakan data sekunder yang didapat dari Badan Pusat Statistik. Data yang diambil mengenai Indeks Pembangunan Manusia (IPM) di Kalimantan tahun 2023 beserta faktor-faktor yang diduga berpengaruh. Berdasarkan kategori yang diperoleh dari BPS, IPM di kabupaten/kota  $< 70$  dikategorikan sebagai daerah dengan IPM sedang dan rendah dan dilambangkan dengan 0, sedangkan IPM di kabupaten/kota  $\geq 70$  dikategorikan sebagai daerah dengan IPM tinggi dan sangat tinggi dan dilambangkan dengan 1. Variabel-variabel bebas yang diduga memiliki pengaruh terhadap IPM adalah harapan lama sekolah (HLS) (Tahun) ( $X_1$ ), persentase penduduk miskin (PPM) (Persen) ( $X_2$ ), laju pertumbuhan penduduk (LPP) ( $X_3$ ) dan Persentase Penduduk (PP) (Persen) ( $X_4$ ).

Tabel 3. Data Penelitian

No.	Prov/Kab/Kota	Y		X1	X2	X3	X4
		Indeks Pembangunan Manusia (IPM)		Harapan Lama Sekolah (HLS) (Tahun)	Persentase Penduduk Miskin (PPM) Tahun	Laju Pertumbuhan Penduduk (LPP)	Persentase Penduduk (PP) (Persen)
1	Kalimantan Barat	69,41	0	12.67	6,71	1,38	100
2	Sambas	68.69	0	12.72	7,08	1,1	11.53
3	Bengkayang	69.53	0	12.16	6,28	0,98	5.32
4	Landak	68.14	0	12.53	9,97	1,32	7.35
5	Mempawah	67.92	0	12.88	5,21	0,52	5.56
6	Sanggau	67.77	0	11.87	4,79	0,55	8.96
7	Ketapang	68.68	0	11.96	9,25	1,84	10.54
8	Sintang	68.67	0	12.3	8,18	1,16	7.79
9	Kapuas Hulu	67.67	0	12.2	8,16	1,51	4.67
10	Sekadau	66.33	0	11.92	5,9	0,91	3.91
11	Melawi	67.76	0	11.34	11,12	2,47	4.26
12	Kayong Utara	64.79	0	12.12	9,13	1,71	2.36
13	Kubu Raya	69.6	0	13.89	4,23	0,56	11.37
14	Kota Pontianak	81.03	1	15.04	4,45	0,49	12.01
15	Kota Singkawang	73.39	1	12.94	4,7	0,74	4.38
16	Kalimantan Tengah	72,2	1	12,76	5,11	0,66	100
17	Kotawaringin Barat	74.04	1	12.77	4,18	0,55	10.09
18	Kotawaringin Timur	72.21	1	12.88	5,69	0,75	15.98
19	Kapuas	70.75	1	12.96	5,21	0,73	15.33
20	Barito Selatan	71.65	1	12.8	4,72	0,41	4.86
21	Barito Utara	71.84	1	12.54	5,35	0,57	5.79
22	Sukamara	69.44	0	12.16	3,96	0,29	2.4
23	Lamandau	71.81	1	12.52	3,12	0,33	3.68
24	Seruyan	68.68	0	12.03	7,12	0,6	6.18
25	Katingan	70.28	1	12.98	4,99	0,68	6.11
26	Pulang Pisau	69.6	0	12.51	4,58	0,53	4.98
27	Gunung Mas	72.09	1	12.05	5,47	0,49	5.15
28	Barito Timur	72.7	1	12.87	6,63	0,75	4.23
29	Murung Raya	69.42	0	12.04	6,44	0,9	4.2
30	Kota Palangka Raya	81.68	1	14.99	3,44	0,27	11.02
31	Kalimantan Selatan	72,5	1	12.86	4,29	0,61	100

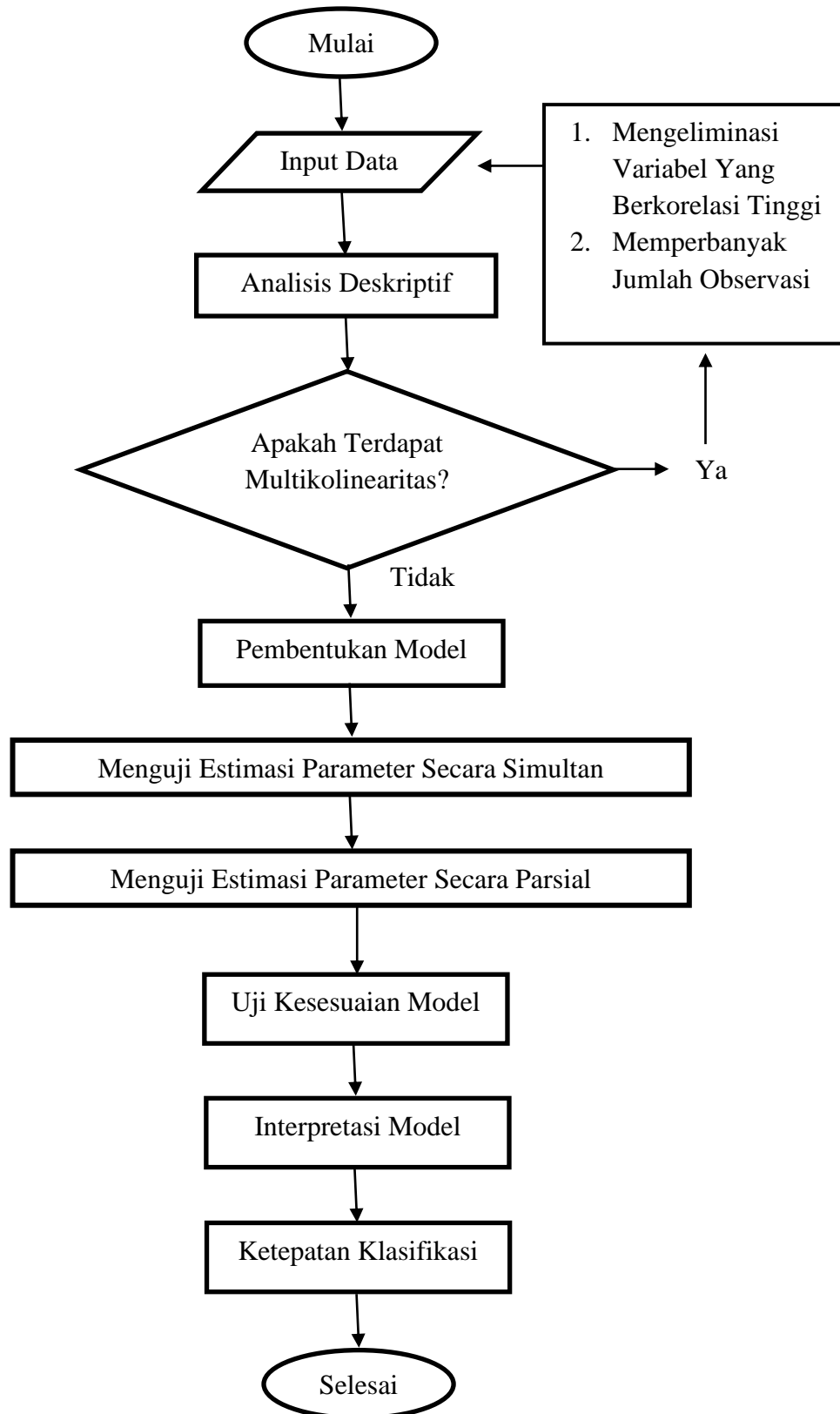
Tabel 3. Data Lanjutan

No.	Prov/Kab/Kota	Y		X1	X2	X3	X4
		Indeks Pembangunan Manusia (IPM)		Harapan Lama Sekolah (HLS) (Tahun)	Persentase Penduduk Miskin (PPM) Tahun)	Laju Pertumbuhan Penduduk (LPP)	Persentase Penduduk (PP) (Persen)
32	Tanah Laut	70.98	1	12.41	3,73	0,3	8.55
33	Kotabaru	70.18	1	12.07	4,32	0,92	8.03
34	Banjar	71.46	1	13.03	2,44	0,38	14.01
35	Barito Kuala	68.11	0	12.55	4,6	1,14	7.73
36	Tapin	72	1	12.33	3,19	0,31	4.65
37	Hulu Sungai Selatan	70.5	1	12.46	4,01	0,69	5.59
38	Hulu Sungai Tengah	70.13	1	12.24	5,84	0,63	6.3
39	Hulu Sungai Utara	67.66	0	13.19	6,25	0,63	5.55
40	Tabalong	73.85	1	12.91	5,77	0,46	6.24
41	Tanah Bumbu	72.31	1	12.6	4,12	0,42	7.99
42	Balangan	70.49	1	12.82	5,22	0,55	3.22
43	Kota Banjarmasin	78.5	1	13.97	4,63	0,55	15.78
44	Kota Banjar Baru	80.11	1	14.85	3,92	0,62	6.35
45	Kalimantan Timur	78,2	1	14.02	6,11	0,77	100
46	Paser	74.14	1	13.37	9,11	1,04	7.27
47	Kutai Barat	73.72	1	13.25	9,72	1,29	4.53
48	Kutai Kartanegara	75.3	1	13.64	7,61	0,91	19.36
49	Kutai Timur	74.98	1	13.01	9,06	1,2	11.65
50	Berau	76.21	1	13.36	5,54	0,16	6.61
51	Penajam Paser Utara	73.3	1	12.86	6,97	0,75	5.05
52	Mahakam Ulu	69.59	0	12.63	11,38	1,02	0.86
53	Kota Balikpapan	81.66	1	14.24	2,31	0,25	18.16
54	Kota Samarinda	82.32	1	15.39	4,81	0,75	21.76
55	Kota Bontang	81.56	1	13.43	4,11	0,39	4.75
56	Kalimantan Utara	72,49	1	13.2	6,45	0,64	100
57	Malinau	73.19	1	13.34	6,54	0,37	11.72
58	Bulungan	72.72	1	13.04	8,99	0,84	21.65
59	Tana Tidung	69.2	0	12.39	4,62	0,14	3.76
60	Nunukan	67.7	0	12.68	5,53	0,59	28.48
61	Kota Tarakan	77.53	1	14.21	6,1	0,79	34.39

### 3.3. Metode Penelitian

Dalam melakukan penelitian ini, penulis menggunakan *software Rstudio* 4.4.1. dalam proses perhitungannya. Sesuai dengan tujuan yang ingin dicapai, tahapan yang dilaksanakan pada penelitian ini ialah sebagai berikut :

1. Merumuskan masalah dan tujuan penelitian.
2. Mengumpulkan data Indeks Pembangunan Manusia di Kalimantan tahun 2023.
3. Mendeskripsikan karakteristik data dengan statistika deskriptif untuk variabel harapan lama sekolah (HLS) (Tahun) ( $X_1$ ), persentase penduduk miskin (PPM) (Persen) ( $X_2$ ), laju pertumbuhan penduduk (LPP) ( $X_3$ ) dan Persentase Penduduk (PP) (Persen) ( $X_4$ ).
4. Melakukan pengujian multikolinearitas dengan nilai *Variance Inflation Factor* (VIF).
5. Membentuk model awal regresi logistik biner dengan menggunakan penduga parameter *Maximum Likelihood*.
6. Melakukan uji signifikansi parameter regresi logistik biner secara simultan dengan uji *Likelihood Ratio* untuk mengetahui apakah variabel independen mempengaruhi variabel dependen secara serentak.
7. Melakukan uji signifikansi parameter regresi logistik biner secara parsial dengan uji *Wald* untuk mengetahui apakah variabel independen mempengaruhi variabel dependen.
8. Melakukan uji kesesuaian model regresi logistik biner dengan *Hosmer-Lemeshow Goodness of Fit* yang bertujuan mengetahui apakah model yang telah dibentuk sesuai atau tidak.
9. Menginterpretasi pemodelan regresi logistik biner dengan *odds ratio* untuk menjelaskan setiap perubahan variabel dependen yang disebabkan oleh variabel independen.
10. Menghitung ketepatan klasifikasi pada model yang dihasilkan dengan nilai *Apparent Error Rate* (APER) untuk mengetahui apakah data diklasifikasikan dengan benar atau tidak.



Gambar 1. Diagram Alir Analisis Regresi Logistik Biner

## V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, didapatkan beberapa kesimpulan sebagai berikut:

Faktor yang berpengaruh terhadap indeks pembangunan manusia di Kalimantan pada tahun 2023 secara signifikan, yaitu variabel harapan lama sekolah (HLS) ( $X_1$ ) dan persentase penduduk miskin (PPM) ( $X_2$ ). Sehingga model regresi logistik biner indeks pembangunan manusia di Kalimantan pada tahun 2023 adalah sebagai berikut:

$$\pi(x) = \frac{\exp(-27,9688 + 2,4074X_1 - 0,3472X_2)}{1 + \exp(-27,9688 + 2,4074X_1 - 0,3472X_2)}$$

Probabilitas Indeks Pembangunan Manusia (IPM) untuk kategori 1 (tinggi dan sangat tinggi) di Kalimantan Barat: 0,5507; Kalimantan Tengah: 0,7262; Kalimantan Selatan: 0,8012; Kalimantan Timur: 0,9749; dan Kalimantan Utara: 0,8276. Variabel harapan lama sekolah (HLS) ( $X_1$ ) sangat mempengaruhi probabilitas ini, karena pada saat  $X_1 = 0$ , nilai probabilitas Indeks Pembangunan Manusia (IPM) untuk kategori 1 (tinggi dan sangat tinggi) pada 5 provinsi di Kalimantan mendekati 0. Sedangkan pada saat variabel  $X_2 = 0$ , probabilitas untuk masing-masing provinsi adalah: Kalimantan Barat: 0,9264; Kalimantan Tengah: 0,9398; Kalimantan Selatan: 0,9521; Kalimantan Timur: 0,9969; dan Kalimantan Utara: 0,9783.

Analisis regresi logistik biner berhasil digunakan untuk memodelkan indeks pembangunan manusia di Kalimantan pada tahun 2023 dengan nilai ketepatan klasifikasi model sebesar 80,36%.

## DAFTAR PUSTAKA

- Abdulqader, Q.M., 2017. Applying The Binary Logistic Regression Analysis On The Medical Data. *Science Journal Of University Of Zakho*. **5**(4): 330-334.
- Agresti. 2007. *Categorical Data Analysis*. 2<sup>nd</sup> Edition. John Willey & Sons, New York.
- Azagi, A.I., 2022. Pemodelan Regresi Logistik Biner Pada Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia di Pulau Jawa. *Jurnal Statistika dan Aplikasinya*. **6**(1): 2620-8369.
- Badan Pusat Statistik. 2022. *Indeks Pembangunan Manusia 2021*. Badan Pusat Statistik, Jakarta.
- Hosmer, D.W., dan Lemeshow. 2000. *Applied Logistic Regression*. USA: John Willey and Sons.
- Kleinbaum, D.G. dan Klein, M. 2010. *Logistic Regression A Self-Learning Text*. 3<sup>rd</sup> Edition. Springer, New York.
- Menezes, F.S., Liska, G.R., Cirillo, M.A., dan Vivanco, M.J.F. 2017. Data Classification With Binary Response Through The Boosting Algorithm And Logistic Regression. *Expert System With Application*. **69**: 62-73.
- Myers, R.H. 1990. *Classical and Modern Regression with Application*. PWS-KENT Publishing Company, Boston.

- Peeters, B., Dewil, R., dan Smets, I.Y. 2012. Improved Process Control Of An Industrial Sludge Centrifuge-Dryer Installation Through Binary Logistic Regression Modeling Of The Fouling Issues. *Journal Of Process Control*. **22**: 1387-1396.
- Peng, C.J., Lee, K.L., dan Ingersoll, G.M. 2002. An Introducing To Logistic Regression Analysis And Reporting. The *Journal Of Educational Research*. **96**(1): 3-14.
- Prasetyo, R.A., dan Helma. 2022. Analisis Regresi Linear Berganda Untuk Melihat Faktor Yang Berpengaruh Terhadap Kemiskinan Di Provinsi Sumatera Barat. *Journal Of Mathematics UNP*. **7**(2): 62-68.
- Ramandhani, R., Sudarno, dan Safitri, D. 2017. Metode Bootstrap Aggregating Regresi Logistik Biner Untuk Ketepatan Klasifikasi Kesejahteraan Rumah Tangga Di Kota Pati. *Journal Gaussian*. **6**(1): 121-124.
- Sari, P.P., Susilawati, M., dan Srinadi, I.G. 2016. Bootsrap Aggregating (BAGGING) Regresi Logistik Ordinal Untuk Mengklasifikasikan Status Gizi Balita di Kabupaten Klungkung. *E-Jurnal Matematika*. **5**(3): 103.
- Srimaneekarn, N., Hayter, A., Liu, W., dan Tantipoj, C. 2022. Binary Response Analysis Using Logistic Regression In Dentistry. *International Journal Of Dentistry*. **22**.
- Wulandari, S.P.,dkk. 2021. Pemodelan Literasi Membaca Siswa di Daerah Terpencil Menggunakan Regresi Logistik Biner. *Jurnal Leverage, Engagement, Empowerment of Community*. **3**(1): 33-42.