

ANALISIS REGRESI LOGISTIK BINER UNTUK MELIHAT FAKTOR-FAKTOR YANG MEMPENGARUHI PARAMETER INDEKS STANDAR PENCEMAR UDARA

BINARY LOGISTIC REGRESSION ANALYSIS TO EXAMINE FACTORS AFFECTING AIR POLLUTION STANDARD INDEX PARAMETERS

Bela Ikhwansyah¹, Lilis Harianti Hasibuan^{2§}, Ilham Dangu Rianjaya³

¹Program Studi Matematika, Universitas Islam Negeri Imam Bonjol Padang [Email: ikhwansyahbela@gmail.com]

^{2§}Program Studi Matematika, Universitas Islam Negeri Imam Bonjol Padang [Email: lilisharianti@uinib.ac.id]

³Program Studi Matematika, Universitas Islam Negeri Imam Bonjol Padang [Email: ilham.rianjaya@uinib.ac.id]

[§]*Corresponding Author*

Received 10th Dec 2025; Accepted 11th Dec 2025; Published 17th Dec 2025;

Abstrak

Pencemaran udara adalah masalah lingkungan yang serius dan dapat berdampak pada kesehatan manusia serta kelestarian lingkungan. Nilai Indeks Standar Pencemar Udara (ISPU) digunakan untuk penentuan kualitas udara berdasarkan beberapa parameter. Penelitian ini bertujuan untuk mengetahui faktor-faktor meteorologi yang mempengaruhi parameter ISPU di Kota Padang dengan menggunakan metode regresi logistik biner. Data yang dianalisis adalah data tahun 2022 yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika Stasiun Meteorologi Kelas II Minangkabau Padang Pariaman dan Dinas Lingkungan Hidup Kota Padang. Dalam penelitian ini, variabel dependennya berupa kategori ISPU (Baik dan Sedang), sedangkan variabel independen yang diduga berpengaruh pada parameter ISPU adalah kecepatan angin, arah angin, suhu udara, kelembaban udara, curah hujan dan tekanan udara. Hasil penelitian ini menunjukkan bahwa kecepatan angin (X_1), suhu udara (X_3), curah hujan (X_5), dan tekanan udara (X_6) memiliki pengaruh yang signifikan terhadap parameter ISPU. Model regresi logistik biner yang dihasilkan memiliki ketepatan klasifikasi sebesar 75,0685%, yang menunjukkan bahwa model ini cukup baik dalam mengklasifikasikan kategori ISPU.

Kata kunci: Ketepatan Klasifikasi, Parameter Indeks Standar Pencemar Udara, Regresi Logistik Biner

Abstract

Air pollution is a serious environmental problem that can impact human health and environmental sustainability. The Air Pollutant Standard Index (APSI) value is used to determine air quality based on several parameters. This study aims to determine the meteorological factors that affect the APSI parameters in Padang city using the binary logistic regression method. The data analysed is the 2022 data obtained from the Meteorology, Climatology and Geophysics Agency, Class II Minangkabau Padang Pariaman Meteorological Station and the Padang City Environment Office. In this study, the dependent variable is the category of APSI (Good and Moderate), while the independent variables that are thought to affect APSI parameters are wind speed, wind direction, air temperature, air humidity, rainfall and air pressure. The results of this study indicate that wind speed (X_1), air temperature (X_3),

rainfall (X_5), and air pressure (X_6) have a significant influence on the APSI parameters. The resulting binary logistic regression model has a classification accuracy of 75.0685%, which indicates that this model is quite good at classifying the APSI category.

Keywords: *Air Pollution Standard Index Parameters, Binary Logistic Regression, Classification Accuracy*

1. Pendahuluan

Isu kualitas udara selalu relevan dengan perkembangan kota-kota besar di dunia saat ini. Demikian pula kota-kota besar di Indonesia juga mengalami pencemaran udara [1]. Contohnya, di Kota Padang indeks kualitas udara pada Tahun 2023-2024 berkisar 73-74 $\mu\text{g}/\text{m}^3$ artinya angka kualitas udara tersebut termasuk kategori sedang (BPS Sumatra Barat, 2024). Sementara itu, pencemaran udara terjadi akibat perkembangan teknologi, peningkatan jumlah kendaraan pribadi dan pembakaran bahan bakar fosil saat ini sehingga turut mempengaruhi kualitas udara [3].

Dampak dari pencemaran udara yang dapat dirasakan berupa terganggunya kesehatan, rusaknya ekosistem, dan juga terjadinya hujan asam. Pencemaran udara terjadi akibat faktor-faktor meteorologi maupun buatan berupa aktivitas makhluk hidup yang menyebabkan terkontaminasinya atmosfer bumi [3]. Sesuai dengan Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor P.14/MENLHK/SETJEN/KUM.1/7/2020, mengenai Indeks Standar Pencemar Udara Pasal 5 Ayat 3 menyatakan bahwa data meteorologi yang berpengaruh pada konsentrasi udara lingkungan mencakup kecepatan dan arah angin, suhu udara, kelembaban udara, intensitas sinar matahari, serta curah hujan

[4].

Standar kualitas udara lingkungan

dinyatakan dalam Indeks Standar Pencemar Udara (ISPU). Hal ini didasarkan pada dampak terhadap kesehatan manusia di tempat dan waktu tertentu. ISPU ini dijadikan acuan dalam melaksanakan langkah-langkah pengendalian pencemaran udara. Menurut Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor P.14/MENLHK/SETJEN/KUM.1/7/2020, mengenai Indeks Standar Pencemar Udara Pasal 2 Ayat 2 mengemukakan bahwa parameter yang dipakai untuk menetapkan ISPU adalah *Particulate Matter* (PM_{10}), *Particulate Matter* ($PM_{2.5}$), Karbon Monoksida (CO), Nitrogen Dioksida (NO_2), Sulfur Dioksida (SO_2), Ozon (O_3), Hidrokarbon (HC) [4].

Kualitas udara dipantau secara harian melalui stasiun pemantauan kualitas udara lingkungan dan dinilai berdasarkan parameter-parameter ISPU [5]. Kategori parameter ISPU diberi nilai 1 apabila konsentrasi parameter ISPU berada pada kategori Baik dengan angka rentang 1-50. Namun apabila diberi nilai 0 maka konsentrasi parameter ISPU diudara berada pada kategori Sedang dengan angka rentang 51-100. Dengan demikian, data kualitas udara yang diperoleh berdasarkan kategori ISPU dapat dikategorikan sebagai data biner.

Analisis regresi logistik biner merupakan analisis untuk mengevaluasi hubungan antara satu variabel dependen yang bersifat biner

(dikotomus) dengan satu atau lebih variabel independen [6]. Faktor-faktor meteorologi yang mempengaruhi parameter ISPU akan dianalisis menggunakan metode regresi logistik biner untuk menentukan apakah terdapat pengaruh yang signifikan dalam memprediksi hubungan antara faktor-faktor meteorologi (variabel independen) dan parameter ISPU (variabel dependen).

2. Landasan Teori

Analisis Regresi

Analisis regresi merupakan suatu metode statistik yang digunakan untuk menjelaskan hubungan matematis antara dua jenis variabel, yaitu variabel dependen (Y) yang dipengaruhi dan variabel independen (X) yang mempengaruhi [7],[8],[9]. Metode ini bertujuan untuk mempelajari keterkaitan atau pola hubungan antar variabel, yang kemudian dapat dinyatakan dalam bentuk persamaan. Jika hubungan tersebut hanya melibatkan satu variabel dependen (Y) dan satu variabel independen (X), itu dikategorikan sebagai regresi sederhana [10], [11], [12], [13]. Sebaliknya, jika melibatkan satu variabel dependen dan lebih dari satu variabel independen, istilah yang digunakan adalah regresi berganda [14], [15]. Secara umum persamaan dalam analisis regresi dapat dituliskan sebagai berikut [16], [17]:

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad (1)$$

di mana:

- Y : variabel dependen
- β_0 : nilai konstanta
- β_1 : parameter model
- X : variabel independen

ε : galat

Analisis Regresi Logistik Biner

Analisis regresi logistik biner merupakan suatu metode statistik yang digunakan untuk menjelaskan atau melihat hubungan antara variabel dependen (Y) yang bersifat biner/dikotomus dengan satu atau lebih variabel independen (X) yang dapat berupa kategori atau numerik. Variabel dependen memiliki dua kategori, misalnya $Y = 1$ untuk "sukses" dan $Y = 0$ untuk "gagal", di mana setiap observasi mengikuti distribusi Bernoulli. Fungsi distribusi peluangnya dirumuskan sebagai berikut [16]:

$$f(y_i; \pi_i) = \begin{cases} \pi_i^{y_i} (1 - \pi_i)^{1-y_i}, & \text{untuk } y_i = 0, 1 \\ 0, & \text{untuk } y_i \text{ yang lain} \end{cases}$$

Keterangan :

π_i : peluang sukses

$1 - \pi_i$: peluang gagal

n : banyaknya pengulangan

y_i : banyaknya sukses dalam n kali pengulangan

diperoleh :

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

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

maka probabilitas yang digunakan $P(Y = 1 | X) = \pi(x_i)$ dan $P(Y = 0 | X) = 1 - \pi(x_i)$. dengan $E(Y | X) = \pi$, dan $0 \leq \pi \leq 1$.

untuk regresi logistik ganda, yang melibatkan lebih dari satu variabel independen, Fungsi peluang model persamaan regresi logistik biner ganda dapat dinyatakan dalam persamaan berikut

[18]:

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}, \quad (2)$$

Model regresi logistik berganda setelah ditransformasi dalam Persamaan 2 menggunakan bentuk logit diperoleh:

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n), \quad (3)$$

dengan $g(x) = (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)$ disebut dengan fungsi logit model regresi logistik biner ganda karena model ini memiliki n buah variabel independen.

Uji Multikolinearitas

Dalam analisis regresi logistik biner, sangat penting untuk memiliki data yang bebas dari masalah multikolinearitas. Untuk menentukan apakah terdapat multikolinearitas antara variabel independen, kita dapat menggunakan nilai *Variance Inflation Factor (VIF)* dengan rumus sebagai berikut [19]:

$$VIF_i = \frac{1}{1 - R_i^2}, \quad (4)$$

di mana R_i^2 merupakan koefisien determinasi

kriteria pengujiannya adalah jika nilai *VIF* < 10 maka dapat disimpulkan bahwa tidak terjadi multikolinearitas antar variabel independen [20].

Pendugaan Parameter

Parameter yang belum diketahui dapat dihitung menggunakan metode *Maximum Likelihood Estimation (MLE)*. Pada dasarnya metode ini mencari nilai estimasi β yang membuat fungsi *likelihood* bernilai maksimum.

Secara lebih terstruktur, fungsi *likelihood* pada model regresi logistik biner dapat dituliskan sebagai berikut [21]:

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}, \quad (5)$$

Da

agar perhitungan menjadi lebih sederhana digunakan metode pendekatan *log likelihood* yang didefinisikan sebagai berikut:

$$l(\beta) = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\}, \quad (6)$$

estimasi koefisien regresi logistik $\tilde{\beta}$ diperoleh dengan menurunkan persamaan $l(\beta)$ terhadap β kemudian hasil turunan tersebut dibuat sama dengan 0.

Uji Simultan

Uji simultan yang sering disebut sebagai uji serentak, bertujuan untuk menguji apakah semua variabel independen secara bersama-sama berpengaruh signifikan terhadap parameter β . Dalam proses ini pengujian dilakukan menggunakan statistik uji *G* atau *Likelihood Ratio Test* dengan hipotesis yang telah ditentukan sebagai berikut [22]:

$H_0: \beta_1 = \beta_2 = \beta \dots = \beta_k = 0$ (tidak terdapat pengaruh).

H_1 : paling sedikit ada satu $\beta_j \neq 0$ ($j = 1, 2, \dots, k$) (ada pengaruh minimal satu).

dengan statistik uji *G*:

$$G = -2 \ln \left[\frac{l_0}{l_k} \right], \quad (7)$$

di mana l_0 adalah nilai *likelihood* Ketika model tidak menggunakan variabel independen dan l_k adalah *likelihood* dengan seluruh variabel independen. Nilai *G* mengikuti distribusi *Chi*-

Square dengan k derajat kebebasan. Keputusan menolak H_0 diambil jika $G > X^2_{(db,\alpha)}$ atau nilai $p_{value} < \alpha$. Penolakan H_0 mengindikasikan bahwa variabel independen secara kolektif mempengaruhi variabel dependen.

Uji Parsial

Uji parsial dilaksanakan guna menguji peranan sejauh mana masing-masing variabel independen berpengaruh dalam model. Pengujian ini dilakukan secara terpisah dengan menggunakan uji Wald (Hosmer & Lemeshow, 2000) yang didasarkan pada hipotesis [23]:

$H_0: \beta_j = 0$ (tidak ada pengaruh variabel independen terhadap variabel dependen)

$H_1: \beta_j \neq 0, j = 1, 2, \dots, k$ (ada pengaruh variabel independen terhadap variabel dependen)

Statistik uji Wald:

$$W = \left[\frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \right]^2, \quad (8)$$

dimana $\hat{\beta}_j$ adalah penduga β_j dan $Se(\hat{\beta}_j)$ adalah standar erornya. Uji ini mengikuti distribusi normal dan H_0 ditolak jika $W > X^2_{(1,\alpha)}$ atau nilai $p_{value} < \alpha$.

Uji Kesesuaian Model

Untuk menentukan apakah model dengan variabel dependen tersebut merupakan model yang tepat, perlu dilaksanakan uji kesesuaian model menggunakan statistik *Chi-square*:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)}, \quad (9)$$

dimana :

g : total kelompok

o_k : total variabel dependen pada kelompok ke-k

$\bar{\pi}$: estimasi rata-rata probabilitas

n_k : total observasi pada kelompok ke-k

Hipotesis yang diuji adalah

H_0 : model sesuai

H_1 : model tidak sesuai

tolak H_0 jika $\hat{C} > X^2_{(\alpha, g-2)}$ dengan $db = g - 2$

atau H_0 ditolak jika nilai $p_{value} < \alpha = 0,05$ [24].

Interpretasi Model

Untuk mempermudah dalam memahami model digunakan nilai *Odds Ratio*, *OR* menggambarkan perbandingan kemungkinan suatu kejadian pada variabel dependen saat nilai $x = 1$ dibandingkan ketika $x = 0$.

$$OR = \frac{\frac{\pi(1)}{[1-\pi(1)]}}{\frac{\pi(0)}{[1-\pi(0)]}}, \quad (10)$$

diketahui bahwa $\pi(1) = \frac{\exp(\beta_0 + \beta_j)}{1 + \exp(\beta_0 + \beta_j)}$ dan

$\pi(0) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$ dimana $j = 1, 2, \dots, p$.

Berdasarkan Persamaan tersebut diperoleh bahwa:

$$OR = \exp(\beta_j) \quad (11)$$

Jika nilai *OR* sama dengan 1, berarti tidak terdapat hubungan antara variabel independen dengan variabel dependen. Nilai *OR* yang < 1 menunjukkan hubungan bersifat negatif, artinya peningkatan variabel independen menurunkan peluang kejadian pada variabel dependen. Sebaliknya, nilai *OR* yang > 1 menunjukkan hubungan bersifat positif, yaitu kenaikan variabel independen meningkatkan peluang kejadian pada

variabel dependen [25].

Ketepatan Klasifikasi

Prosedur klasifikasi yang dapat diterapkan untuk menilai akurasi klasifikasi adalah *Apparent Error Rate (APER)* [26]. Nilai *APER* merupakan indikator untuk mengidentifikasi kesalahan dalam pengklasifikasian suatu objek. Tingkat ketepatan klasifikasi dapat dilihat pada Tabel berikut [22].

Tabel 1. Ketepatan Klasifikasi

Hasil Observasi	Prediksi	
	y_1	y_2
y_1	n_{11}	n_{12}
y_2	n_{21}	n_{22}

Keterangan:

n_{11} : total subjek y_1 yang diklasifikasikan dengan benar sebagai y_1

n_{12} : total subjek y_1 yang salah diklasifikasikan sebagai y_2

n_{21} : total subjek y_2 yang salah diklasifikasikan sebagai y_1

n_{22} : total subjek y_2 yang diklasifikasikan dengan benar sebagai y_2

Maka nilai *APER* dapat dihitung dengan rumus:

$$APER = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \quad (12)$$

3. Metode Penelitian

Data yang digunakan pada riset ini adalah data sekunder yang didapatkan dari Badan Meteorologi Klimatologi dan Geofisika (BMKG) Stasiun Meteorologi Kelas II Minangkabau Padang Pariaman, dan Dinas Lingkungan Hidup (DLH) Kota Padang pada Tahun 2022 dengan jumlah data sebanyak 356. Variabel dependen

yaitu parameter ISPU yang dikategorikan menjadi Baik = 1 dan Sedang = 0, sedangkan variabel independen adalah faktor-faktor meteorologi.

Adapun Tabel variabel yang digunakan dalam riset ini adalah sebagai berikut:

Tabel 2. Variabel-variabel yang digunakan dalam penelitian

Variabel	Nama Variabel	Satuan
Dependen (Y)	Parameter Indeks Standar Pencemar Udara (Y)	$\mu\text{g}/\text{m}^3$
	Kecepatan Angin (X_1)	Knots
	Arah Angin (X_2)	Derajat
Independen (X)	Suhu Udara (X_3)	Celsius
	Kelembaban Udara (X_4)	Persen
	Curah Hujan (X_5)	Milimeter
	Tekanan Udara (X_6)	Milibar

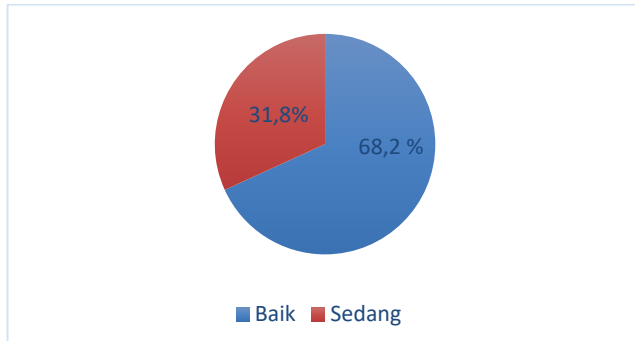
Teknik Analisis Data

Proses analisis data pada riset ini menggunakan bantuan *software* RStudio. Adapun tahapan analisis yang dilakukan meliputi:

1. Mengumpulkan data
2. Melakukan analisis statistik deskriptif pada data yang diteliti.
3. Melakukan uji multikolinearitas.
4. Melakukan pendugaan atau estimasi parameter menggunakan metode *maximum likelihood estimation*.
5. Melakukan uji *G* secara simultan.
6. Melakukan uji *Wald* secara parsial.
7. Melakukan uji kesesuaian model.
8. Menginterpretasi Model.

4. Hasil dan Pembahasan

Statistik deskriptif bertujuan untuk memberikan pandangan keseluruhan terkait karakteristik data yang digunakan dalam suatu analisis. Berikut menyajikan deskripsi dari variabel dependen.



Gambar 1. Diagram Persentase ISPU

Pada Gambar 1 parameter indeks standar pencemar udara merupakan variabel dependen dimana kategori Baik memiliki persentase sebesar 68,2%, sedangkan kategori Sedang memiliki persentase sebesar 31,8%. Dapat disimpulkan bahwa mayoritas parameter ISPU di kota Padang Tahun 2022 berada dalam kategori baik.

Uji Multikolinearitas

Uji Multikolinearitas bertujuan untuk mengidentifikasi adanya hubungan atau korelasi yang linear antar variabel independen. Deteksi dilakukan dengan memanfaatkan nilai *Variance Inflation Factor (VIF)*. Apabila nilai $VIF < 10$, berarti variabel independen tersebut tidak mengalami masalah multikolinearitas.

Tabel 3. *Variance Inflation Factor (VIF)*

Variabel Independen	VIF
Kecepatan Angin (X_1)	1,059
Arah Angin (X_2)	1,190
Suhu Udara (X_3)	1,183
Kelembaban Udara (X_4)	1,185
Curah Hujan (X_5)	1,099

Tekanan Udara (X_6)	1,009
-------------------------	-------

Pada Tabel 3 dapat disimpulkan bahwa tidak ada variabel independen yang menunjukkan nilai $VIF < 10$. Oleh karena itu, dapat dinyatakan dengan tegas bahwa tidak terdapat multikolinearitas antar variabel independen.

Pendugaan Parameter Regresi Logistik Biner

Estimasi parameter merupakan tahap awal dalam menganalisis regresi logistik biner, di mana hasil estimasi dari semua variabel independen menghasilkan parameter $\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_n$, yang diperlukan untuk mendapatkan model regresi logistik biner.

Tabel 4. Pendugaan Parameter Regresi Logistik Biner

Variabel Independen	β (Nilai Pendugaan Parameter)
<i>Intercept</i>	-298,5
Kecepatan Angin (X_1)	0,593
Arah Angin (X_2)	0,001
Suhu Udara (X_3)	-0,541
Kelembaban Udara (X_4)	0,001
Curah Hujan (X_5)	0,015
Tekanan Udara (X_6)	0,301

Menurut Tabel 4 hasil estimasi Parameter β diperoleh model awal regresi logistik biner dari faktor-faktor meteorologi yang memengaruhi parameter indeks standar pencemar udara yaitu:

$$\pi(x) = \frac{\exp\left(-289,5 + 0,593x_1 + 0,001x_2 - 0,541x_3 + 0,001x_4 + 0,015x_5 + 0,301x_6\right)}{1 + \exp\left(-289,5 + 0,593x_1 + 0,001x_2 - 0,541x_3 + 0,001x_4 + 0,015x_5 + 0,301x_6\right)}$$

dengan model logit:

$$g(x) = -289,5 + 0,593x_1 + 0,001x_2 - 0,541x_3 + 0,001x_4 + 0,015x_5 + 0,301x_6$$

Interpretasi Model

Untuk mempermudah interpretasi hasil regresi logistik, dipakai nilai *OR* yang menggambarkan seberapa besar peluang hubungan antara variabel independen dengan variabel dependen. Nilai *OR* tersebut bisa dilihat pada Tabel berikut.

Tabel 7. Nilai *Odds ratio* Model Regresi Logistik

Variabel Independen	<i>Exp</i> (β)
Kecepatan Angin (X_1)	1,810
Suhu Udara (X_3)	0,583
Curah Hujan (X_5)	1,016
Tekanan Udara (X_6)	1,352

Seperti yang terlihat di Tabel 7 nilai *odds ratio* pada variabel Kecepatan Angin (X_1) sebesar 1,810. Artinya jika kecepatan angin meningkat satu satuan maka peluang terjadinya parameter ISPU berada pada kategori baik sebesar 1,810 kali dibandingkan dengan parameter ISPU berada pada kategori sedang, dengan asumsi variabel lainnya tetap konstan. Nilai *odds ratio* untuk Suhu Udara (X_3) sebesar 0,583, menunjukkan bahwa jika suhu udara naik satu satuan akan menurunkan peluang terjadinya parameter ISPU berada pada kategori baik sebesar 0,583 kali dibandingkan dengan parameter ISPU berada pada kategori sedang, dengan asumsi variabel lainnya tetap konstan. Nilai *odds ratio* untuk Curah Hujan (X_5) sebesar 1,016, menunjukkan bahwa jika curah hujan meningkat satu satuan maka peluang terjadinya parameter ISPU berada pada kategori baik sebesar 1,016 kali dibandingkan dengan parameter ISPU berada

pada kategori sedang, dengan asumsi variabel lainnya tetap konstan. Nilai *odds ratio* untuk Tekanan Udara (X_6) sebesar 1,352, menunjukkan bahwa jika tekanan udara meningkat satu satuan maka peluang terjadinya parameter ISPU berada pada kategori baik sebesar 1,352 kali dibandingkan dengan parameter ISPU berada pada kategori sedang, dengan asumsi variabel lainnya tetap konstan.

Ketepatan Klasifikasi

Untuk menghitung ketepatan klasifikasi, dapat di lihat dari nilai ketepatan klasifikasi seperti pada tabel 8.

Tabel 8. Ketepatan Klasifikasi

Observasi	Prediksi	
	Sedang	Baik
Sedang	48	68
Baik	23	226

Menurut Tabel 8 diperoleh nilai ketepatan klasifikasi sebagai berikut:

$$\begin{aligned}
 APER &= \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \\
 &= \frac{68 + 23}{365} \times 100\% = 24,9315\%
 \end{aligned}$$

berdasarkan nilai *APER* yang diperoleh tingkat ketepatan klasifikasi dapat dihitung sebagai $(1 - APER)$ yang menghasilkan nilai sebesar 75,0685%. Ini menandakan bahwa model regresi logistik biner memiliki performa yang cukup baik dalam mengklasifikasikan data ke dalam kategori ISPU Baik dan Sedang. Hasil ini menyimpulkan bahwa model yang dibangun layak digunakan dalam penelitian ini.

5. Kesimpulan

Setelah data dianalisis, diperoleh kesimpulan riset ini yaitu faktor-faktor meteorologi yang paling berpengaruh terhadap parameter indeks standar pencemar udara adalah Kecepatan Angin (X_1), Suhu Udara (X_3), Curah Hujan (X_5), dan Tekanan Udara (X_6) dilihat dari nilai $p_{value} < 0,05$. Model regresi logistik biner yang diperoleh adalah:

$$\pi(x) = \frac{\exp\left(\frac{-289,5 + 0,593x_1 - 0,541x_3 + 0,015x_5}{+0,301x_6}\right)}{1 + \exp\left(\frac{-289,5 + 0,593x_1 - 0,541x_3 + 0,015x_5}{+0,301x_6}\right)}$$

dengan model logit:

$$g(x) = (-289,5 + 0,593x_1 - 0,541x_3 + 0,015x_5 + 0,301x_6)$$

Dengan Ketepatan klasifikasi yang didapatkan pada model regresi logistik biner sebesar 75,0685%. Ini menunjukkan bahwa model pengklasifikasian faktor-faktor meteorologi terhadap parameter indeks standar pencemar udara cukup akurat dan baik dalam memprediksi data penelitian ini.

Daftar Pustaka

- [1] A. A. Nababan, M. Jannah, M. Aulina, and D. Andrian, "Prediksi Kualitas Udara Menggunakan Xgboost Dengan Synthetic Minority Oversampling Technique (Smote) Berdasarkan Indeks Standar Pencemaran Udara (Ispu)," *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 7, no. 1, pp. 214–219, 2023, doi: 10.59697/jtik.v7i1.66.
- [2] BPS and S. Barat, "Indeks Kualitas Udara Menurut Kabupaten/Kota," Barat, Sumatera Barat BPS.
- [3] R. Pratiwi, R. Widyasari, and M. Fathonni, "Parameter Kritis Indeks Standar Pencemar Udara," vol. 5, no. 1, pp. 499–513, 2024.
- [4] D. Chaniago, Z. Annisa, and I. S. Ramadhani, "Indeks Standar Pencemar Udara (ISPU) Sebagai Informasi Mutu Udara Ambien Di Indonesia," MENLHK.
- [5] D. Kartikasari, "Analisis Faktor-Faktor Yang Mempengaruhi Level Polusi Udara Dengan Metode Regresi Logistik Biner," *MATHunesa J. Ilm. Mat.*, vol. 8, no. 1, pp. 55–59, 2020, doi: 10.26740/mathunesa.v8n1.p55-59.
- [6] L. Anisa and N. A. K. Rifai, "Analisis Regresi Logistik Biner dengan Metode Penalized Maximum Likelihood pada Penyakit Covid-19 di RSUD Pringsewu," *J. Ris. Stat.*, pp. 129–136, 2022, doi: 10.29313/jrs.v2i2.1425.
- [7] L. H. Hasibuan and S. Musthofa, "Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang," *JOSTECH J. Sci. Technol.*, vol. 2, no. 1, pp. 85–95, Mar. 2022, doi: 10.15548/jostech.v2i1.3802.
- [8] L. H. Hasibuan, F. Yanuar, D. Devianto, and M. Maiyastri, "Quantile Regression Analysis; Simulation Study With Violation of Normality Assumption," *JOSTECH J. Sci. Technol.*, vol. 4, no. 2, pp. 133–142, 2024.
- [9] L. H. Hasibuan, D. M. Putri, and M. Jannah, "Simple Linear Regression Method to Predict Cooking Oil Prices in the Time of Covid-19," *Logaritma J. Ilmu-ilmu Pendidik.*

- dan Sains*, vol. 10, no. 01, pp. 81–94, 2022.
- [10] N. I. E. Harahap, L. H. Hasibuan, and S. Musthofa, “Model Klasifikasi Tinggi Badan Balita Stunting Menggunakan Regresi Logistik Biner,” *JOSTECH J. Sci. Technol.*, vol. 5, no. 2, pp. 257–270, 2025.
- [11] L. H. Hasibuan, F. Yanuar, and D. Devianto, “Hypertension Status Model with Bayesian Adaptive LASSO Binary Quantile Regression Using Asymmetric Laplace Distribution,” *IAENG Int. J. Comput. Sci.*, vol. 52, no. 8, 2025.
- [12] L. H. Hasibuan, F. Yanuar, D. Devianto, and M. Maiyastri, “MODELING CLASSIFICATION OF STUNTING TODDLER HEIGHT USING BAYESIAN BINARY QUANTILE REGRESSION WITH PENALIZED LASSO,” *Mathline J. Mat. dan Pendidik. Mat.*, vol. 10, no. 2, pp. 373–385, 2025, doi: <https://doi.org/10.31943/mathline.v10i2.928>.
- [13] L. H. Hasibuan, F. Yanuar, D. Devianto, M. Maiyastri, and R. Rudiyanto, “COMPARISON BETWEEN BAYESIAN QUANTILE REGRESSION AND BAYESIAN LASSO QUANTILE REGRESSION FOR MODELING POVERTY LINE WITH PRESENCE OF HETEROSCEDASTICITY IN WEST SUMATRA,” *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 19, no. 3, pp. 1587–1596, 2025, doi: <https://doi.org/10.30598/barekengvol19iss3pp1587-1596>.
- [14] A. Varamita, “Analisis regresi logistik dan aplikasinya pada penyakit anemia untuk ibu hamil di RSKD Ibu dan Anak Siti Fatimah Makassar,” 2017, *FMIPA*.
- [15] L. H. Hasibuan, F. Yanuar, V. P. Harahap, and L. Qalbi, “Performance Quantile Regression and Bayesian Quantile Regression in Dealing with Non-normal Errors (Case Study on Simulated Data),” *Numer. J. Mat. Pendidik. Mat.*, vol. 8, 2024.
- [16] W. Alwi, E. Ermawati, and S. Husain, “Analisis Regresi Logistik Biner Untuk Memprediksi Kepuasan Pengunjung Pada Rumah Sakit Umum Daerah Majene,” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 6, no. 1, p. 20, 2018, doi: [10.24252/msa.v6i1.4783](https://doi.org/10.24252/msa.v6i1.4783).
- [17] L. H. Hasibuan and S. Musthofa, “Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang | Hasibuan | JOSTECH: Journal of Science and Technology.” Accessed: Apr. 26, 2022. [Online]. Available: <https://www.ejournal.uinib.ac.id/jurnal/index.php/jostech/article/view/3802/pdf>
- [18] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013.
- [19] A. S. Sholih, L. H. Hasibuan, and I. D. Rianjaya, “ANALISIS REGRESI LOGISTIK ORDINAL TERHADAP FAKTOR-FAKTOR YANG MEMPENGARUHI PREDIKAT KELULUSAN MAHASISWA SARJANA UIN IMAM BONJOL PADANG,” *Map (Mathematics Appl. J.*, vol. 6, no. 2,

- pp. 99–110, 2024.
- [20] S. W. Rizki, “ANALISIS REGRESI LOGISTIK MULTINOMIAL PADA PEMILIHAN ALAT KONTRASEPSI WANITA (Studi Kasus di Puskesmas Sungai Kakap),” vol. 08, no. 4, pp. 751–758, 2019.
- [21] Y. Tampil, H. Komaliq, and Y. Langi, “Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado,” *d’CARTESIAN*, vol. 6, no. 2, p. 56, 2017, doi: 10.35799/dc.6.2.2017.17023.
- [22] D. L. W. Putri, S. Mariani, and S. Sunarmi, “Peningkatan Ketepatan Klasifikasi Model Regresi Logistik Biner dengan Metode Bagging (Bootstrap Aggregating),” *Indones. J. Math. Nat. Sci.*, vol. 44, no. 2, pp. 61–72, 2021, doi: 10.15294/ijmns.v44i2.33144.
- [23] A. R. Aditya, S. Suparti, and S. Sudarno, “Ketepatan Klasifikasi Pemilihan Metode Kontrasepsi Di Kota Semarang Menggunakan Booststrap Aggregatting Regresi Logistik Multinomial,” *J. Gaussian*, vol. 4, no. 1, pp. 11–20, 2015.
- [24] A. Yudissanta and M. Ratna, “Analisis Pemakaian Kemoterapi pada Kasus Kanker Payudara dengan Menggunakan Metode,” *Sains Dan Seni Its*, vol. 1, no. 1, 2012.
- [25] T. M. T. Nisva and V. Ratnasari, “Analisis Regresi Logistik Biner pada Faktor-Faktor yang Mempengaruhi Jenis Perceraian di Kabupaten Lumajang,” *Inferensi*, vol. 3, no. 1, p. 29, 2020, doi: 10.12962/j27213862.v3i1.6879.
- [26] M. A. Suhendra, D. Ispriyanti, and S. Sudarno, “Ketepatan Klasifikasi Pemberian Kartu Keluarga Sejahtera Di Kota Semarang Menggunakan Metode Regresi Logistik Biner Dan Metode Chaid,” *J. Gaussian*, vol. 9, no. 1, pp. 64–74, 2020, doi: 10.14710/j.gauss.v9i1.27524.