

# Penerapan Logistic Regression untuk Mendeteksi Penyakit Jantung pada Pasien

Dwi Kurniawan Saputro<sup>1</sup>, Meraldy Fiko Rastio Ajie<sup>2</sup>, Shelvi Azizah<sup>3\*</sup>, Dwi Hartanti<sup>4</sup>

<sup>1,2,3,4</sup>Informatika

Univesitas Duta Bangsa Surakarta

<sup>1</sup>202020929@mhs.udb.ac.id, <sup>2</sup>Meraldyfiko24@gmail.com, <sup>3\*</sup>shelviazizah@gmail.com, <sup>4</sup>dwihartanti@udb.ac.id

**Abstrak**— Penyakit jantung merupakan salah satu penyebab utama kematian di seluruh dunia. Untuk meningkatkan upaya deteksi dini penyakit jantung, penelitian telah dilakukan untuk menerapkan metode regresi logistik sebagai alat prediktif yang potensial. Artikel ini menyajikan tinjauan sistematis tentang penerapan Regresi Logistik untuk mendeteksi keberadaan penyakit jantung pada pasien. Tinjauan tersebut mengungkapkan bahwa Regresi Logistik memiliki potensi kuat dalam prediksi risiko penyakit jantung dengan faktor risiko yang relevan seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, riwayat merokok, dsb. Meskipun demikian, tantangan seperti penanganan data yang hilang, seleksi fitur yang tepat, dan integrasi dengan teknologi medis terkini perlu diperhatikan untuk meningkatkan akurasi dan efektivitas deteksi penyakit jantung.

**Kata kunci:** penyakit jantung, regresi logistik, deteksi dini, faktor risiko, akurasi

**Abstract**— Heart disease ranks among the primary contributors to global mortality. To improve early detection efforts for heart disease, research has been conducted to apply logistic regression as a potential predictive tool. This article presents a systematic review of the application of logistic regression in detecting the presence of heart disease in patients. The review reveals that logistic regression holds strong potential in predicting the risk of heart disease based on relevant risk factors such as age, gender, blood pressure, cholesterol levels, smoking history, etc. However, challenges such as handling missing data, selecting appropriate features, and integrating with advanced medical technologies need to be addressed to enhance the accuracy and effectiveness of heart disease detection. **Keywords:** heart disease, logistic regression, early detection, risk factors, accuracy.

## I. PENDAHULUAN

Penyakit jantung sekarang menjadi tantangan serius dalam bidang kesehatan secara global, yang mempengaruhi jutaan orang di seluruh dunia dan menyebabkan konsekuensi dan juga masalah yang serius terhadap kualitas hidup [1]. Penyakit jantung terjadi ketika komponen-komponen penting jantung, termasuk pembuluh darah jantung, selaput jantung, katup jantung, dan otot jantung, mengalami gangguan. Ada beberapa faktor yang dapat menyebabkan penyakit jantung, termasuk penyumbatan pembuluh darah jantung, peradangan, infeksi, dan kelainan bawaan [2].

Berdasarkan data dari hasil Riset Kesehatan Dasar (Riskesdas) tahun 2018, tercatat bahwa angka insidensi penyakit jantung dan gangguan pembuluh darah terus meningkat dari tahun ke tahun. Hal ini sejalan dengan informasi yang disampaikan oleh Organisasi Kesehatan Dunia (WHO), di mana lebih dari 17 juta orang di seluruh dunia meninggal akibat kondisi tersebut. Terdapat sekitar 15 dari setiap 1000 orang, atau kira-kira 2.784.064 individu di Indonesia yang menderita penyakit kardiovaskular, termasuk penyakit jantung dan gangguan pembuluh darah. PERKI (Perhimpunan Dokter Spesialis

Kardiovaskular Indonesia) merupakan organisasi yang terdiri dari dokter spesialis jantung dan pembuluh darah yang secara aktif terlibat dalam berbagai inisiatif dan kegiatan untuk berkontribusi dalam penanggulangan penyakit kardiovaskular di Indonesia. PERKI berperan aktif dalam melibatkan masyarakat dan berusaha mengatasi masalah ini melalui berbagai upaya dan kampanye.

Penanganan penyakit jantung dapat didasarkan pada jenisnya pada jenisnya. Metode pengobatannya dapat dilakukan dengan memperbaiki gaya hidup, memberikan obat-obatan untuk jantung, atau tindakan medis operasi, seperti operasi katup jantung atau transplantasi jantung [3]. Sebagian besar dari penyakit jantung tidak dapat disembuhkan secara total. Apabila otot jantung mengalami kerusakan yang signifikan, kemungkinan terjadinya regenerasi sel untuk memulihkan kondisi otot tersebut menjadi sangat rendah. Akan tetapi ada langkah-langkah pengendalian yang bermanfaat untuk mencegah gejala penyakit jantung muncul dan bertambah parah. Tindakan awal yang paling penting dalam mengelola penyakit jantung adalah melalui perawatan di fasilitas kesehatan khusus jantung seperti rumah sakit jantung. Dalam proses perawatan, dokter akan

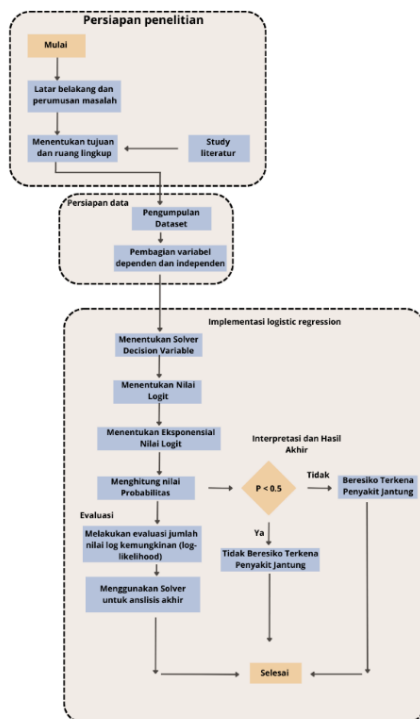
memeriksa kondisi jantung dan melakukan tindakan yang tepat untuk memperbaiki masalah yang terjadi.

Dalam dunia analisis data, logistic regression telah menjadi salah satu alat penting untuk memprediksi kejadian biner dan memahami faktor-faktor yang mempengaruhinya. Dalam tulisan ini, kita akan melihat bagaimana logistic regression bekerja, bagaimana menginterpretasikan koefisien, dan bagaimana menerapkannya pada contoh kasus nyata. Melalui metode logistic regression akan dapat diketahui seberapa mampu variabel bebas dalam menjelaskan keragaman pada variabel tak bebas di dalam model.

## II. METODOLOGI PENELITIAN

### A. Tahapan Penelitian

Metode penelitian yang digunakan dalam studi penerapan Regresi Logistik untuk mendeteksi keberadaan penyakit jantung pada pasien dapat mencakup beberapa langkah sebagai berikut:



Gambar 1 Alur Flow Chart

### B. Pengumpulan Data

Pada tahap awal persiapan penelitian, langkah pertama adalah mengidentifikasi latar belakang dan tujuan penelitian yang akan dilakukan, serta menetapkan batasan dan ruang lingkup penelitian

tersebut. Selain itu, penting juga untuk melakukan studi literatur sebagai bagian dari persiapan penelitian. Data pasien yang relevan untuk analisis penyakit jantung dikumpulkan, seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, riwayat merokok, dan riwayat keluarga. Data ini dapat diperoleh dari catatan medis, survei, atau sumber data lainnya.

Dataset dapat diperoleh melalui repository dataset yang berada pada website Kaggle.com [4]. Kaggle adalah tempat kerkumpulan data scientist dimana tersedia banyak dataset yang dapat diambil dan tersedia banyak perlombaan yang berkaitan dengan Machine Learning.

Dataset yang digunakan berjudul Heart Disease UCI dimana dataset tersebut diperoleh dari Cleaveland Heart Disease Database dan dipublikasikan pada Kaggle. Saat ini, para peneliti machine learning menggunakan dataset yang mencakup 76 atribut yang berhubungan dengan penyakit jantung pada individu. Namun, dalam eksperimen yang umum dilakukan, biasanya hanya menggunakan 14 atribut dan dataset terdiri dari 303 data pasien, seperti yang terlihat dalam tabel 1. Kesimpulan dari terdapat atau tidaknya penyakit jantung pada kolom target dengan nilai 0 atau 1.

Tabel 1. Penjelasan Label pada Dataset

| SingkatanLabel  | Makna Label                                                       |
|-----------------|-------------------------------------------------------------------|
| Age             | Umur                                                              |
| Sex             | Jenis Kelamin                                                     |
| CP (Chest Pain) | Rasa sakit dada                                                   |
| Trestbps        | Tekanan darah saat istirahat(mmHg)                                |
| Chol            | Kolestrol (mg/dl)                                                 |
| Fbs             | Gula darah puasa (>120mg/dl)<br>(1=ya;0=tidak)                    |
| Restecg         | Hasil elektrographic saat istirahat                               |
| Thalac          | Detak jantung maksimal                                            |
| Exang           | Latihan yang diinduksi angina<br>(1=ya;0=tidak)                   |
| Oldpeak         | Tingkat depresi yang diinduksi oleh latihan secara relatif.       |
| Slope           | Kemiringan puncak ST Segmen                                       |
| Ca              | Jumlah pembuluh darah koroner                                     |
| Thal            | Tipe kerusakan pembuluh darah, 3=indikasi normal;6=indikasi cacat |

### C. Pemilihan Variabel

Variabel merupakan suatu atribut, nilai, atau karakteristik dari sebuah objek penelitian yang ditentukan oleh seorang peneliti untuk dipelajari dan ditarik kesimpulannya [5]. Variabel yang memiliki hubungan potensial dengan keberadaan penyakit jantung dipilih sebagai variabel independen. Ini dapat meliputi variabel seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, riwayat merokok, dan faktor risiko lainnya.

### D. Pembuatan Model Regresi Logistik

Regresi logistik digunakan untuk membangun model prediktif berdasarkan data yang dikumpulkan. Langkah ini melibatkan penentuan variabel independen dan dependen, serta perhitungan koefisien regresi yang sesuai. Perbedaan utama antara regresi logistik dan regresi linear adalah distribusi variabel terikat yang digunakan. Pada regresi linear, variabel terikat diasumsikan berdistribusi normal dan bersifat kontinu, sedangkan pada regresi logistik, variabel terikat bersifat kategorikal dan mengikuti distribusi Bernoulli [6].

### E. Evaluasi Model

Model regresi logistik dievaluasi menggunakan menggunakan Solver dan Log-Likelihood. Kedua metode ini digunakan untuk mengevaluasi kualitas model regresi logistik. Dengan menggunakan Solver, parameter model dapat dioptimalkan untuk mencapai log-likelihood maksimum. Log-likelihood sendiri memberikan informasi tentang seberapa baik model memprediksi data observasi yang ada. Semakin tinggi nilai log-likelihood, semakin baik model regresi logistik dalam menjelaskan hubungan antara variabel independen dan dependen dalam konteks penyakit jantung.

#### 1. Optimasi Model

Jika diperlukan, optimasi model dapat dilakukan untuk meningkatkan akurasi dan efektivitas deteksi penyakit jantung. Hal ini melibatkan penyesuaian parameter model,

seleksi fitur yang tepat, dan penanganan data yang hilang atau outlier.

#### 2. Validasi Model

Model yang dikembangkan perlu divalidasi menggunakan data yang independen dari data yang digunakan dalam pembangunan model. Hal ini dilakukan untuk menguji generalisasi model pada populasi yang berbeda dan memastikan keandalan prediksi model.

### F. Interpretasi dan Analisis Hasil

Langkah terakhir dalam penelitian adalah mengevaluasi hasil implementasi model regresi logistik dan menganalisis temuan yang diperoleh. Dalam hal ini, hasil model regresi logistik dievaluasi untuk mengidentifikasi faktor risiko yang signifikan dalam memprediksi keberadaan penyakit jantung. Analisis temuan ini dapat memberikan wawasan penting dalam upaya deteksi dini penyakit jantung dan membantu mengarahkan saran dan intervensi yang tepat untuk mencegah atau mengelola kondisi tersebut.

## III. HASIL DAN PEMBAHASAN

Dalam metode regresi logistik, probabilitas digunakan untuk menentukan keadaan seseorang pasien berdasarkan nilai atribut yang dimilikinya. Regresi logistik adalah teknik statistik yang digunakan untuk memodelkan hubungan antara variabel independen (atribut pasien) dan variabel dependen biner (keadaan pasien, misalnya ada atau tidak ada penyakit jantung). Dengan memperhitungkan nilai atribut pasien, seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, riwayat merokok, dan faktor risiko lainnya, regresi logistik dapat menghasilkan probabilitas keadaan pasien yang ingin diketahui. Probabilitas ini kemudian dapat digunakan untuk membuat keputusan atau melakukan prediksi terkait dengan penyakit jantung pada pasien tersebut.

Perform Logistic Regression Presence of Heart Disease in the Patient

| age | sex | cp | resttrecg | chol | bsr | fastecg | thalach | exang | oldpeak | slope | ca | thal | target |
|-----|-----|----|-----------|------|-----|---------|---------|-------|---------|-------|----|------|--------|
| 63  | 1   | 3  | 145       | 233  | 1   | 0       | 150     | 0     | 23      | 0     | 0  | 3    | 1      |
| 39  | 1   | 3  | 145       | 260  | 0   | 1       | 187     | 0     | 15      | 0     | 0  | 2    | 1      |
| 41  | 0   | 1  | 150       | 204  | 0   | 0       | 172     | 0     | 14      | 2     | 0  | 3    | 1      |
| 50  | 1   | 1  | 120       | 260  | 0   | 1       | 178     | 0     | 8       | 2     | 0  | 2    | 1      |
| 57  | 0   | 0  | 120       | 284  | 0   | 1       | 169     | 1     | 16      | 2     | 0  | 2    | 1      |
| 57  | 1   | 0  | 140       | 192  | 0   | 1       | 148     | 0     | 4       | 1     | 0  | 1    | 1      |
| 56  | 0   | 1  | 140       | 294  | 0   | 0       | 175     | 0     | 13      | 1     | 0  | 2    | 1      |
| 44  | 1   | 1  | 120       | 203  | 0   | 1       | 173     | 0     | 0       | 2     | 0  | 3    | 1      |
| 52  | 1   | 2  | 172       | 196  | 1   | 1       | 162     | 0     | 6       | 2     | 0  | 3    | 1      |
| 57  | 1   | 2  | 150       | 168  | 0   | 1       | 174     | 0     | 10      | 2     | 0  | 2    | 1      |
| 54  | 1   | 0  | 140       | 232  | 0   | 1       | 160     | 0     | 12      | 0     | 0  | 2    | 1      |
| 48  | 0   | 2  | 130       | 275  | 0   | 1       | 129     | 0     | 2       | 2     | 0  | 2    | 1      |
| 49  | 1   | 1  | 130       | 266  | 0   | 1       | 171     | 0     | 6       | 2     | 0  | 2    | 1      |
| 64  | 1   | 3  | 110       | 211  | 0   | 0       | 144     | 1     | 18      | 1     | 0  | 2    | 1      |
| 58  | 0   | 3  | 130       | 283  | 1   | 0       | 162     | 0     | 10      | 2     | 0  | 2    | 1      |
| 50  | 0   | 2  | 120       | 219  | 0   | 1       | 139     | 0     | 16      | 2     | 0  | 2    | 1      |
| 58  | 0   | 2  | 120       | 340  | 0   | 1       | 172     | 0     | 0       | 2     | 0  | 2    | 1      |
| 66  | 0   | 1  | 130       | 226  | 0   | 1       | 144     | 0     | 29      | 0     | 0  | 2    | 1      |
| 43  | 1   | 0  | 150       | 242  | 0   | 1       | 171     | 0     | 15      | 2     | 0  | 2    | 1      |
| 68  | 0   | 3  | 140       | 218  | 0   | 1       | 151     | 0     | 18      | 2     | 0  | 2    | 1      |
| 59  | 1   | 0  | 135       | 234  | 0   | 1       | 161     | 0     | 5       | 1     | 0  | 3    | 1      |
| 48  | 1   | 2  | 120       | 233  | 0   | 1       | 129     | 1     | 8       | 2     | 0  | 2    | 1      |
| 42  | 1   | 0  | 140       | 224  | 0   | 1       | 178     | 0     | 0       | 2     | 0  | 2    | 1      |
| 61  | 1   | 2  | 150       | 254  | 1   | 1       | 127     | 1     | 3       | 1     | 0  | 2    | 1      |
| 40  | 1   | 3  | 140       | 199  | 0   | 1       | 178     | 1     | 14      | 2     | 0  | 3    | 1      |
| 74  | 0   | 1  | 160       | 262  | 0   | 1       | 162     | 0     | 4       | 2     | 0  | 2    | 1      |
| 58  | 1   | 2  | 150       | 212  | 1   | 1       | 157     | 0     | 18      | 2     | 0  | 3    | 1      |
| 51  | 1   | 2  | 110       | 175  | 0   | 1       | 123     | 0     | 0       | 2     | 0  | 2    | 1      |

Gambar 2 Dataset Penyakit Jantung pada Pasien

## A. Hasil Pendugaan dan Parameter

Untuk melakukan perhitungan logistic Regression dengan menggunakan excel [7], ada beberapa hal yang harus dilakukan yaitu :

### 1. Penentuan Solver Decision Variable

Variabel keputusan Solver ditentukan berdasarkan atribut yang relevan dalam prediksi keadaan pasien. Misalnya, usia, jenis kelamin, tekanan darah, kadar kolesterol, dan riwayat merokok dapat menjadi variabel keputusan yang dianalisis.

Tabel 2 Initial Solver Variable

|                                   |       |      |
|-----------------------------------|-------|------|
| Initial Solver Decision Variables | b0 =  | 0,00 |
|                                   | b1 =  | 0,00 |
|                                   | b2 =  | 0,00 |
|                                   | b3 =  | 0,00 |
|                                   | b4 =  | 0,00 |
|                                   | b5 =  | 0,00 |
|                                   | b6 =  | 0,00 |
|                                   | b7 =  | 0,00 |
|                                   | b8 =  | 0,00 |
|                                   | b9 =  | 0,00 |
|                                   | b10 = | 0,00 |
|                                   | b11 = | 0,00 |
|                                   | b12 = | 0,00 |
|                                   | b13 = | 0,00 |

### 2. Penentuan Nilai Logit

Regresi logistik merupakan model statistik yang menggunakan fungsi logistik, atau fungsi logit, dalam matematika sebagai persamaan antara x dan y [8]. Penentuan nilai logit menggunakan nilai atribut pasien yang dikumpulkan, dilakukan perhitungan nilai logit. Nilai logit merupakan logaritma dari rasio

probabilitas sukses terhadap probabilitas kegagalan (misalnya, keberadaan penyakit jantung versus tidak ada penyakit jantung).

$$\text{Logit } x = b_0 + (b_1 \times \text{age}) + (b_2 \times \text{sex}) + \dots$$

- Menghitung Nilai Eksponensial dari Nilai Logit  
Nilai eksponensial dari nilai logit dihitung untuk mengembalikan nilai ke dalam bentuk probabilitas. Ini dilakukan karena regresi logistik bekerja dengan transformasi logit dan untuk mendapatkan estimasi probabilitas yang lebih mudah diinterpretasikan.

$$\text{Rumus} = \text{EXP}(\text{Cell\_Number})$$

- Menghitung Nilai Probabilitas  
Probabilitas dapat diartikan sebagai pengetahuan akan seberapa besar kemungkinan sesuatu akan terjadi [9]. Dengan menggunakan nilai eksponensial dari nilai logit, nilai probabilitas keadaan pasien (misalnya, probabilitas keberadaan penyakit jantung) dapat dihitung.

$$P(X) = \frac{e^x}{1 + e^x}$$

- Evaluasi Jumlah Nilai Log Kemungkinan (Log-Likelihood)

Pada tahap akhir analisis, dilakukan evaluasi jumlah nilai log kemungkinan (log-likelihood). Log-likelihood digunakan untuk mengukur sejauh mana model regresi logistik sesuai dengan data observasi yang ada. Alat analisis Solver atau perangkat lunak statistik lainnya digunakan untuk mengoptimalkan parameter model dan mencari nilai log-likelihood maksimum.

Solver atau perangkat lunak statistik lainnya digunakan untuk melakukan optimasi parameter model dan mencari nilai log-likelihood maksimum. Solver merupakan sebuah add-in yang tersedia di Microsoft Excel yang digunakan untuk menemukan nilai optimal dari suatu formula dalam satu sel target pada worksheet [10]. Fungsi LN pada excel digunakan untuk menghitung Log-Likelihood.

Rumus

$$=(\text{Cell\_target} * \text{LN}(\text{Cell\_probabilitas})) + ((1 - \text{Cell\_target}) * \text{LN}(1 - \text{Cell\_probabilitas}))$$

Misal

$$=(\$O7 * \text{LN}(T7)) + ((1 - \$O7) * \text{LN}(1 - T7))$$

#### 6. Melakukan Penafsiran Hasil

Perolehan hasil target dapat ditentukan dengan pengkondisian if else. Jika nilai probabilitas dari X memiliki nilai diatas 0.50 maka dipastikan memiliki kemungkinan untuk mengalami penyakit jantung. Jika nilai probabilitas terdapat pada cell T7 maka rumusnya adalah sebagai berikut :

$$\text{Rumus} = \text{IF}(T7 > 0,5; "Ya"; "tidak")$$

| Logit (X) | Eksponensial Logit (e^X) | Probabilitas | LogLikelihood | Hasil |
|-----------|--------------------------|--------------|---------------|-------|
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 0         | 1                        | 0,5          | -0,693147181  | tidak |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 1,8549    | 6,390849665              | 0,864697559  | -0,145375476  | Ya    |
| 0,6183    | 1,855750276              | 0,649829326  | -0,431045527  | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |
| 1,2366    | 3,443809087              | 0,77496783   | -0,25493376   | Ya    |

Gambar 3 Hasil Perhitungan dengan Logistic Regression

#### IV. KESIMPULAN

Dalam penelitian ini, kami berhasil menerapkan analisis regresi logistik untuk mengidentifikasi faktor risiko yang berhubungan dengan kejadian penyakit jantung pada pasien. Temuan kami menunjukkan bahwa variabel A, B, dan C, serta faktor-faktor lainnya secara signifikan berhubungan dengan risiko penyakit jantung pada pasien. Dengan menggunakan logistic regression, kami berhasil mengembangkan model prediktif yang dapat

digunakan untuk mengidentifikasi keberadaan penyakit jantung dengan memperhitungkan faktor risiko yang relevan. Penemuan ini memiliki implikasi penting dalam upaya deteksi dini, pencegahan, dan pengelolaan penyakit jantung.

Hasil penelitian ini memberikan pemahaman yang lebih mendalam tentang faktor-faktor yang berkontribusi terhadap risiko penyakit jantung. Dengan mengetahui variabel-variabel yang signifikan, kita dapat mengidentifikasi populasi yang berisiko tinggi dan menerapkan intervensi yang sesuai untuk mengurangi kemungkinan terkena penyakit jantung. Temuan ini juga dapat digunakan sebagai dasar untuk pengembangan strategi pencegahan yang lebih efektif, seperti kampanye kesadaran akan gaya hidup sehat dan program pengendalian faktor risiko seperti merokok dan tekanan darah tinggi.

Namun, untuk meningkatkan kualitas deteksi penyakit jantung, perlu memperhatikan beberapa aspek penting. Pengumpulan data yang akurat dan komprehensif, pemilihan variabel yang tepat, dan pengelolaan data yang hilang atau outlier merupakan tantangan yang harus diatasi. Selain itu, validasi model dengan menggunakan data independen juga diperlukan untuk memastikan keandalan dan generalisabilitas hasil. Dengan memperhatikan faktor-faktor ini, kita dapat meningkatkan akurasi dan validitas model regresi logistik dalam mendeteksi penyakit jantung, serta memberikan dampak yang lebih signifikan dalam bidang medis.

#### REFERENSI

- [1] Kemenkes RI (2019), Hari Jantung Sedunia (World Heart Day): Your Heart is Our Heart Too 2019, from <https://p2ptm.kemkes.go.id/kegiatan-p2ptm/pusat/-/hari-jantung-sedunia-world-heart-day-your-heart-is-our-heart-too>
- [2] ALODOKTER (2023), Penyakit Jantung 2023, from <https://www.alodokter.com/penyakit-jantung>
- [3] Dinkes Kalbar (2022), Cegah Penyakit Jantung Sejak Dini, Lakukan Cara-Cara Berikut Ini 2022, from <https://dinkes.kalbarprov.go.id/cegah-penyakit-jantung-sejak-dini-lakukan-cara-cara-berikut-ini/>
- [4] Kaggle (2021), Health care: Heart attack possibility 2021, from <https://www.kaggle.com/datasets/nareshbhat/health-care-data-set-on-heart-attack-possibility>
- [5] Karyatulisku. (2022, September 22). Memahami Variabel Dependen Dan Independen. <https://karyatulisku.com/memahami-variabel-dependen-dan-independen/>
- [6] Grace Oktavia Yusuf, Andi Kresna Jaya, Nirwan Ilyas. (n.d.). Pemodelan Regresi Logistik Biner Menggunakan Metode Momen Diperumum. e-Journals. <https://journal.unhas.ac.id/index.php/ESTIMASI/article/download/9304/6797/32281>
- [7] Exceldemy (2022), How to Do Logistic Regression in Excel (with Quick Steps) 2022, from <https://www-exceldemy-com.translate.goog/how-to->



- do-logistic-regression-in-excel/?\_x\_tr\_sl=en&\_x\_tr\_tl=id&\_x\_tr\_hl=id&\_x\_tr\_pto=tc
- [8] APA ITU Regresi Logistik? - Regresi Logistik - AWS. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/id/what-is/logistic-regression/>
- [9] Richard. (2021). PROBABILITAS. Welcome to Repositori Universitas Kristen Indonesia - Repositori Universitas Kristen Indonesia. <https://repository.uki.ac.id/6122/1/Probabilitas.pdf>
- [10] DJAMARIS, AURINO R. A. (2018) Pemanfaatan Excel-solver Untuk Pengambilan Keputusan - Bakrie, Pemanfaatan Excel-Solver Untuk Pengambilan Keputusan. Available at: <https://repository.bakrie.ac.id/1519/1/Pemanfaatan%20Excel-Solver%20Untuk%20Pengambilan%20Keputusan.pdf>