



**TUGAS AKHIR - KS184822**

**ANALISIS GEJALA KLINIS KANKER PAYUDARA  
MENGUNAKAN REGRESI LOGISTIK BINER-*GENETIC*  
ALGORITHM DI RS NUR HIDAYAH BANTUL**

**DINDA GALUH GUMINTA  
NRP 062115 4000 0093**

**Dosen Pembimbing  
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**TUGAS AKHIR - KS184822**

**ANALISIS GEJALA KLINIS KANKER PAYUDARA  
MENGUNAKAN REGRESI LOGISTIK BINA-  
GENETIC ALGORITHM DI RS NUR HIDAYAH  
BANTUL**

**DINDA GALUH GUMINTA  
NRP 062115 4000 0093**

**Dosen Pembimbing  
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**FINAL PROJECT - KS184822**

**BREAST CANCER CLINICAL SYMPTOMS ANALYSIS  
USING BINARY LOGISTIC REGRESSION - GENETIC  
ALGORITHM IN NUR HIDAYAH HOSPITAL  
BANTUL**

**DINDA GALUH GUMINTA  
SN 062115 4000 0093**

**Supervisor  
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME  
DEPARTEMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**



## LEMBAR PENGESAHAN

# ANALISIS GEJALA KLINIS KANKER PAYUDARA MENGUNAKAN REGRESI LOGISTIK BINER - GENETIC ALGORITHM DI RS NUR HIDAYAH BANTUL

### TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Statistika  
pada

Program Studi Sarjana Departemen Statistika  
Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember

Oleh :

**Dinda Galuh Guminta**  
NRP. 062115 4000 0093

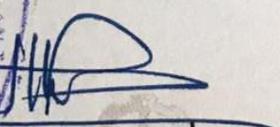
Disetujui oleh Pembimbing :

**Dr. Santi Wulan Purnami, S.Si., M.Si.**

NIP. 19720923 199803 2 001



Mengetahui,  
Kepala Departemen Statistika

  
**Dr. Suhartono**  
NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

*(Halaman ini sengaja dikosongkan)*

# **ANALISIS GEJALA KLINIS KANKER PAYUDARA MENGUNAKAN REGRESI LOGISTIK BINER-GENETIC ALGORITHM DI RS NUR HIDAYAH BANTUL**

**Nama Mahasiswa : Dinda Galuh Guminta**  
**NRP : 062115 4000 00493**  
**Departemen : Statistika**  
**Dosen Pembimbing : Dr. Santi Wulan Purnami, S.Si., M.Si.**

## **Abstrak**

*Kanker payudara adalah penyakit seluler yang ditandai oleh pertumbuhan sel abnormal pada jaringan payudara. Penyakit ini menempati peringkat pertama jenis kanker penyebab kematian di Indonesia. Pada umumnya, wanita melakukan pemeriksaan kanker payudara pada stadium lanjut yang menyebabkan rendahnya angka harapan hidup, padahal kanker payudara dapat dicegah dengan mengenali gejala-gejala klinis seperti benjolan di payudara atau ketiak, bentuk dan ukuran payudara yang tidak sama, penarikan puting susu, dan lain-lain. Disamping itu, gangguan tidur perlu dipertimbangkan sebagai risiko kanker payudara. Pada penelitian ini gangguan tidur diukur dengan jam tidur responden setiap malam. Oleh karena itu, dilakukan analisis gejala klinis dan gangguan tidur dengan regresi logistik biner. Pemilihan variabel dilakukan menggunakan seleksi backward, forward, stepwise, dan genetic algorithm dengan tujuan mendapatkan gejala-gejala yang berpengaruh. Hasil penelitian menggunakan semua metode seleksi menunjukkan bahwa nyeri di bagian tertentu payudara berpengaruh terhadap status pemeriksaan kanker payudara pada taraf signifikan 5% dengan odds ratio 5,98. Gejala penarikan puting susu berpengaruh terhadap kanker payudara pada taraf signifikan 10% dengan odds ratio 6,13 menggunakan seleksi backward, forward, dan stepwise. Sementara itu, genetic algorithm menghasilkan gejala yang berpengaruh terhadap kanker payudara yaitu perubahan bentuk dan ukuran payudara pada tingkat signifikan 12% dengan odds ratio 4,19.*

***Kata Kunci: Durasi Tidur, Gejala Klinis, Genetic Algorithm, Kanker Payudara, Regresi Logistik Biner***

*(Halaman ini sengaja dikosongkan)*

# **BREAST CANCER CLINICAL SYMPTOMS ANALYSIS USING BINARY LOGISTIC REGRESSION-GENETIC ALGORITHM IN NUR HIDAYAH HOSPITAL BANTUL**

**Name** : Dinda Galuh Guminta  
**Student Number** : 062115 4000 0093  
**Departement** : Statistics  
**Supervisor** : Dr. Santi Wulan Purnami, S.Si., M.Si.

## **Abstract**

*Breast cancer is a disease caused by the uncontrolled growth of abnormal cells in breast tissue. This cancer is the most frequent cancer and the first leading cause of cancer death among women in Indonesia. A majority of women with breast cancer are diagnosed at an advanced stage which causes their survival rates to be low, whereas breast cancer can be detected through early diagnosis. Early diagnosis can be done by recognizing the clinical symptoms of breast cancer such as breast lump or a lump under the armpit, change the shape and size of the breast, nipple retraction, and so on. Besides that, sleep disorders also need to be considered as a risk of breast cancer. Sleep disorders were measured by the sleep duration of the respondent. Therefore, this research was conducted to analyze breast cancer symptoms and sleep disorders using binary logistic regression with variable selection using backward, forward, stepwise, and genetic algorithm. This study showed that all selection methods produce breast pain as a significant variable of breast cancer examination ( $\alpha=5\%$ ) with odds ratio of 5,98. Backward, forward, and stepwise selection showed that nipple retraction whose correlation with the outcome is statistically significant ( $\alpha=10\%$ ) and provide odds ratio of 6,13. Meanwhile, genetic algorithm showed that change shape and size of the breast have a statistically significant correlation with breast cancer examination ( $\alpha=12\%$ ) and odds ratio of 4,19.*

**Keywords:** *Binary Logistic Regression, Breast Cancer, Clinical symptoms, Genetic Algorithm, Sleep Duration*

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**Analisis Gejala Klinis Kanker Payudara Menggunakan Regresi Logistik Biner-*Genetic Algorithm* di RS Nur Hidayah Bantul**” dengan lancar dan tepat waktu.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan dengan baik, tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada :

1. Orang tua dan keluarga yang selalu memberikan dukungan, nasihat, dan doa.
2. Dr. Santi Wulan Purnami, S.Si., M.Si. selaku dosen pembimbing sekaligus Ketua Program Studi Sarjana yang telah memberikan pengarahan dan bimbingan selama proses penulisan laporan Tugas Akhir.
3. Dr. Bambang Widjanarko Otok, M.Si. dan Dr. Purhadi, M.Sc. selaku dosen penguji yang telah memberikan saran dalam penyelesaian laporan Tugas Akhir.
4. dr. Azzam Hizburrahman dan segenap pihak rumah terapi Nur Hidayah Bantul yang turut membantu.
5. Dr. Suhartono selaku Kepala Departemen yang telah memberikan arahan dan dukungan selama perkuliahan di Departemen Statistika ITS.
6. Dr. rer. pol Dedy Dwi Prastyo, S.Si., M.Si. selaku dosen wali yang telah memberikan saran, motivasi, dan arahan selama proses belajar di Departemen Statistika ITS.
7. Seluruh dosen Departemen Statistika ITS yang telah memberikan ilmu dan pengajaran kepada penulis serta segenap karyawan Departemen Statistika ITS yang telah membantu kelancaran dalam proses perkuliahan.
8. Teman-teman Vivacious, Statistika ITS angkatan 2015, yang telah membantu dan mendukung selama masa perkuliahan berlangsung hingga proses penyelesaian Tugas Akhir.

9. Pihak-pihak lain yang tidak bisa penulis sebutkan satu persatu yang turut membantu dalam penulisan laporan Tugas Akhir.

Terlepas dari semua itu, penulis menyadari laporan Tugas Akhir ini masih jauh dari sempurna. Penulis berharap laporan ini dapat bermanfaat bagi masyarakat dan ilmu pengetahuan. Saran dan kritik yang bersifat membangun akan sangat berguna sebagai bahan perbaikan penulisan selanjutnya.

Surabaya, Juli 2019

Penulis

# DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>TITLE PAGE</b> .....	iii
<b>LEMBAR PENGESAHAN</b> .....	v
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR LAMPIRAN</b> .....	xix
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan.....	6
1.4 Manfaat.....	6
1.5 Batasan Masalah.....	7
<b>BAB II TINJAUAN PUSTAKA</b> .....	9
2.1 Tabel Kontingensi.....	9
2.2 Regresi Logistik Biner.....	9
2.2.1 Estimasi Parameter.....	10
2.2.2 Uji Signifikansi Parameter.....	13
2.2.3 Uji Kesesuaian Model.....	18
2.2.4 Interpretasi Model.....	18
2.2.5 <i>Akaike's Information Criterion</i> .....	19
2.3 <i>Genetic Algorithm</i> .....	19
2.4 Evaluasi Metode.....	23
2.5 Gejala Kanker Payudara.....	24
<b>BAB III METODOLOGI PENELITIAN</b> .....	27
3.1 Sumber Data.....	27
3.2 Desain Penelitian.....	27
3.3 Kerangka Konsep.....	28
3.4 Variabel Penelitian.....	28

3.5	Struktur Data .....	31
3.6	Langkah Analisis .....	31
3.7	Diagram Alir.....	33
<b>BAB IV</b>	<b>ANALISIS DAN PEMBAHASAN.....</b>	<b>35</b>
4.1	Karakteristik Gejala Berdasarkan Status Pemeriksaan Kanker Payudara .....	35
4.2	Regresi Logistik Biner-Seleksi <i>Backward</i> .....	45
4.3	Regresi Logistik Biner-Seleksi <i>Forward</i> .....	50
4.4	Regresi Logistik Biner-Seleksi <i>Stepwise</i> .....	52
4.5	Regresi Logistik Biner- <i>Genetic Algorithm</i> .....	54
4.6	Evaluasi Metode Terbaik.....	61
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN.....</b>	<b>63</b>
5.1	Kesimpulan.....	63
5.2	Saran.....	64
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>65</b>
<b>LAMPIRAN</b>	<b>.....</b>	<b>69</b>
<b>BIODATA PENULIS</b>	<b>.....</b>	<b>87</b>

## DAFTAR TABEL

		Halaman
<b>Tabel 3.1</b>	Variabel Penelitian .....	29
<b>Tabel 4.1</b>	Tabulasi Silang Status Kanker Payudara dan Cekungan Pada Kulit Payudara .....	36
<b>Tabel 4.2</b>	Tabulasi Silang Status Kanker Payudara dan Nyeri di Bagian Tertentu Payudara.....	37
<b>Tabel 4.3</b>	Tabulasi Silang Status Kanker Payudara dan Kerutan di Sekitar Payudara .....	38
<b>Tabel 4.4</b>	Tabulasi Silang Status Kanker Payudara dan Benjolan di Bawah Ketiak .....	39
<b>Tabel 4.5</b>	Tabulasi Silang Status Kanker Payudara dan Perubahan Bentuk dan Ukuran Payudara .....	40
<b>Tabel 4.6</b>	Tabulasi Silang Status Kanker Payudara dan Keluar Cairan dari Puting Susu.....	41
<b>Tabel 4.7</b>	Tabulasi Silang Status Kanker Payudara dan Penarikan Puting Susu .....	42
<b>Tabel 4.8</b>	Estimasi Parameter Regresi Logistik Biner .....	46
<b>Tabel 4.9</b>	Seleksi Variabel <i>Backward</i> .....	47
<b>Tabel 4.10</b>	Estimasi Parameter Regresi Logistik Biner dengan Seleksi <i>Backward</i> .....	48
<b>Tabel 4.11</b>	<i>Odds Ratio</i> Regresi Logistik dengan Seleksi <i>Backward</i> .....	49
<b>Tabel 4.12</b>	<i>Confusion Matrix</i> Regresi Logistik dengan Seleksi <i>Backward</i> .....	50
<b>Tabel 4.13</b>	Seleksi Variabel <i>Forward</i> .....	50
<b>Tabel 4.14</b>	Estimasi Parameter Regresi Logistik Biner dengan Seleksi <i>Forward</i> .....	50
<b>Tabel 4.15</b>	<i>Odds Ratio</i> Regresi Logistik dengan Seleksi <i>Forward</i> .....	52
<b>Tabel 4.16</b>	<i>Confusion Matrix</i> Regresi Logistik dengan Seleksi <i>Forward</i> .....	52
<b>Tabel 4.17</b>	Estimasi Parameter Regresi Logistik Biner dengan Seleksi <i>Stepwise</i> .....	53

<b>Tabel 4.18</b>	<i>Odds Ratio</i> Regresi Logistik dengan Seleksi Stepwise .....	54
<b>Tabel 4.19</b>	<i>Confusion Matrix</i> Regresi Logistik dengan Seleksi Stepwise.....	54
<b>Tabel 4.20</b>	Ilustrasi Kromosom Pada Populasi Awal .....	55
<b>Tabel 4.21</b>	Ilustrasi Nilai <i>Fitness</i> Tiap Kromosom .....	56
<b>Tabel 4.22</b>	Ilustrasi Seleksi <i>Roulette Wheel</i> .....	56
<b>Tabel 4.23</b>	Estimasi Parameter Regresi Logistik Biner- <i>Genetic Algorithm</i> .....	58
<b>Tabel 4.24</b>	Pemodelan Kembali Regresi Logistik Biner- <i>Genetic Algorithm</i> .....	59
<b>Tabel 4.25</b>	<i>Odds Ratio</i> Regresi Logistik Biner- <i>Genetic Algorithm</i> .....	60
<b>Tabel 4.26</b>	<i>Confusion Matrix</i> Regresi Logistik Biner- <i>Genetic Algorithm</i> .....	60
<b>Tabel 4.27</b>	Perbandingan Kebaikan Metode Seleksi .....	61

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Representasi Gen, Kromosom, dan Individu .....	19
<b>Gambar 2.2</b> Ilustrasi <i>Single Point Crossover</i> .....	22
<b>Gambar 2.3</b> Ilustrasi Mutasi.....	22
<b>Gambar 3.1</b> Kerangka Konsep Risiko Prevalensi Kanker Payudara.....	28
<b>Gambar 3.2</b> Diagram Alir <i>Genetic Algorithm</i> .....	34
<b>Gambar 4.1</b> Persentase Status Pemeriksaan Kanker Payudara	35
<b>Gambar 4.2</b> Persentase Responden dengan Gejala Cekungan Pada Kulit Payudara.....	36
<b>Gambar 4.3</b> Persentase Responden dengan Gejala Nyeri di Bagian Tertentu Payudara .....	37
<b>Gambar 4.4</b> Persentase Responden dengan Gejala Kerutan di Sekitar Payudara.....	38
<b>Gambar 4.5</b> Persentase Responden dengan Gejala Benjolan di Bawah Ketiak .....	39
<b>Gambar 4.6</b> Persentase Responden dengan Gejala Perubahan Bentuk dan Ukuran Payudara.....	40
<b>Gambar 4.7</b> Persentase Responden dengan Gejala Keluar Cairan dari Puting Susu.....	41
<b>Gambar 4.8</b> Persentase Responden dengan Gejala Penarikan Puting Susu.....	42
<b>Gambar 4.9</b> Perbandingan Gejala Klinis pada Status Pemeriksaan Kanker Payudara.....	43
<b>Gambar 4.10</b> Perbandingan Usia Responden Berdasarkan Status Pemeriksaan Kanker Payudara .....	44
<b>Gambar 4.11</b> Perbandingan Status Pemeriksaan Kanker Payudara-Durasi Tidur .....	45
<b>Gambar 4.12</b> Ilustrasi Kromosom Awal .....	55
<b>Gambar 4.13</b> Ilustrasi Proses <i>Crossover</i> .....	57
<b>Gambar 4.14</b> Ilustrasi Proses Mutasi .....	58

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1.</b> Data Penelitian .....	69
<b>Lampiran 2.</b> <i>Syntax</i> R Regresi Logistik Biner .....	70
<b>Lampiran 3.</b> <i>Syntax</i> R <i>Genetic Algorithm</i> .....	71
<b>Lampiran 4.</b> Tahapan Seleksi <i>Backward</i> .....	72
<b>Lampiran 5.</b> Tahapan Pemodelan dengan Seleksi <i>Backward</i> .....	74
<b>Lampiran 6.</b> Tahapan Seleksi <i>Forward</i> .....	75
<b>Lampiran 7.</b> Pemodelan dengan Seleksi <i>Forward</i> .....	76
<b>Lampiran 8.</b> Tahapan Seleksi <i>Stepwise</i> .....	77
<b>Lampiran 9.</b> Pemodelan dengan Seleksi <i>Stepwise</i> .....	78
<b>Lampiran 10.</b> Perhitungan Probabilitas Tiap Responden .....	79
<b>Lampiran 11.</b> Hasil <i>Genetic Algorithm</i> .....	80
<b>Lampiran 12.</b> Pemodelan dengan <i>Genetic Algorithm</i> .....	80
<b>Lampiran 13.</b> Pemodelan Kembali dengan <i>Genetic Algorithm</i> .....	81
<b>Lampiran 14.</b> Perhitungan Probabilitas pada Model dengan <i>Genetic Algorithm</i> .....	82
<b>Lampiran 15.</b> Contoh Hasil Tes Pemeriksaan Responden.....	83
<b>Lampiran 16.</b> Surat Izin Pengambilan Data.....	84
<b>Lampiran 17.</b> Surat Pernyataan Data.....	85
<b>Lampiran 18.</b> Surat Keterangan .....	86

*(Halaman ini sengaja dikosongkan)*

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penyakit kanker saat ini masih menjadi salah satu permasalahan global dan isu internasional penting di bidang kesehatan. Berdasarkan data Globocan pada tahun 2018, sebanyak 9,6 dari 18,1 juta jiwa di dunia meninggal akibat kanker. Angka tersebut mengalami peningkatan yang cukup tinggi jika dibandingkan pada tahun 2008 dengan jumlah kematian sebesar 7,6 juta dari 12,7 juta jiwa yang mengidap penyakit kanker (IARC, 2018). Indonesia menempati peringkat ke-8 dunia dan peringkat pertama di antara negara-negara Asia Tenggara untuk perkiraan jumlah kejadian kanker dan kematian akibat kanker berturut-turut sebesar 313.673 kejadian dan 174.508 kematian (IARC, 2018). Kanker payudara menjadi jenis kanker dengan jumlah kejadian tertinggi di Indonesia yaitu 55.424 (18 persen dari total kejadian kanker) dan 20.473 kematian (12 persen dari total kematian) (IARC, 2018). Penyakit seluler yang ditandai oleh pertumbuhan sel abnormal pada jaringan payudara ini tumbuh secara destruktif dan tidak terkendali. Sel-sel tersebut tumbuh dan melakukan invasi ke jaringan sekitar seperti kelenjar getah bening, pembuluh darah, tulang, paru-paru, hati hingga otak sehingga menyebabkan kegagalan fungsi organ-organ tersebut yang dapat berujung kematian (Rasjidi, 2009).

Proses terbentuknya sel kanker merupakan proses yang sangat kompleks. Satu sel kanker harus melepaskan diri dari kelompoknya (*primary tumor*) untuk melakukan invasi ke daerah sekitarnya, sel berusaha menembus pembuluh darah dan melawan proses pertahanan tubuh, hingga berhenti pada organ yang dituju untuk memulai berkembang biak di lingkungan barunya (*secondary tumor*) (Fidler, 1990). Kanker payudara merupakan jenis kanker yang paling umum di diagnosis pada wanita. Penyakit ini juga dapat menyerang laki-laki dengan frekuensi sekitar satu persen. Disamping itu, sebanyak 42,7 dari 100.000 penduduk di

Indonesia terkena penyakit kanker payudara dengan angka kematian sebesar 15,8 dari 100.000 penduduk di Indonesia (IARC, 2018). Kanker payudara saat ini tidak hanya menyerang perempuan usia tua sekitar lebih dari 50 tahun, perempuan usia produktif juga banyak mengidap kanker payudara. Disamping itu, kejadian kanker payudara banyak terjadi pada populasi wanita yang mengalami menopause. Sebanyak 78 persen pasien kanker payudara berusia lebih dari 50 tahun dan hanya 6 persen pasien kanker payudara berusia kurang dari 40 tahun. Hal ini menunjukkan bahwa usia menjadi salah satu faktor risiko kejadian kanker payudara. Wanita yang menopause setelah usia 55 tahun mempunyai risiko kanker payudara dua kali lebih tinggi dibandingkan wanita yang menopause sebelum usia 45 tahun (Sabiston, 1995).

Pada umumnya, wanita yang menderita kanker payudara mengalami keluhan seperti benjolan di payudara yang mengalami pertumbuhan dengan/tanpa rasa sakit, payudara mengalami perubahan bentuk dan ukuran akibat pembengkakan, penarikan puting susu dan krusta, mulai timbul luka pada payudara dan puting susu seperti koreng, kulit payudara menjadi berkerut seperti kulit jeruk, kelainan kulit, benjolan di ketiak dan edema lengan serta keluhan tambahan seperti keluar cairan atau darah berwarna merah kehitam-hitaman dari puting susu, nyeri tulang, sesak dan lain sebagainya (Komite Penanggulangan Kanker Nasional, 2017). Para wanita di Rajasthan, India sadar bahwa benjolan pada payudara, perubahan bentuk atau ukuran, puting kemerahan, nyeri pada payudara adalah gejala-gejala penyakit kanker payudara. Namun, gejala lain seperti benjolan di bawah ketiak dan perubahan posisi puting susu kurang banyak diketahui sebagai gejala kanker payudara (Yadav & Jaroli, 2010). Sebuah studi mengenai gejala kanker payudara mengatakan bahwa terdapat perbedaan rasa nyeri antara penderita dan bukan penderita kanker payudara pada umur 41 tahun ke atas (Lumachi, et al., 2002).

Penelitian serupa menunjukkan empat gejala yang berhubungan dengan kanker payudara yaitu benjolan pada

payudara atau ketiak, nyeri pada payudara, keluar cairan dari puting susu, dan penarikan puting susu (Walker, Hyde, & Hamilton, 2014). Penelitian terkait keluhan kanker payudara menyebutkan bahwa keluhan gejala kanker payudara paling tinggi terjadi pada wanita berusia 25-44 tahun dan wanita berusia 65 tahun ke atas. Benjolan payudara tercatat menjadi gejala dengan tingkat diagnosis kanker payudara tertinggi (Eberl, et al., 2008). Kemudian, kejadian kanker payudara banyak terjadi pada pasien berusia 50 tahun ke atas dengan gejala keluar cairan selain air susu dari puting susu dan pasien berusia 70 tahun ke atas dengan gejala nyeri (Seltzer, 2004). Penelitian lain juga menunjukkan bahwa benjolan payudara dan perubahan kulit/puting susu berhubungan meningkatkan risiko kanker payudara (Ryerson, Miller, & Ehem, 2015).

Pada penelitian terdahulu juga menjelaskan bahwa benjolan, perubahan kulit/puting susu, dan nyeri menjadi gejala yang dapat mengindikasikan kejadian kanker payudara (Barton, Elmore, & Fletcher, 1999). Disamping itu, diperoleh persentase gejala yang paling sering dirasakan responden yaitu benjolan tanpa rasa nyeri, benjolan di bawah ketiak, dan perubahan ukuran payudara (Grunfeld, Ramirez, & Richards, 2002). Para wanita seringkali melakukan pemeriksaan kanker payudara dengan hasil diagnosa pada stadium lanjut yang menyebabkan angka harapan hidup penderita menjadi rendah karena kurang penyakit yang di derita sudah terlalu berbahaya. Padahal kanker payudara dapat diketahui melalui deteksi dini dengan mengenali gejala-gejala klinis yang menandakan adanya kanker. Sebanyak 60 hingga 70 persen penderita kanker payudara pertama kali melakukan pengobatan ketika telah berada dalam stadium lanjut (III dan IV). Kemudian, penderita yang berobat ke Rumah Sakit Kanker Dharmas Jakarta tiga kali lebih banyak pada kondisi metastasis jauh (Suzanna, et al., 2012).

Tidur adalah fenomena biologis dan keadaan alami yang terjadi berulang untuk memenuhi kebutuhan fisiologis dan psikologis yang penting bagi manusia. Namun, dari waktu ke

waktu, prevalensi gangguan tidur telah meningkat di seluruh dunia (Fang, et al., 2015). Negara Jepang, memiliki angka prevalensi gangguan tidur sebesar 23,5 persen dari total populasi Jepang dan hasil serupa dilaporkan untuk populasi Asia (Kaneita, et al., 2006). Gangguan tidur dapat berupa durasi tidur pendek atau gangguan seperti kualitas tidur buruk. Efisiensi tidur yang rendah dan banyak gangguan tidur secara signifikan menjadi faktor risiko independen pada wanita dengan kanker payudara stadium lanjut (Paresh, et al., 2014). Studi antara durasi tidur dan risiko kanker payudara menunjukkan bahwa risiko kanker payudara menurun dengan meningkatnya durasi tidur (Wu, et al., 2008). Sebuah penelitian mengatakan sebanyak 56 persen pasien kanker paru-paru menderita gangguan tidur dan 60 persen mengalami tekanan psikologis (Nishiura, et al., 2015). Pasien dengan gangguan tidur menunjukkan risiko kanker payudara yang lebih tinggi dibandingkan dengan pasien tanpa gangguan tidur (Fang, et al., 2015).

Kondisi ini menandakan bahwa permasalahan kanker payudara di Indonesia membutuhkan perhatian yang lebih serius. Kondisi telat diagnosis akan mempersulit proses penyembuhan dan kemungkinan untuk sembuh sangat kecil. Apabila kasus kanker payudara dapat diketahui secara dini maka pengobatan yang dilakukan dapat lebih mudah. Oleh karena itu, pencegahan primer dan sekunder dapat dilakukan untuk meningkatkan kewaspadaan terhadap kanker payudara seperti Pemeriksaan Payudara Sendiri (SADARI) yang dilakukan secara mandiri oleh setiap perempuan dan Pemeriksaan Payudara Klinis (SADANIS) oleh tenaga kesehatan terlatih. Tercatat masing-masing sebesar 53,7 persen dan 95,6 persen masyarakat tidak pernah melakukan SADARI dan SADANIS. Pemeriksaan ini bertujuan untuk menemukan benjolan dan tanda-tanda abnormal pada payudara sedini mungkin. SADARI dan SADANIS dapat dilakukan setiap bulan pada hari ke-7 hingga ke-10 terhitung dari hari pertama haid; atau pada tanggal yang sama setiap bulan bagi perempuan yang sudah menopause (KEMENKES, 2017). Disamping itu, pengobatan

kanker sangat tergantung pada jenis dan tingkat penyebarannya. Ada beberapa jenis pengobatan pada pasien kanker payudara diantaranya yaitu, operasi, kemoterapi, dan terapi herbal (Shaharudin, et al., 2011).

Berdasarkan uraian sebelumnya maka penelitian ini dilakukan untuk menganalisis pengaruh gejala klinis kanker payudara terhadap hasil tes pemeriksaan kanker payudara di RS Nur Hidayah, Bantul, Yogyakarta. Penelitian mengenai klasifikasi penderita kanker payudara dan bukan kanker payudara berdasarkan faktor risiko dilakukan oleh Purwantaka (2010) menggunakan regresi logistik biner, menunjukkan bahwa usia dan usia pertama kali menstruasi berpengaruh signifikan terhadap hasil klasifikasi dengan akurasi sebesar 63 persen (Purwantaka, 2010). Johnson (2014) menerapkan *genetic algorithm* pada regresi logistik untuk memprediksi progress penyakit alzheimer, menunjukkan bahwa sekumpulan kecil variabel memiliki kinerja yang lebih unggul daripada semua variabel tunggal yang signifikan dalam model untuk prediksi perkembangan penyakit (Johnson, et al., 2014). Gayou (2008) menyimpulkan bahwa *genetic algorithm* memberikan cara yang handal dan cepat dalam pemilihan variabel untuk memilih faktor-faktor penting terhadap hasil pengobatan radioterapi menggunakan regresi logistik (Gayou, et al., 2008). Oleh karena itu, penelitian ini dilakukan menggunakan regresi logistik biner dengan seleksi variabel *genetic algorithm*. Regresi logistik merupakan metode yang mudah diaplikasikan karena variabel penjelas atau prediktor tidak harus memiliki distribusi normal, linier, maupun memiliki varian yang sama dalam setiap grup (identik). Penelitian ini diharapkan dapat menjadi *early diagnosis* terhadap diagnosa kanker payudara sehingga dapat diberikan perawatan pada tahap sedini mungkin.

## **1.2 Rumusan Masalah**

Kanker payudara menjadi jenis kanker penyebab kematian tertinggi wanita di Indonesia. Faktor penyebab kanker payudara yang belum pasti, menyebabkan jumlah kejadian kanker payudara semakin bertambah. Kondisi telat diagnosis penyakit kanker

payudara akan berdampak pada telatnya penanganan dan perawatan sehingga pengobatan yang diberikan menjadi kurang maksimal, bahkan dapat berujung pada kematian. Oleh karena itu, penelitian ini menarik permasalahan mengenai gejala klinis yang mempengaruhi hasil pemeriksaan kanker payudara di RS Nur Hidayah Bantul. Gejala-gejala yang dianggap berpengaruh terhadap penyakit kanker payudara dapat dimodelkan menggunakan regresi logistik biner dengan seleksi variabel *backward*, *forward*, *stepwise*, dan *genetic algorithm*. Kemudian, dilakukan perbandingan tingkat ketepatan klasifikasi dari metode tersebut.

### 1.3 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan, tujuan dari penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik gejala klinis kanker payudara pada responden di RS Nur Hidayah Bantul.
2. Mendapatkan gejala-gejala klinis yang berpengaruh signifikan pada status pemeriksaan kanker payudara di RS Nur Hidayah Bantul menggunakan regresi logistik biner dengan seleksi variabel *backward*, *forward*, *stepwise*, dan *genetic algorithm*.
3. Membandingkan model terbaik untuk status pemeriksaan kanker payudara di RS Nur Hidayah Bantul menggunakan regresi logistik biner dengan seleksi variabel *backward*, *forward*, *stepwise*, dan *genetic algorithm*.

### 1.4 Manfaat

Penelitian ini dapat bermanfaat sebagai *early diagnosis* terhadap kejadian kanker payudara sehingga perawatan yang diperoleh dapat bekerja secara efektif dan memberikan hasil yang optimal. Deteksi dini dengan mengenali gejala klinis peringatan kanker dapat menjadi strategi yang penting dalam semua pengaturan kesehatan masyarakat. Bagi pihak rumah sakit, dapat memberikan informasi mengenai gejala-gejala yang berpengaruh terhadap kejadian kanker payudara terhadap para pasien. Dalam

bidang keilmuan, diharapkan berguna sebagai pengembangan ilmu pengetahuan, khususnya *genetic algorithm* dalam bidang kesehatan.

### **1.5 Batasan Masalah**

Pada penelitian ini digunakan data gejala-gejala klinis pada responden berjenis kelamin perempuan di Rumah Terapi Tauhid RS Nur Hidayah Bantul, Yogyakarta. Metode yang digunakan adalah regresi logistik biner dengan seleksi variabel *backward*, *forward*, *stepwise*, dan *genetic algorithm*.

*(Halaman ini sengaja dikosongkan)*

## BAB II TINJAUAN PUSTAKA

### 2.1 Tabel Kontingensi

Dimisalkan terdapat dua variabel berskala kategori yang dinotasikan  $A$  dan  $B$ . Variabel  $A$  terdiri dari  $I$  level dan variabel  $B$  terdiri dari  $J$  level. Kemudian dibentuk tabel berukuran  $I$  baris untuk level variabel  $A$  dan  $J$  kolom untuk level variabel  $B$  sehingga diperoleh sebanyak  $IJ$  kombinasi.

**Tabel 2.1** Tabel Kontingensi  $I \times J$

	$B_1$	$B_2$	...	$B_j$	...	$B_J$	Total
$A_1$	$n_{11}$	$n_{12}$	...	$n_{1j}$	...	$n_{1J}$	$n_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$A_i$	$n_{i1}$	$n_{i2}$	...	$n_{ij}$	...	$n_{iJ}$	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$A_I$	$n_{I1}$	$n_{I2}$	...	$n_{Ij}$	...	$n_{IJ}$	$n_{I.}$
<b>Total</b>	$n_{.1}$	$n_{.2}$		$n_{.j}$		$n_{.J}$	$n_{..}$

**Tabel 2.2** Tabel Kontingensi  $I \times J$  untuk Probabilitas

	$B_1$	$B_2$	...	$B_j$	...	$B_J$	Total
$A_1$	$P_{11}$	$P_{12}$	...	$P_{1j}$	...	$P_{1J}$	$P_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$A_i$	$P_{i1}$	$P_{i2}$	...	$P_{ij}$	...	$P_{iJ}$	$P_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$A_I$	$P_{I1}$	$P_{I2}$	...	$P_{Ij}$	...	$P_{IJ}$	$P_{I.}$
<b>Total</b>	$P_{.1}$	$P_{.2}$		$P_{.j}$		$P_{.J}$	$P_{..}$

Bentuk tabel yang menampilkan jumlah *outcomes* dalam sel disebut tabel kontingensi (Agresti, 2013). Tabel 2.1 merupakan tabel kontingensi  $I \times J$  dengan  $n$  adalah total observasi pada sel ke  $ij$ ,  $i = 1, 2, \dots, I$  dan  $j = 1, 2, \dots, J$ .

### 2.2 Regresi Logistik Biner

Analisis regresi logistik merupakan salah satu metode statistika yang bertujuan untuk menjelaskan hubungan antara variabel respon ( $Y$ ) yang berskala kategori dengan satu atau lebih variabel prediktor ( $X$ ) yang berskala kategori atau kontinu. Apabila variabel respon memiliki dua ( $k = 2$ ) level yang terdiri dari kejadian

sukses ( $y = 1$ ) dan kejadian gagal ( $y = 0$ ) maka digunakan regresi logistik biner. Regresi logistik biner merupakan metode yang menjelaskan hubungan antara variabel respon yang bersifat biner atau dikotomous dengan variabel prediktor (Agresti, 2013). Variabel respon tersebut mengikuti distribusi Bernoulli untuk setiap variabel tunggal dengan fungsi probabilitas adalah sebagai berikut.

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

Apabila  $y_i=0$  maka  $f(y_i) = (1 - \pi)$  atau  $y_i=1$  maka  $f(y_i) = \pi$ . Bentuk persamaan logit adalah sebagai berikut.

$$\text{Logit}(\pi(\mathbf{x}_i)) = \ln\left(\frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)}\right) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (2.2)$$

Dilakukan transformasi logit  $\pi(\mathbf{x}_i)$  agar diperoleh fungsi yang linier dalam parameter sehingga model probabilitas yang terbentuk adalah sebagai berikut.

$$\begin{aligned} \pi(\mathbf{x}_i) &= \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + \dots + \beta_p x_{ip})} \\ &= \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})} \end{aligned} \quad (2.3)$$

dengan  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$  adalah nilai pengamatan ke- $i$  ( $i = 1, 2, \dots, n$ ) dari setiap  $p$  ( $j = 1, 2, \dots, p$ ) variabel prediktor. Model regresi yang terbentuk bila terdapat  $k = 2$  kategori respon adalah sejumlah  $k-1$ .

### 2.2.1 Estimasi Parameter

Estimasi parameter model *logit* dilakukan menggunakan *Maximum Likelihood Estimation* (MLE) yaitu memaksimalkan fungsi *likelihood* sehingga estimasi parameter menjadi penduga yang konsisten dan efisien untuk ukuran sampel yang besar. Pada regresi logistik biner, setiap pengamatan mengikuti distribusi Bernoulli sehingga diperoleh fungsi *likelihood* sebagai berikut.

$$f(y_i) = \pi(\mathbf{x}_i)^{y_i} (1 - \pi(\mathbf{x}_i))^{1-y_i}; y_i = 0,1 \quad (2.4)$$

Jika  $X_i$  dan  $Y_i$  adalah pasangan variabel prediktor dan respon pada pengamatan ke- $i$  yang diasumsikan saling independen pada tiap

pasangan pengamatan sehingga fungsi *likelihood* yang diperoleh merupakan gabungan dari fungsi distribusi masing-masing pasangan. Apabila diambil sampel secara acak sebanyak  $n$  dengan  $i = 1, 2, \dots, n$ , maka fungsi *likelihood* dinyatakan dalam persamaan (2.5) (Hosmer & Lemeshow, 2000).

$$\begin{aligned} l(\boldsymbol{\beta}) &= \prod_{i=1}^n f(y_i) \\ &= \prod_{i=1}^n \pi(\mathbf{x}_i)^{y_i} (1 - \pi(\mathbf{x}_i))^{1-y_i} \end{aligned} \quad (2.5)$$

Kemudian, persamaan (2.5) dimaksimumkan dalam bentuk fungsi *ln-likelihood* sehingga diperoleh persamaan (2.6) sebagai berikut.

$$\begin{aligned} L(\boldsymbol{\beta}) = \ln(l(\boldsymbol{\beta})) &= \ln\left(\prod_{i=1}^n \pi(\mathbf{x}_i)^{y_i} (1 - \pi(\mathbf{x}_i))^{1-y_i}\right) \\ &= \ln\left(\prod_{i=1}^n \left(\frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)}\right)^{y_i} (1 - \pi(\mathbf{x}_i))\right) \\ &= \sum_{i=1}^n \left[ y_i \ln\left(\frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)}\right) + \ln(1 - \pi(\mathbf{x}_i)) \right] \\ &= \sum_{i=1}^n \left[ y_i (\mathbf{x}'_i \boldsymbol{\beta}) + \ln\left(\frac{1}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})}\right) \right] \\ &= \sum_{i=1}^n [y_i (\mathbf{x}'_i \boldsymbol{\beta}) - \ln(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))] \end{aligned} \quad (2.6)$$

dengan  $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \dots \ \beta_p]'$  matriks berukuran  $(p + 1) \times 1$  dan

$\mathbf{x}_i = [1 \ x_{i1} \ x_{i2} \ \dots \ x_{ip}]'$  matriks berukuran  $(p + 1) \times 1$ .

Nilai  $\boldsymbol{\beta}$  maksimum didapatkan melalui turunan  $L(\boldsymbol{\beta})$  terhadap  $\boldsymbol{\beta}$ ,

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} = \frac{\partial \sum_{i=1}^n [y_i (\mathbf{x}'_i \boldsymbol{\beta}) - \ln(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))]}{\partial \boldsymbol{\beta}'}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left[ y_i \mathbf{x}_i - \frac{\mathbf{x}_i \exp(\mathbf{x}'_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})} \right] \\
&= \sum_{i=1}^n [y_i - \pi(\mathbf{x}_i)] \mathbf{x}_i \\
&= \mathbf{X}'(\mathbf{Y} - \boldsymbol{\pi})
\end{aligned} \tag{2.7}$$

dengan

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \text{ serta } \boldsymbol{\pi} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_n \end{bmatrix}.$$

Setelah mendapatkan nilai  $\boldsymbol{\beta}$  yang menyebabkan nilai dari fungsi *likelihood* bernilai ekstrim melalui turunan pertama  $L(\boldsymbol{\beta})$  terhadap  $\boldsymbol{\beta}$ . Berikut penurunan kedua yang ditunjukkan pada persamaan (2.8).

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}' \partial \boldsymbol{\beta}} &= \frac{\partial}{\partial \boldsymbol{\beta}} \left( \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} \right) \\
&= \frac{\partial}{\partial \boldsymbol{\beta}} \left( \sum_{i=1}^n \left[ y_i \mathbf{x}_i - \frac{\mathbf{x}_i \exp(\mathbf{x}'_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})} \right] \right) \\
&= - \sum_{i=1}^n \mathbf{x}_i \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})) - \exp(\mathbf{x}'_i \boldsymbol{\beta})}{(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))^2} \mathbf{x}'_i \\
&= - \sum_{i=1}^n \mathbf{x}_i \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))^2} \mathbf{x}'_i \\
&= - \sum_{i=1}^n \mathbf{x}_i \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))^2} \frac{1}{(1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))^2} \mathbf{x}'_i \\
&= - \sum_{i=1}^n \mathbf{x}_i \pi(\mathbf{x}_i)(1 - \pi(\mathbf{x}_i)) \mathbf{x}'_i \\
&= -\mathbf{X}'\mathbf{W}\mathbf{X}
\end{aligned} \tag{2.8}$$

dengan  $\mathbf{W} = \text{diag}(\hat{\pi}(\mathbf{x}_i)(1 - \hat{\pi}(\mathbf{x}_i)))$ .

Persamaan (2.8) menghasilkan bentuk yang tidak *closed-form* sehingga dilakukan iterasi numerik untuk memperoleh  $\boldsymbol{\beta}$  yang konvergen. Metode iterasi *Newton Raphson* digunakan untuk menyelesaikan persamaan yang tidak linier. Berikut rumus iterasi *Newton Raphson*.

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \mathbf{H}^{-1}(\boldsymbol{\beta}^{(t)})\mathbf{g}(\boldsymbol{\beta}^{(t)}); t = 0, 1, 2, \dots \quad (2.9)$$

dengan

$$\mathbf{g}(\boldsymbol{\beta}) = \begin{pmatrix} \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_0} \\ \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1} \\ \vdots \\ \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_p} \end{pmatrix} \text{ dan } \mathbf{H}(\boldsymbol{\beta}) = \begin{pmatrix} \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_1^2} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_0} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_1} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_p^2} \end{pmatrix}.$$

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

1. Menentukan nilai awal estimasi parameter  $\hat{\boldsymbol{\beta}}^{(0)}$ ,
2. Membentuk vektor gradien  $\mathbf{g}$  dan matriks *Hessian*  $\mathbf{H}$ ,
3. Memasukkan nilai  $\hat{\boldsymbol{\beta}}^{(0)}$  pada elemen  $\mathbf{g}$  dan  $\mathbf{H}$  sehingga diperoleh  $\mathbf{g}(\hat{\boldsymbol{\beta}}^{(0)})$  dan  $\mathbf{H}(\hat{\boldsymbol{\beta}}^{(0)})$ ,
4. Iterasi dimulai  $t = 0$  menggunakan persamaan (2.9). Nilai  $\hat{\boldsymbol{\beta}}^{(t)}$  adalah sekumpulan penaksir parameter yang konvergen pada iterasi ke-  $t$ , dan
5. Apabila estimasi parameter yang diperoleh belum konvergen, maka langkah (3) diulang kembali hingga nilai  $\|\hat{\boldsymbol{\beta}}^{(t+1)} - \hat{\boldsymbol{\beta}}^{(t)}\| \leq \varepsilon$ , dengan  $\varepsilon$  adalah bilangan yang sangat kecil. Hasil estimasi yang diperoleh adalah  $\hat{\boldsymbol{\beta}}^{(t+1)}$  pada iterasi terakhir.

### 2.2.2 Uji Signifikansi Parameter

Pengujian estimasi parameter bertujuan memeriksa pengaruh signifikansi taksiran parameter terhadap model. Uji parameter secara serentak dilakukan menggunakan *Likelihood Ratio Test* dengan tahapan sebagai berikut.

1. Menentukan himpunan parameter untuk populasi,  $y \sim B(1, p)$

$$\ln \left[ \frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)} \right] = \mathbf{x}_i' \boldsymbol{\beta}$$

$$\pi(\mathbf{x}_i) = \frac{\exp(\mathbf{x}_i' \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i' \boldsymbol{\beta})}$$

$$1 - \pi(\mathbf{x}_i) = \frac{1}{1 + \exp(\mathbf{x}_i' \boldsymbol{\beta})} \quad (2.10)$$

Himpunan parameter dibawah populasi ( $\Omega$ ) yaitu

$$\Omega = \{\boldsymbol{\beta}\} = \{\beta_0, \beta_1, \beta_2, \dots, \beta_p\}.$$

2. Menentukan himpunan parameter dibawah  $H_0(\omega)$ ,

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

maka model berikut,

$$\ln \left[ \frac{\pi}{1 - \pi} \right] = \beta_0 \quad (2.11)$$

$$\omega = \{\beta_0\}$$

tidak mengandung  $x_{i1} + x_{i2} + \dots + x_{ip}$ , sehingga  $y \sim B(1, p)$

$$\ln \left[ \frac{\pi}{1 - \pi} \right] = \beta_0$$

$$\frac{\pi}{1 - \pi} = \exp(\beta_0)$$

$$\pi = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$$

$$1 - \pi = \frac{1}{1 + \exp(\beta_0)} \quad (2.12)$$

3. Membuat fungsi *likelihood* dibawah populasi.

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

$$L(\Omega) = \prod_{i=1}^n \pi(\mathbf{x}_i)^{y_i} (1 - \pi(\mathbf{x}_i))^{1 - y_i}$$

$$= \prod_{i=1}^n \left( \frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)} \right)^{y_i} (1 - \pi(\mathbf{x}_i))$$

$$\begin{aligned}
&= \prod_{i=1}^n (\exp(\mathbf{x}_i' \boldsymbol{\beta}))^{y_i} \left( \frac{1}{1 + \exp(\mathbf{x}_i' \boldsymbol{\beta})} \right) \\
&= \prod_{i=1}^n \exp(y_i \mathbf{x}_i' \boldsymbol{\beta}) (1 + \exp(\mathbf{x}_i' \boldsymbol{\beta}))^{-1} \\
&= \sum_{i=1}^n (y_i \mathbf{x}_i' \boldsymbol{\beta} - \ln(1 + \exp(\mathbf{x}_i' \boldsymbol{\beta}))) \quad (2.13)
\end{aligned}$$

Persamaan (2.13) dimaksimumkan menggunakan MLE untuk menghasilkan estimasi  $\hat{\boldsymbol{\beta}}$  dengan iterasi numerik sehingga diperoleh.

$$\ln L(\hat{\Omega}) = \sum_{i=1}^n (y_i \mathbf{x}_i' \hat{\boldsymbol{\beta}} - \ln(1 + \exp(\mathbf{x}_i' \hat{\boldsymbol{\beta}}))) \quad (2.14)$$

4. Membuat fungsi *likelihood* dibawah  $H_0$ .

$$\begin{aligned}
L(\omega) &= L(\beta_0) \\
&= \prod_{i=1}^n f(y_i) \\
&= \prod_{i=1}^n \pi^{y_i} (1 - \pi)^{1-y_i} \\
&= \prod_{i=1}^n \left( \frac{\pi}{1 - \pi} \right)^{y_i} (1 - \pi) \\
&= \prod_{i=1}^n (\exp(\beta_0))^{y_i} \left( \frac{1}{1 + \exp(\beta_0)} \right) \\
&= \prod_{i=1}^n \exp(\beta_0 y_i) (1 + \exp(\beta_0))^{-1} \\
\ln L(\omega) &= \ln \left[ \prod_{i=1}^n \exp(\beta_0 y_i) (1 + \exp(\beta_0))^{-1} \right] \quad (2.15)
\end{aligned}$$

$$\begin{aligned}
&= \ln \left[ \exp \left( \sum_{i=1}^n \beta_0 y_i \right) (1 + \exp(\beta_0))^{-n} \right] \\
&= \ln \left( \exp \left( \sum_{i=1}^n \beta_0 y_i \right) \right) + \ln(1 + \exp(\beta_0))^{-n} \\
&= \beta_0 \sum_{i=1}^n y_i - n \ln(1 + \exp(\beta_0)) \quad (2.16)
\end{aligned}$$

Nilai  $\hat{\beta}_0$  didapatkan melalui turunan  $\ln L(\omega)$  terhadap  $\beta_0$  sama dengan nol,

$$\begin{aligned}
\frac{\partial \ln L(\omega)}{\partial \beta_0} &= \frac{\partial (\beta_0 \sum_{i=1}^n y_i - n \ln(1 + \exp(\beta_0)))}{\partial \beta_0} \\
&= \sum_{i=1}^n y_i - \frac{n}{1 + \exp(\beta_0)} \frac{\partial (1 + \exp(\beta_0))}{\partial \beta_0} \\
&= \sum_{i=1}^n y_i - \frac{n}{1 + \exp(\beta_0)} \exp(\beta_0) = 0 \\
\sum_{i=1}^n y_i &= \frac{n(\exp(\beta_0))}{1 + \exp(\beta_0)} \\
\frac{\sum_{i=1}^n y_i}{n} &= \frac{\exp(\beta_0)}{1 + \exp(\beta_0)} \\
\bar{y} &= \frac{\exp(\hat{\beta}_0)}{1 + \exp(\hat{\beta}_0)}
\end{aligned}$$

dengan  $\hat{\pi} = \bar{y}$ , maka didapatkan  $\hat{\beta}_0$  sebagai berikut,

$$\begin{aligned}
\exp(\hat{\beta}_0) &= \bar{y}(1 + \exp(\hat{\beta}_0)) \\
\exp(\hat{\beta}_0)(1 - \bar{y}) &= \bar{y} \\
\exp(\hat{\beta}_0) &= \frac{\bar{y}}{1 - \bar{y}} \\
\hat{\beta}_0 &= \ln \left( \frac{\bar{y}}{1 - \bar{y}} \right) \quad (2.17)
\end{aligned}$$

sehingga diperoleh.

$$\begin{aligned}
 \ln l(\hat{\omega}) &= \hat{\beta}_0 \sum_{i=1}^n y_i - n \ln(1 + \exp(\hat{\beta}_0)) \\
 &= \left( \sum_{i=1}^n y_i \right) \ln \left( \frac{\bar{y}}{1 - \bar{y}} \right) - n \ln \left( \frac{1}{1 - \bar{y}} \right) \\
 &= n \bar{y} \ln \left( \frac{\bar{y}}{1 - \bar{y}} \right) + n \ln(1 - \bar{y}) \\
 &= n \bar{y} (\ln \bar{y} - \ln(1 - \bar{y})) + n \ln(1 - \bar{y}) \\
 &= n (\bar{y} \ln \bar{y} + (1 - \bar{y}) \ln(1 - \bar{y})) \\
 &= n (\hat{\pi} \ln \hat{\pi} + (1 - \hat{\pi}) \ln(1 - \hat{\pi})) \quad (2.18)
 \end{aligned}$$

Berikut ini hipotesis dalam uji signifikansi parameter secara serentak.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0, \text{ dan}$$

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

Statistik uji yang digunakan yaitu,

$$\begin{aligned}
 G^2 &= -2 \ln \Lambda = -2 \ln \left( \frac{l(\hat{\omega})}{l(\hat{\Omega})} \right) \\
 &= -2 (\ln l(\hat{\omega}) - \ln l(\hat{\Omega})) \quad (2.19)
 \end{aligned}$$

Statistik uji  $G^2$  mengikuti distribusi *Chi-squared* dengan derajat bebas  $p$ . Kriteria penolakan yaitu tolak  $H_0$  jika  $G^2 > \chi_{\alpha, p}^2$ . yang menunjukkan minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon.

Uji parameter secara parsial dilakukan untuk mengetahui signifikansi setiap parameter terhadap variabel respon. Pengujian ini dilakukan dengan hipotesis sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, \text{ dengan } j = 1, 2, \dots, p$$

Statistik uji parameter secara parsial adalah sebagai berikut.

$$W^2 = \left[ \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right]^2 \quad (2.20)$$

dengan  $SE(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)}$ . Daerah penolakan yaitu tolak  $H_0$  apabila  $W^2 > \chi_{\alpha, 1}^2$ . Selain itu, dapat juga menggunakan *p-value*,

yaitu tolak  $H_0$  jika  $p\text{-value} < \alpha$ . Keputusan tolak  $H_0$  menunjukkan bahwa variabel ke- $j$  berpengaruh signifikan di dalam model dengan taraf signifikansi sebesar  $\alpha$  (Hosmer & Lemeshow, 2000).

### 2.2.3 Uji Kesesuaian Model

Pengujian kesesuaian model bertujuan untuk menilai kelayakan model regresi logistik yang dihasilkan yaitu tidak terdapat perbedaan antara hasil pengamatan dan hasil prediksi model regresi logistik. Hal ini biasa disebut sebagai *goodnes of fit*. Pengujian ini dilakukan dengan hipotesis sebagai berikut.

$H_0$  : Model sesuai

$H_1$  : Model tidak sesuai

Persamaan (2.21) berikut merupakan statistik uji dalam melakukan uji kesesuaian model.

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.21)$$

Keterangan:

$O_k$  :  $\sum_{j=1}^{C_k} y_j$ : Jumlah nilai variabel respon pada grup ke- $k$

$\bar{\pi}_k$  :  $\sum_{j=1}^{C_k} \frac{m_j \hat{\pi}_j}{n'_k}$ : Rata-rata taksiran probabilitas

$g$  : Jumlah grup (kombinasi kategori dalam model serentak)

$n'_k$  : Banyak observasi pada grup ke- $k$

$m_j$  : Banyaknya observasi yang memiliki nilai  $\bar{\pi}_j$

$C_k$  : Kategori variabel respon

Daerah penolakan yaitu tolak  $H_0$  jika jika  $\hat{C} > \chi^2_{(\alpha, p-2)}$  yang menunjukkan bahwa model tidak sesuai (Hosmer & Lemeshow, 2000).

### 2.2.4 Interpretasi Model

Hubungan antara variabel respon dan prediktor dalam model regresi logistik dapat diketahui melalui nilai *odds ratio*. Berikut *odds ratio* untuk variabel prediktor yang berskala kategori.

$$\psi_k(a, b) = \frac{Pr(Y = k|x = a) / Pr(Y = 0|x = a)}{Pr(Y = k|x = b) / Pr(Y = 0|x = b)} = e^{\beta_j} \quad (2.22)$$

dengan  $a, b$  adalah kategori dari variabel prediktor tersebut. Nilai  $\psi_k = 1$  menunjukkan kedua variabel tidak memiliki hubungan, sedangkan jika  $0 < \psi_k < 1$  menunjukkan variabel prediktor dan variabel respon memiliki hubungan yang negatif untuk setiap perubahan nilai pada variabel prediktor. Nilai  $\psi_k > 1$  mengartikan bahwa variabel prediktor dan variabel respon memiliki hubungan positif untuk setiap perubahan nilai pada variabel prediktor (Hosmer & Lemeshow, 2000).

### 2.2.5 Akaike's Information Criterion

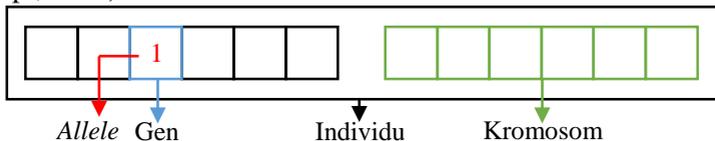
Pemilihan model terbaik diperoleh menggunakan nilai *Akaike's Information Criterion* (AIC) terkecil.

$$AIC = -2 \ln(L(\hat{\beta})) + 2k \quad (2.23)$$

Dengan  $L(\hat{\beta})$  adalah nilai maksimum fungsi *likelihood* dan  $k$  adalah jumlah parameter.

## 2.3 Genetic Algorithm

*Genetic algorithm* bekerja dengan memanfaatkan prinsip genetik dan seleksi alam. Algoritma yang dikembangkan oleh John Holland tahun 1975 dan dipopulerkan oleh David Goldberg tahun 1989 ini terinspirasi oleh konsep evolusi. Algoritma ini berfungsi untuk mencari variabel yang dapat digunakan dalam menentukan jenis kelas dengan meminimalisir variabel yang dianggap tidak informatif sehingga diperoleh variabel utama dalam melakukan deskripsi suatu kelas. Salah satu kebaikan *genetic algorithm* yaitu mampu mengatasi permasalahan kompleks dan paralel (Haupt & Haupt, 2004).



**Gambar 2.1** Representasi Gen, Kromosom, dan Individu

Berikut adalah struktur umum dalam *genetic algorithms*.

1. *Genotype* (gen) adalah sebuah nilai yang menyatakan satuan dasar yang membentuk arti tertentu dalam satu kesatuan gen yang disebut kromosom.

2. *Allele* adalah nilai dari gen. Tipe data *allele* dapat berupa biner, *floating point*, atau *integer* tergantung representasi genetika yang digunakan.
  3. Kromosom adalah gabungan gen yang membentuk nilai tertentu.
  4. Individu menyatakan suatu keadaan yang menunjukkan salah satu kandidat solusi dari permasalahan.
  5. Populasi adalah sekumpulan individu yang akan diproses bersama dalam satu siklus evolusi.
- Komponen penting dalam *genetic algorithm* adalah sebagai berikut.

1. Representasi dan inisialisasi  
*Genetic algorithm* bekerja pada sejumlah solusi dari suatu populasi. Pada awalnya gen-gen yang diinisialisasi sebagai estimasi nilai mengandung informasi dalam bentuk kode tertentu. Satu gen mewakili satu variabel. Pada kasus seleksi variabel, jumlah gen pada setiap kromosom yaitu sebanyak jumlah variabel prediktor pada penelitian. *Allele* mempunyai nilai yang bersifat biner yaitu 0 dan 1, dengan variabel yang tidak terpilih bernilai 0 dan variabel yang terpilih bernilai 1.
2. Fungsi *fitness*  
 Fungsi *fitness* digunakan untuk mengukur tingkat kebaikan dari solusi yang didapatkan. Fungsi *fitness* yang digunakan adalah nilai AUC (Gayou, et al., 2008).
3. *Elitism*  
*Elitism* merupakan proses pengopian individu agar individu dengan *fitness* tertinggi tidak hilang selama proses evolusi. Suatu individu yang memiliki nilai *fitness* tertinggi tidak akan selalu terpilih karena proses genetika dilakukan secara random sehingga untuk menjaga individu tersebut perlu dibuat satu atau beberapa salinan.
4. *Parent Selection*  
*Selection* bertujuan untuk memberikan kesempatan reproduksi bagi anggota populasi yang memiliki *fitness* tinggi. Individu dengan *fitness* lebih tinggi akan memiliki

kesempatan lebih besar untuk terpilih dalam proses reproduksi. Pada seleksi *roulette wheel*, tiap kromosom menempati potongan lingkaran pada roda *roulette* secara proporsional sesuai dengan nilai *fitness* masing-masing sehingga semua kromosom memiliki kesempatan untuk dipilih. Sebuah kromosom akan terpilih, jika bilangan random yang dibangkitkan berada di dalam interval kumulatifnya (Gayou, et al., 2008). Berikut adalah tahapan seleksi *roulette wheel* (Jadaan, Rajamani, & Rao, 2005).

- a. Mengitung nilai *fitness* tiap kromosom  $v_i$ . Berikut ilustrasi pada perhitungan nilai *fitness*.

Populasi	$x_1$	$x_2$	...	$x_9$	Nilai <i>Fitness</i>
1	1	0	...	1	AUC <sub>1</sub>
2	1	1	...	0	AUC <sub>2</sub>
⋮	⋮	⋮	⋮	⋮	
$n$	0	0	...	1	AUC <sub><math>n</math></sub>

- b. Menghitung total nilai *fitness* dalam populasi,

$$F = \sum_{h=1}^{n_{pop}} f(h) \quad (2.24)$$

- c. Menghitung proporsi masing-masing kromosom,

$$P_h = \frac{f(h)}{F}, h = 1, 2, \dots, n_{pop} \quad (2.25)$$

- d. Menghitung nilai kumulatif proporsi untuk masing-masing kromosom,

$$S_h = \sum_{q=1}^h P_q, h = 1, 2, \dots, n_{pop} \quad (2.26)$$

- e. Membangkitkan sebuah bilangan random ( $r$ ) berdistribusi Uniform antara 0 sampai 1.

- f. Jika  $r \leq S_h$ , maka dipilih kromosom  $v_h$ . Jika tidak, cek apabila  $r \leq S_{h+1}$ , maka dipilih kromosom  $v_{h+1}$ .

- g. Ulangi tahapan e hingga terpilih jumlah kromosom sebanyak  $n$ .

Keterangan:

$F$  : Total nilai *fitness* semua kromosom dalam populasi

$N$  : Jumlah kromosom dalam satu populasi

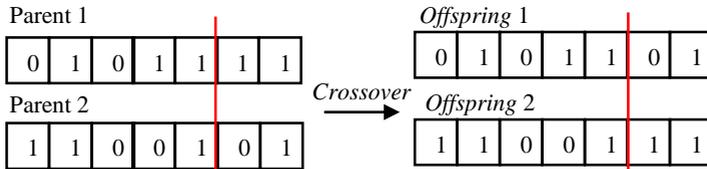
$f(h)$  : Nilai *fitness* kromosom ke- $h$

$P_h$  : Nilai proporsi *fitness* kromosom ke- $h$

$S_q$  : Nilai *fitness* kumulatif kromosom  $q$  ke- $h$

5. *Crossover*

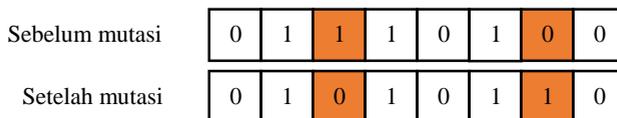
*Crossover* adalah proses penyilangan dua kromosom untuk membentuk keturunan kromosom baru. Proses *crossover* dapat dilakukan jika suatu bilangan random berdistribusi Uniform antara 0 sampai 1 kurang dari probabilitas *crossover* ( $P_c$ ), umumnya mendekati 1, misalnya 0,8 dan 0,9. Berikut ilustrasi *single point crossover*.



**Gambar 2.2** Ilustrasi *Single Point Crossover*

6. Mutasi (*Mutation*)

Mutasi dilakukan agar algoritma tidak terjebak pada solusi lokal dan memperluas kemampuan untuk mencari solusi optimal. Mutasi dilakukan dengan membalik arah dari setiap gen, yaitu menukar 0 menjadi 1, atau sebaliknya. Namun proses mutasi dibatasi dengan ketentuan yaitu, jika bilangan random berdistribusi Uniform antara 0 sampai 1 yang dibangkitkan kurang dari probabilitas mutasi ( $P_m$ ) maka terjadi mutasi, umumnya mendekati 0, misalnya 0,1;0,01; dan 0,05.



**Gambar 2.3** Ilustrasi Mutasi

7. Penggantian Populasi (*Replacement*)

Penggantian populasi digunakan untuk menentukan individu yang bertahan pada populasi dan individu yang harus diganti

karena mempengaruhi konvergensi. Beberapa teknik penggantian populasi, diantaranya yaitu melakukan penggantian individu yang paling tua, mengganti individu dengan nilai *fitness* terkecil, dan membandingkan anak dengan kedua induk, yaitu anak menggantikan induk yang memiliki *fitness* terendah. Skema penggantian populasi berarti mengganti sejumlah  $N$  individu dalam satu populasi dari suatu generasi dengan  $N$  individu baru (Suyanto, 2005).

## 2.4 Evaluasi Metode

Evaluasi prosedur klasifikasi bertujuan untuk mengetahui persentase kesalahan hasil klasifikasi atau persentase ketepatan hasil klasifikasi yang dilakukan oleh metode klasifikasi tertentu. Ketepatan hasil klasifikasi diperoleh melalui nilai akurasi yang dihasilkan oleh *confusion matrix*. Variabel respon yang memiliki dua kelas memiliki empat kemungkinan hasil prediksi klasifikasi yang berbeda yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP) dan *false negatif* (FN).

**Tabel 2.3** *Confusion Matrix*

		<i>Predicted Membership</i>	
		<i>Positive class</i>	<i>Negative class</i>
<i>Actual Membership</i>	<i>Positive class</i>	TP	FN
	<i>Negative class</i>	FP	TN

Tabel 2.3 merupakan *confusion matrix* dengan dua kelas pada variabel respon (Johnson & Winchern, 2007). Dapat dihitung ketepatan hasil prediksi menggunakan *Area Under ROC (Receive Operating Characteristic) Curve (AUC)*. *Area Under ROC (Receive Operating Characteristic) Curve (AUC)* bertujuan untuk menggambarkan probabilitas antara *sensitivity* dan *specificity* dengan batas 0 sampai 1. Apabila nilai AUC mendekati 1, maka model klasifikasi yang terbentuk semakin akurat. Nilai AUC dapat dihitung dengan persamaan berikut (Bekkar, Djemaa, & Alitouche, 2013).

$$AUC = \frac{1}{2}(\text{sensitivity} + \text{specificity}) \quad (2.27)$$

dengan.

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (2.28)$$

$$\text{specificity} = \frac{TN}{FP + TN} \quad (2.29)$$

Keterangan:

- TP : Jumlah prediksi positif yang tepat diklasifikasikan ke kelas positif
- TN : Jumlah prediksi negatif yang tepat diklasifikasikan ke kelas negatif
- FP : Jumlah prediksi negatif yang salah diklasifikasikan ke kelas positif
- FN : Jumlah prediksi positif yang salah diklasifikasikan ke kelas negatif

## 2.5 Gejala Kanker Payudara

Kanker payudara merupakan penyakit neoplasma ganas yang berasal dari parenkim. Penyakit ini muncul sebagai akibat sel-sel abnormal yang terbentuk dan tumbuh pada jaringan payudara secara tidak terkendali dan beraturan. Sel-sel tersebut adalah hasil mutasi gen dengan perubahan-perubahan bentuk, ukuran, maupun fungsi. Penyakit ini dapat menyebar ke organ tubuh lain seperti paru-paru, hati, dan otak melalui pembuluh darah. Kelenjar getah bening aksila atau supraklavikula membesar akibat dari penyebaran kanker payudara melalui pembuluh getah bening dan tumbuh di kelenjar getah bening. Gejala-gejala umum kanker payudara adalah sebagai berikut (Komite Penanggulangan Kanker Nasional, 2017).

1. Benjolan di payudara atau ketiak dengan/tanpa rasa sakit
2. Ukuran payudara yang tidak sama
3. Retraksi atau penarikan puting susu
4. Puting susu mengeluarkan darah
5. Perubahan kulit payudara seperti edema, kerutan, kulit menyerupai permukaan jeruk
6. Penebalan kulit payudara
7. Nyeri tulang, sesak, dan lain sebagainya

Kualitas tidur merupakan ukuran seseorang dapat dengan mudah dalam memulai tidur dan mempertahankan tidur. Kualitas tidur itu digambarkan dengan lama waktu tidur dan keluhan yang dirasakan saat tidur ataupun setelah bangun tidur. Peningkatan risiko gangguan tidur mungkin berhubungan dengan immunosupresi. Hubungan antara gangguan tidur dan penekanan sistem kekebalan tubuh dapat menyebabkan keganasan terbentuk dan meningkat yang menunjukkan bahwa pasien gangguan tidur memiliki peningkatan risiko kanker payudara (Besedovsky, Lange, & Born, 2012). Gangguan tidur seperti sulit tidur atau melakukan kegiatan hingga larut malam memungkinkan hormon melatonin menurun karena kekebalan tubuh yang menurun sehingga berpengaruh terhadap kejadian kanker payudara. Disamping itu, penurunan produksi hormon melatonin akan meningkatkan produksi hormon estrogen yang dapat memicu risiko kanker (Dumitrescu & Cotarla, 2005). Saat kekebalan tubuh menurun, sel-sel normal berubah menjadi sel-sel kanker. Gangguan fungsi kekebalan tubuh, seperti peningkatan kadar kortisol bersamaan dengan aktivitas sel pembunuh alami, limfosit pengintai tumor yang dominan, dapat memengaruhi risiko kanker payudara (McElroy, et al., 2006).

*(Halaman ini sengaja dikosongkan)*

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh dari Rumah Terapi Tauhid RS Syariah Nur Hidayah Bantul, Yogyakarta. Pengambilan data dilakukan pada bulan April 2019 dengan jumlah sampel yang terkumpul sebanyak 49 responden.

#### **3.2 Desain Penelitian**

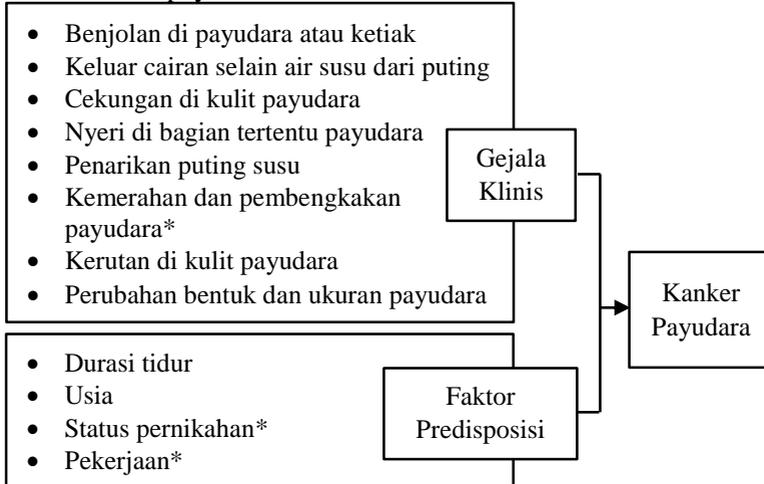
Responden terbagi ke dalam dua kelompok yaitu penderita dan bukan penderita kanker payudara. Kelompok penderita kanker payudara yaitu responden yang telah melakukan tes pemeriksaan kanker payudara seperti USG Mammae dan patologi anatomi dengan hasil yang positif. Kelompok bukan penderita kanker payudara yaitu responden yang belum pernah melakukan pemeriksaan kanker payudara. Pada penelitian, kelompok bukan penderita kanker payudara disebut dengan responden negatif. Kriteria yang dibutuhkan pada subyek penelitian ini adalah sebagai berikut.

- a. Kriteria inklusi:
  1. Memiliki jenis kelamin perempuan.
  2. Responden dengan hasil tes pemeriksaan kanker payudara positif atau responden yang belum melakukan pemeriksaan kanker payudara.
  3. Responden yang melakukan terapi/pengobatan selain medis di Rumah Terapi Tauhid RS Nur Hidayah.
- b. Kriteria eksklusi yaitu responden yang sebelumnya sudah pernah melakukan terapi/pengobatan selain medis di Rumah Terapi Tauhid RS Syariah Nur Hidayah Bantul pada jangka waktu pengambilan data.

Pengambilan dan pemilihan sampel dilakukan dengan mempertimbangkan kesesuaian responden dengan kriteria inklusi dan eksklusi yang telah ditentukan.

### 3.3 Kerangka Konsep

Gejala klinis pada kanker payudara dipertimbangkan bersama dengan faktor predisposisi. Pada penelitian ini, dilakukan peninjauan mengenai gejala klinis dan faktor predisposisi pada status kanker payudara.



\* variabel yang tidak digunakan dalam penelitian

**Gambar 3.1** Kerangka Konsep Risiko Prevalensi Kanker Payudara  
Gejala-gejala klinis pada penelitian ini dimasukkan dengan mempertimbangkan gejala umum yang dijelaskan oleh Komite Penanggulangan Kanker Nasional dan beberapa penelitian terdahulu. Durasi tidur seseorang diduga memiliki hubungan dengan risiko kejadian kanker payudara. Sebuah penelitian menunjukkan bahwa efisiensi tidur yang rendah dan banyak gangguan tidur secara signifikan menjadi faktor risiko independen pada wanita dengan kanker payudara stadium lanjut (Palesh, et al., 2014). Penelitian serupa juga dilakukan dengan hasil bahwa risiko kanker payudara menurun dengan meningkatnya durasi tidur (Wu, et al., 2008).

### 3.4 Variabel Penelitian

Variabel yang digunakan pada penelitian ini terdiri dari variabel respon ( $Y$ ) dan prediktor ( $X$ ). Variabel respon dibagi ke

dalam dua kategori yaitu status pemeriksaan kanker payudara positif dengan hasil tes kanker payudara positif dan status kanker payudara negatif yaitu responden yang belum melakukan pemeriksaan kanker payudara. Tabel 3.1 berikut menampilkan variabel yang digunakan pada penelitian.

**Tabel 3.1** Variabel Penelitian

<b>Simbol</b>	<b>Variabel</b>	<b>Skala Data</b>	<b>Kategori</b>
Y	Status pemeriksaan kanker payudara	Nominal	1 = Positif 0 = Negatif
X <sub>1</sub>	Keluar cairan dari puting susu	Nominal	1 = Ya; 0 = Tidak
X <sub>2</sub>	Cekungan di kulit payudara	Nominal	1 = Ya; 0 = Tidak
X <sub>3</sub>	Nyeri di bagian tertentu payudara	Nominal	1 = Ya; 0 = Tidak
X <sub>4</sub>	Penarikan puting susu ke dalam	Nominal	1 = Ya; 0 = Tidak
X <sub>5</sub>	Kerutan di kulit payudara	Nominal	1 = Ya; 0 = Tidak
X <sub>6</sub>	Benjolan di bawah ketiak	Nominal	1 = Ya; 0 = Tidak
X <sub>7</sub>	Perubahan bentuk dan ukuran payudara	Nominal	1 = Ya; 0 = Tidak
X <sub>8</sub>	Durasi tidur	Ordinal	3 = < 5 Jam 2 = ≥ 5-6 Jam 1 = ≥ 6-7 Jam 0 = ≥ 7 Jam
X <sub>9</sub>	Usia	Ordinal	1 = ≥ 45 tahun 0 = < 45 tahun

Berikut ini adalah definisi operasional variabel penelitian.

1. Keluar cairan selain air susu dari puting susu. Cairan yang keluar dari puting susu selain air susu dapat berupa cairan berwarna kuning, hijau, hingga merah. Wanita menyusui yang mengeluarkan air susu adalah kondisi normal.
2. Cekungan pada kulit payudara (lesung pipi). Sel-sel kanker dapat menyebabkan penumpukan cairan getah bening di payudara yang menyebabkan kulit membengkak dan

berbentuk seperti lesung pipi. Perubahan penampilan kulit payudara ini menyerupai permukaan jeruk.

3. Nyeri di bagian tertentu payudara. Kanker payudara dapat menyebabkan perubahan pada sel-sel kulit yang menyebabkan rasa sakit, nyeri saat ditekan, dan ketidaknyamanan pada payudara.
4. Penarikan puting susu ke dalam. Kanker payudara dapat menyebabkan perubahan sel di balik puting sehingga mengakibatkan kondisi puting susu tertarik ke dalam atau biasa disebut retraksi. Pada umumnya, wanita dengan retraksi puting susu adalah keadaan normal, namun jika hanya terjadi pada satu sisi maka perlu dilakukan pemeriksaan.
5. Kerutan di kulit payudara. Kanker payudara dapat menyebabkan perubahan dan peradangan pada sel-sel kulit yang dapat menyebabkan perubahan tekstur. Contoh dari perubahan tekstur ini termasuk, kulit bersisik di sekitar puting dan areola, seolah kulitnya terbakar sinar matahari atau sangat kering serta penebalan kulit di payudara.
6. Benjolan di bawah ketiak. Kelenjar getah bening adalah kumpulan kecil dari jaringan sistem kekebalan yang menyaring cairan dan menangkap sel-sel yang berpotensi bahaya seperti bakteri, virus, dan sel kanker. Jika sel kanker meninggalkan payudara, sel tersebut akan menyebar ke daerah kelenjar getah bening ketiak di sisi yang sama dengan payudara yang terkena sehingga dapat menyebabkan pembengkakan di area ini.
7. Perubahan bentuk dan ukuran payudara. Kanker payudara dapat menyebabkan seluruh payudara atau area payudara membesar. Pembengkakan payudara ini juga menyebabkan ukuran payudara satu berbeda dari payudara lainnya.
8. Durasi tidur. Durasi tidur adalah waktu yang dibutuhkan seseorang ketika mulai tertidur saat malam hari hingga

bangun di pagi hari dalam satuan jam (Lumbantobing, 2008).

9. Usia. Usia yang digunakan yaitu usia saat pengambilan data.

### 3.5 Struktur Data

Struktur data dari variabel-variabel yang digunakan dalam penelitian dijelaskan dalam Tabel 3.2 berikut.

Tabel 3.2 Struktur Data Penelitian

Responden ke-	Y	X <sub>1</sub>	X <sub>2</sub>	...	X <sub>9</sub>
1	Y <sub>1</sub>	X <sub>1,1</sub>	X <sub>2,1</sub>	...	X <sub>9,1</sub>
2	Y <sub>2</sub>	X <sub>1,2</sub>	X <sub>2,2</sub>	...	X <sub>9,2</sub>
⋮	⋮	⋮	⋮	⋮	⋮
i	Y <sub>i</sub>	X <sub>1,i</sub>	X <sub>2,i</sub>	...	X <sub>9,i</sub>
⋮	⋮	⋮	⋮	⋮	⋮
49	Y <sub>49</sub>	X <sub>1,49</sub>	X <sub>2,49</sub>	...	X <sub>9,49</sub>

### 3.6 Langkah Analisis

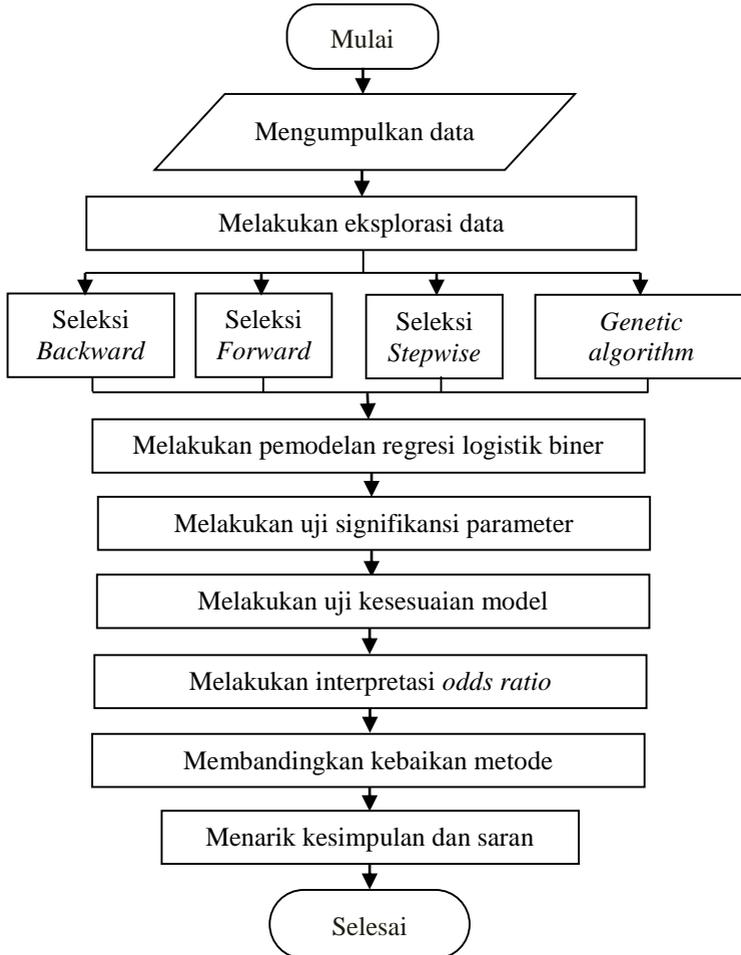
Langkah-langkah analisis dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan data.
2. Melakukan eksplorasi data.
3. Melakukan seleksi variabel menggunakan metode *backward*, *forward*, *stepwise*, dan *genetic algorithm*. Berikut adalah langkah-langkah dalam *genetic algorithm*.
  - a. Membentuk populasi awal yang terdiri dari beberapa kromosom yang memuat gen. Pada penelitian ini gen berisi variabel prediktor dengan *binary encoding* yaitu 0 untuk variabel prediktor yang tidak masuk ke dalam model dan 1 untuk variabel prediktor yang masuk ke dalam model. Inisialisasi dilakukan dengan membangkitkan 80 individu secara acak.
  - b. Melakukan evaluasi kemampuan masing-masing kromosom dengan nilai *fitness* menggunakan AUC
  - c. Melakukan proses *selection* sebanyak 80 kromosom. Apabila sebuah kromosom memiliki nilai *fitness* lebih optimum dari nilai inisialnya, maka kromosom dihentikan namun apabila tidak dilanjutkan ke tahap

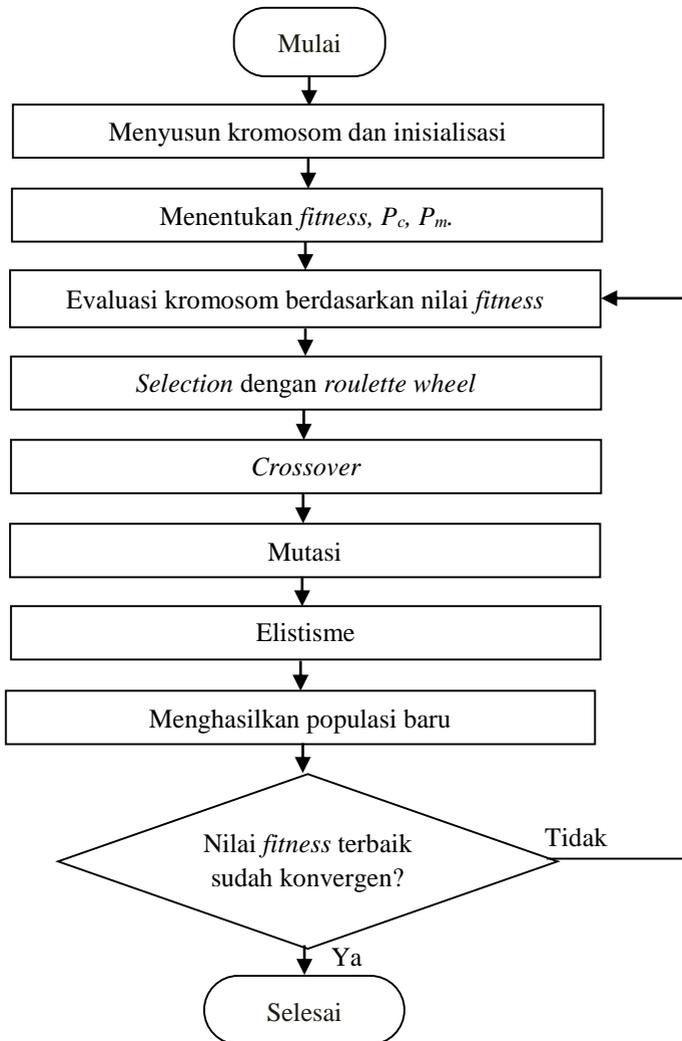
- berikutnya. Nilai *fitness* tertinggi dipilih sebagai solusi permasalahan pada penelitian ini.
- d. Melakukan proses *crossover*. *Crossover* terjadi apabila nilai bilangan *random*  $r$  (0,1) yang dibangkitkan kurang dari probabilitas *crossover* yaitu 0,8. Dua induk secara acak dipilih dan digunakan untuk membentuk dua kromosom baru.
  - e. Melakukan mutasi untuk memperkenalkan unsur gen baru pada kromosom secara acak. Mutasi terjadi jika nilai bilangan *random*  $r$  (0,1) yang dibangkitkan kurang dari probabilitas mutasi yaitu 0,1.
  - f. Melakukan proses elitisme, lima kromosom dengan nilai *fitness* terbaik akan bertahan ke generasi selanjutnya.
  - g. Melakukan pergantian populasi lama dengan generasi baru melalui pemilihan kromosom terbaik dari orang tua dan anak baru yang memiliki nilai *fitness* tertinggi setelah proses *selection*, *crossover*, mutasi, dan elitisme.
  - h. Melakukan pemeriksaan solusi yang didapatkan. Solusi yang diperoleh dikatakan sudah memenuhi kriteria apabila nilai *fitness* terbaik sudah konvergen. Apabila solusi yang didapatkan belum mencapai kriteria maka kembali ke-b.
4. Melakukan pemodelan kanker payudara menggunakan regresi logistik biner.
  5. Melakukan uji parameter secara serentak dan parsial.
  6. Melakukan uji kesesuaian model.
  7. Melakukan interpretasi *odds ratio*.
  8. Membandingkan model terbaik antar metode seleksi variabel.
  9. Menarik kesimpulan.

### 3.7 Diagram Alir

Berdasarkan langkah analisis sebelumnya, berikut merupakan diagram alir dari penelitian ini.



Gambar 3.2 Diagram Alir

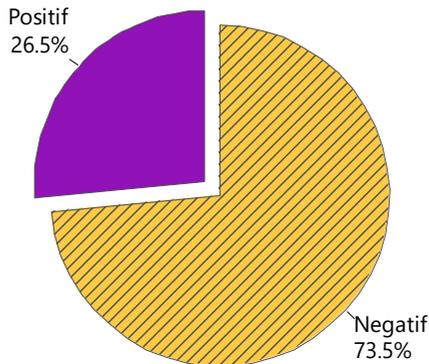


**Gambar 3.2** Diagram Alir *Genetic Algorithm*

## BAB IV ANALISIS DAN PEMBAHASAN

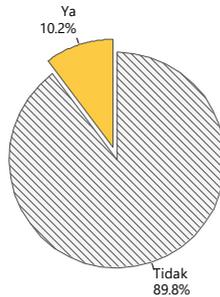
### 4.1 Karakteristik Gejala Berdasarkan Status Pemeriksaan Kanker Payudara

Nyeri di bagian tertentu payudara, keluar cairan selain air susu dari puting susu, terdapat cekungan seperti lesung pipi pada kulit payudara, penarikan puting susu, kerutan di kulit payudara, benjolan di bawah ketiak serta perubahan bentuk dan ukuran pada salah satu payudara dapat menjadi tanda-tanda dan gejala-gejala bagi perempuan agar sadar akan penyakit kanker payudara sedini mungkin. Berikut analisis statistika deskriptif dan visualisasi karakteristik responden.



**Gambar 4.1** Persentase Status Pemeriksaan Kanker Payudara

Gambar 4.1 menunjukkan persentase status pemeriksaan kanker payudara positif dan negatif di RS Nur Hidayah Bantul. Diperoleh hampir sebanyak 75 persen atau 13 dari 49 reponden berstatus pemeriksaan kanker payudara positif sedangkan sekitar 26,5 persen responden memiliki status negatif. Adanya cekungan pada kulit payudara dapat dikaitkan dengan kejadian kanker payudara. Sel-sel kanker yang berkembang di dalam tubuh menyebabkan adanya penumpukan cairan getah bening sehingga terjadi pembengkakan kulit payudara dan memiliki tampilan menyerupai permukaan jeruk.



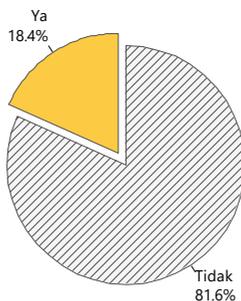
**Gambar 4.2** Persentase Responden dengan Gejala Cekungan Pada Kulit Payudara

Terlihat pada Gambar 4.2, hampir 90 persen atau sebanyak 44 responden tidak memiliki riwayat gejala cekungan pada payudara. Sebanyak 44 responden tersebut, sejumlah 35 diantaranya merupakan penderita dengan status pemeriksaan kanker payudara negatif. Lebih jelas lagi, dapat dilihat pada tabulasi silang Tabel 4.1 berikut.

**Tabel 4.1** Tabulasi Silang Status Kanker Payudara dan Cekungan Pada Kulit Payudara

		Status Kanker Payudara		Total	
		Negatif	Positif		
<b>Cekungan pada kulit payudara</b>	<b>Tidak</b>	<i>Count</i>	35	9	44
		<i>% Total</i>	71,4	18,4	89,8
	<b>Ya</b>	<i>Count</i>	1	4	5
		<i>% Total</i>	2	8,2	10,2
<b>Total</b>		<i>Count</i>	36	13	49
		<i>% Total</i>	73,5	26,5	100

Berdasarkan Tabel 4.1, diketahui bahwa sebanyak 9 responden atau 69,2 persen dari total responden dengan status pemeriksaan kanker payudara positif tidak mengalami gejala cekungan pada kulit payudara. Pada responden dengan status negatif, ditemui hanya 1 dari 36 responden yang mengalami gejala cekungan pada kulit payudara sedangkan pada status kanker payudara positif terdapat 4 responden yang memiliki riwayat gejala tersebut. Dalam kata lain, responden yang memiliki gejala cekungan pada kulit payudara banyak yang terdiagnosis status pemeriksaan kanker payudara positif.



**Gambar 4.3** Persentase Responden dengan Gejala Nyeri di Bagian Tertentu Payudara

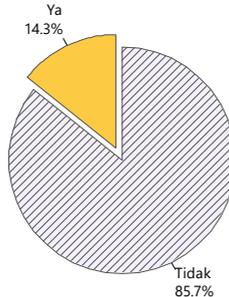
Gejala lain seperti rasa nyeri pada payudara dapat disebabkan oleh aktivitas sel kanker yang menimbulkan perubahan pada sel-sel kulit sehingga muncul rasa sakit dan nyeri saat ditekan. Gejala nyeri di bagian tertentu payudara menunjukkan persentase responden yang mengalami nyeri lebih banyak dibandingkan cekungan pada kulit payudara. Sebanyak 18,4 persen dari 49 responden menjawab mengalami nyeri di bagian tertentu payudara.

**Tabel 4.2** Tabulasi Silang Status Kanker Payudara dan Nyeri di Bagian Tertentu Payudara

		Status Kanker Payudara		Total	
		Negatif	Positif		
<b>Nyeri di bagian tertentu payudara</b>	<b>Tidak</b>	<i>Count</i>	33	7	40
		<i>% Total</i>	67,3	14,3	81,6
	<b>Ya</b>	<i>Count</i>	3	6	9
		<i>% Total</i>	6,1	12,2	18,4
<b>Total</b>		<i>Count</i>	36	13	49
		<i>% Total</i>	73,5	26,5	100

Terlihat pada Tabel 4.2, responden dengan status pemeriksaan kanker payudara positif dengan riwayat gejala nyeri di bagian tertentu payudara memiliki jumlah dan persentase yang hampir sama dengan responden status negatif. Berbeda pada status kanker payudara negatif, diketahui hanya 3 responden dengan status kanker payudara negatif yang mengalami nyeri di bagian tertentu payudara sedangkan sebanyak hampir 70 persen dari total responden status negatif tidak memiliki gejala nyeri di bagian tertentu payudara.

Peradangan pada sel-sel kulit payudara yang menyebabkan perubahan tekstur kulit seperti timbul kerutan, kulit bersisik, atau kulit kering dapat menjadi salah satu indikasi penyakit kanker payudara. Berikut adalah persentase responden yang mengalami gejala kerutan di sekitar payudara.



**Gambar 4.4** Persentase Responden dengan Gejala Kerutan di Sekitar Payudara Diagram lingkaran pada Gambar 4.4 menunjukkan bahwa hanya sekitar 15 persen responden yang menjawab pernah mengalami gejala kerutan di sekitar payudara. Lebih dari 70 persen di antaranya merupakan responden dengan status pemeriksaan kanker payudara positif. Sementara itu, 81 persen dari 42 responden yang tidak mengalami gejala kerutan di sekitar payudara merupakan responden dengan status negatif.

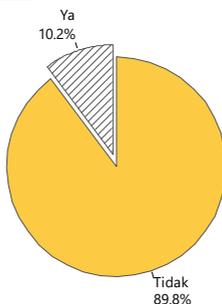
**Tabel 4.3** Tabulasi Silang Status Kanker Payudara dan Kerutan di Sekitar Payudara

		Status Kanker Payudara		Total
		Negatif	Positif	
<b>Kerutan di Tidak sekitar payudara</b>	<i>Count</i>	34	8	42
	<i>% Total</i>	69,4	16,3	85,7
<b>Ya</b>	<i>Count</i>	2	5	7
	<i>% Total</i>	4,1	10,2	14,3
<b>Total</b>	<i>Count</i>	36	13	49
	<i>% Total</i>	73,5	26,5	100

Tabel 4.3 menampilkan secara jelas jumlah responden pada setiap status pemeriksaan kanker payudara yang memiliki gejala kerutan di sekitar payudara. Diketahui hanya 2 responden atau sebesar lima persen responden dengan status pemeriksaan kanker payudara

negatif yang menjawab mengalami gejala kerutan di sekitar payudara. Sebanyak hampir 70 persen responden status pemeriksaan negatif tidak mengalami kerutan di sekitar payudara sedangkan responden status pemeriksaan kanker payudara positif yang memiliki gejala tersebut hanya 10 persen dari total responden.

Benjolan di bawah ketiak juga memiliki potensi gejala kanker payudara. Sel-sel kanker pada payudara yang telah berkembang dapat menyebar ke daerah sekitarnya seperti daerah kelenjar getah bening pada ketiak di sisi yang sama dengan payudara yang terkena sehingga dapat terjadi menyebabkan pembengkakan di area ini.



**Gambar 4.5** Persentase Responden dengan Gejala Benjolan di Bawah Ketiak Diperoleh hanya 10,2 persen responden yang memiliki gejala benjolan di bawah ketiak sedangkan sebanyak 44 responden menjawab tidak mengalami gejala tersebut.

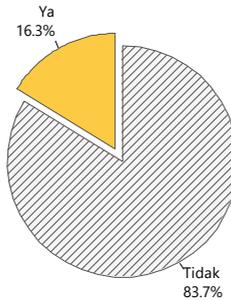
**Tabel 4.4** Tabulasi Silang Status Kanker Payudara dan Benjolan di Bawah Ketiak

		Status Kanker Payudara		Total
		Negatif	Positif	
<b>Benjolan di Tidak bawah ketiak</b>	<i>Count</i>	34	10	44
	<i>% Total</i>	69,4	20,4	89,8
<b>Ya</b>	<i>Count</i>	2	3	5
	<i>% Total</i>	4,1	6,1	10,2
<b>Total</b>	<i>Count</i>	36	13	49
	<i>% Total</i>	73,5	26,5	100

Terlihat pada Tabel 4.4, responden dengan gejala benjolan di bawah ketiak memiliki perbandingan jumlah yang hampir sama

antara status positif dan negatif. Sementara itu, responden yang tidak memiliki benjolan di bawah ketiak lebih banyak terdiagnosis status kanker payudara negatif dibandingkan negatif yaitu dengan persentase 69,4 persen atau sebanyak 34 responden.

Pembengkakan payudara akibat adanya aktivitas sel kanker dapat menyebabkan ukuran payudara satu berbeda dari payudara lainnya. Terlihat pada Gambar 4.6 persentase responden yang mengalami perubahan bentuk dan ukuran payudara sebesar 16,3 persen sedangkan responden yang tidak mengalami gejala ini sebanyak 41 responden.



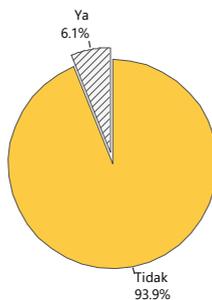
**Gambar 4.6** Persentase Responden dengan Gejala Perubahan Bentuk dan Ukuran Payudara

Tabel 4.5 berikut menunjukkan bahwa sebanyak 61,5 persen dari total responden dengan status pemeriksaan positif, tidak mengalami perubahan bentuk dan ukuran payudara. Sebanyak 33 dari 41 responden responden yang tidak mengalami perubahan bentuk dan ukuran payudara merupakan responden dengan status negatif.

**Tabel 4.5** Tabulasi Silang Status Kanker Payudara dan Perubahan Bentuk dan Ukuran Payudara

		<b>Status Kanker Payudara</b>		<b>Total</b>
		<b>Negatif</b>	<b>Positif</b>	
<b>Perubahan Tidak bentuk dan ukuran payudara</b>	<i>Count</i>	33	8	41
	<i>% Total</i>	67,3	16,3	89,8
<b>Ya</b>	<i>Count</i>	3	5	8
	<i>% Total</i>	6,1	10,2	16,3
<b>Total</b>	<i>Count</i>	36	13	49
	<i>% Total</i>	73,5	26,5	100

Wanita yang mengeluarkan air susu karena menyusui adalah kondisi normal, namun apabila ditemukan gejala keluar cairan berwarna kuning, hijau, hingga merah perlu dilakukan pemeriksaan.



**Gambar 4.7** Persentase Responden dengan Gejala Keluar Cairan dari Puting Susu

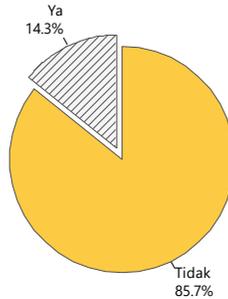
Terlihat perbedaan jumlah responden dengan gejala keluar cairan dari puting susu yang cukup jauh daripada gejala-gejala yang telah dijelaskan sebelumnya. Sebanyak 93,9 persen dari total responden yang tidak mengalami gejala keluar cairan dari puting susu dengan 11 diantaranya merupakan responden dengan status pemeriksaan kanker payudara positif. Sementara itu, diperoleh persentase hanya 6,1 persen responden yang mengalami gejala keluar cairan dari puting susu dengan 1 diantaranya adalah responden dengan status pemeriksaan kanker payudara negatif.

**Tabel 4.6** Tabulasi Silang Status Kanker Payudara dan Keluar Cairan dari Puting Susu

		<b>Status Kanker Payudara</b>		<b>Total</b>	
		<b>Negatif</b>	<b>Positif</b>		
<b>Keluar Cairan dari Puting Susu</b>	<b>Tidak</b>	<i>Count</i>	35	11	41
		<i>% Total</i>	71,43	22,45	93,88
	<b>Ya</b>	<i>Count</i>	1	2	3
		<i>% Total</i>	2,04	4,08	16,3
<b>Total</b>		<i>Count</i>	36	13	49
		<i>% Total</i>	73,5	26,5	100

Perubahan sel abnormal di balik puting dapat mengakibatkan kondisi puting susu tertarik ke dalam. Diperoleh sekitar 14 persen total responden atau sebanyak 7 wanita

mengalami gejala penarikan puting susu. Lima dari 7 responden tersebut merupakan responden dengan status pemeriksaan kanker payudara positif.



**Gambar 4.8** Persentase Responden dengan Gejala Penarikan Puting Susu

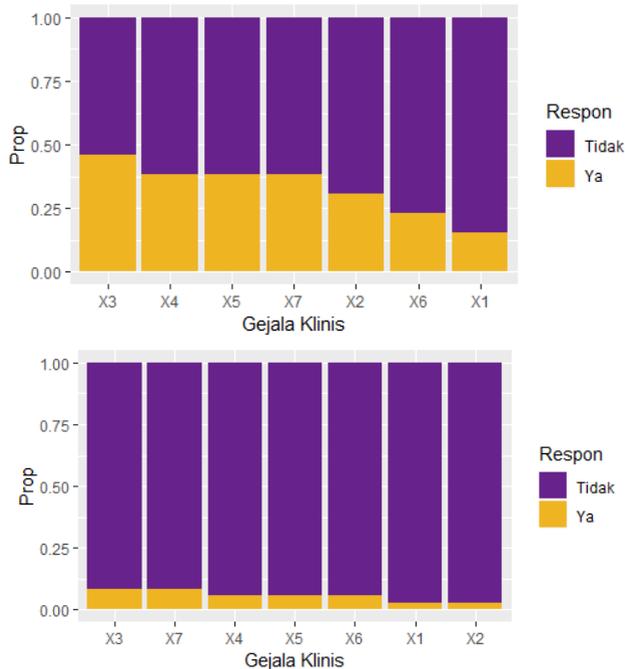
Gambar 4.8 juga menampilkan bahwa 86 persen responden tidak mengalami penarikan puting susu. Disamping itu, terlihat pada Tabel 4.7 sebanyak 62,5 persen total responden status pemeriksaan kanker payudara positif mengalami gejala penarikan puting susu.

**Tabel 4.7** Tabulasi Silang Status Kanker Payudara dan Penarikan Puting Susu

		<b>Status Kanker Payudara</b>		<b>Total</b>	
		<b>Negatif</b>	<b>Positif</b>		
<b>Penarikan Puting Susu</b>	<b>Tidak</b>	<i>Count</i>	34	8	42
		<i>% Total</i>	69,39	16,33	85,71
	<b>Ya</b>	<i>Count</i>	2	5	7
		<i>% Total</i>	4,08	10,2	14,29
<b>Total</b>		<i>Count</i>	36	13	49
		<i>% Total</i>	73,5	26,5	100

Gambar 4.9 berikut menampilkan perbandingan status pemeriksaan kanker payudara positif dan negatif yang mengalami gejala klinis dan tidak mengalami gejala klinis. Terlihat pada status pemeriksaan kanker payudara positif Gambar 4.9 (a), nyeri di bagian tertentu payudara ( $X_3$ ) menjadi gejala yang sering dikeluhkan oleh para responden berstatus pemeriksaan kanker payudara positif yaitu hampir setengah total penderita status positif menjawab mengalami gejala tersebut. Kemudian, sekitar 40 persen dari responden status pemeriksaan kanker payudara positif juga mengalami gejala penarikan puting susu ( $X_4$ ), kerutan pada kulit

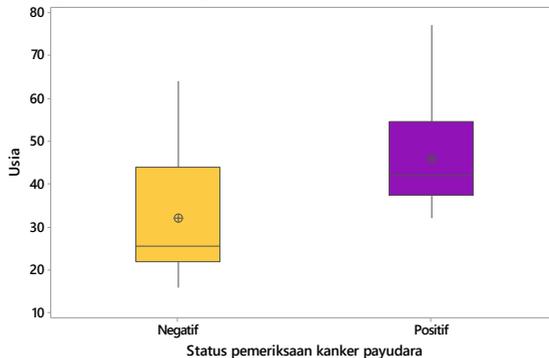
payudara ( $X_5$ ), dan perubahan bentuk dan ukuran payudara ( $X_7$ ). Sekitar 30 persen responden dengan status pemeriksaan kanker payudara positif mengalami gejala cekungan pada kulit payudara. Sementara itu, kurang dari 25 persen dari total responden dengan status pemeriksaan kanker payudara positif mengalami gejala keluar cairan dari puting susu ( $X_1$ ).



**Gambar 4.9** Perbandingan Gejala Klinis pada Status Pemeriksaan Kanker Payudara: (a) Positif; (b) Negatif

Tanda dan gejala klinis kanker payudara tersebut tidak menjadi ukuran pasti bagi seseorang untuk terkena kanker payudara. Terlihat diagram batang gejala klinis pada status pemeriksaan negatif Gambar 4.9 (b), responden dengan status negatif juga mengalami gejala klinis kanker payudara seperti yang dialami oleh responden dengan status pemeriksaan positif meskipun persentasenya sangat kecil. Sekitar 10 persen dari total

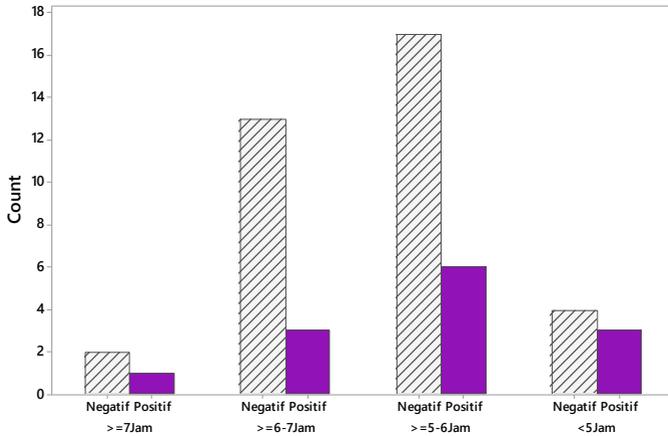
responden status kanker payudara negatif mengalami gejala nyeri di bagian tertentu payudara ( $X_3$ ) dan perubahan bentuk dan ukuran payudara ( $X_7$ ) sedangkan cekungan pada kulit payudara ( $X_2$ ) menjadi gejala yang paling jarang ditemui oleh responden dengan status kanker payudara negatif.



**Gambar 4.10** Perbandingan Usia Responden Berdasarkan Status Pemeriksaan Kanker Payudara

Perbedaan rata-rata usia antar responden dengan status pemeriksaan kanker payudara positif dan negatif dapat dilihat pada Gambar 4.10. Responden dengan status pemeriksaan positif memiliki rata-rata usia yang lebih tua yaitu 45 tahun dibandingkan responden dengan status kanker payudara negatif yaitu 32 tahun. Hal itu ditunjukkan oleh garis tengah pada boxplot yaitu garis median pada status positif berada pada posisi yang lebih tinggi daripada status negatif. Disamping itu, tidak ditemukan adanya outlier atau nilai ekstrem pada kedua kelas status pemeriksaan kanker payudara. Panjang boxplot menggambarkan ukuran penyebaran data, diperoleh responden dengan status kanker payudara negatif memiliki *varians* usia yang lebih beragam daripada status pemeriksaan kanker payudara positif. Usia responden dengan status pemeriksaan kanker payudara positif dan negatif masing-masing menunjukkan pola distribusi yang tidak simetris. Hal itu ditunjukkan oleh garis median yang tidak berada di tengah kotak dan garis vertikal bagian atas dan bawah kotak

yang memiliki panjang berbeda. Hal itu menunjukkan bahwa usia responden cenderung berkumpul di satu sisi yaitu responden dengan status pemeriksaan positif cenderung berkumpul di usia 30 hingga 50 tahun sedangkan responden dengan status negatif cenderung berusia 20 hingga 30 tahun.



**Gambar 4.11** Perbandingan Status Pemeriksaan Kanker Payudara-Durasi Tidur

Terlihat pola sama pada tiap kategori durasi tidur yaitu responden dengan status negatif memiliki frekuensi yang lebih tinggi daripada status positif. Mayoritas responden memiliki durasi tidur  $\geq 5-6$  jam baik responden dengan status positif maupun negatif. Responden jarang memiliki durasi tidur  $\geq 7$  jam. Sekitar 46.15 persen total responden status pemeriksaan positif memiliki durasi tidur  $\geq 5-6$  jam. Begitu juga, sebanyak 47.22 persen total responden status negatif memiliki durasi tidur  $\geq 5-6$  jam. Diperoleh sebanyak 14.3 persen responden memiliki durasi tidur  $< 5$  jam dengan 75 persen diantaranya merupakan responden dengan status positif.

#### 4.2 Regresi Logistik Biner-Seleksi *Backward*

Dalam mengetahui gejala klinis yang berpengaruh signifikan terhadap status pemeriksaan kanker payudara dilakukan analisis menggunakan regresi logistik biner. Kategori 1 menunjukkan status pemeriksaan kanker payudara positif

sedangkan kategori 0 menunjukkan status negatif. Pada pemodelan ini, kategori 0 digunakan sebagai kategori pembanding.

**Tabel 4.8** Estimasi Parameter Regresi Logistik Biner

Variabel	$\hat{\beta}$	SE ( $\hat{\beta}$ )	Wald	P-value
Konstan	-1,959	2,498	0,615	0,433
Keluar cairan dari puting susu (1)	-2,247	2,514	0,799	0,371
Cekungan di kulit payudara (1)	-0,255	2,854	0,008	0,929
Nyeri di bagian tertentu payudara (1)	1,809	1,272	2,022	0,155
Penarikan puting susu ke dalam (1)	2,188	1,755	1,555	0,213
Kerutan di kulit payudara (1)	1,635	1,674	0,953	0,329
Benjolan di bawah ketiak (1)	-1,843	2,835	0,423	0,516
Perubahan bentuk dan ukuran payudara (1)	1,665	1,711	0,947	0,330
Durasi tidur malam (1)	-0,325	2,664	0,015	0,903
Durasi tidur malam (2)	-0,305	2,462	0,015	0,901
Durasi tidur malam (3)	0,670	2,752	0,059	0,808
Usia (1)	0,351	1,035	0,115	0,735

Pemodelan menggunakan seluruh variabel prediktor menghasilkan nilai estimasi parameter pada Tabel 4.8. Diperoleh statistik uji  $G^2$  yaitu 16,16, jika dibandingkan dengan  $\chi^2_{0,1;11}(17,28)$  dihasilkan keputusan gagal tolak  $H_0$  yang menandakan bahwa tidak terdapat variabel prediktor yang berpengaruh signifikan terhadap status pemeriksaan kanker payudara pada taraf signifikansi  $\alpha (0,1)$ .

Kemudian dilakukan seleksi variabel untuk menghindari hal multikolinearitas dan mendapatkan model terbaik. Seleksi variabel menggunakan *backward elimination* dilakukan dengan memasukkan semua variabel prediktor ke dalam model terlebih dahulu, kemudian melakukan eliminasi satu persatu hingga tersisa variabel prediktor yang signifikan. Model terbaik yaitu model yang memiliki nilai AIC terkecil. Langkah-langkah pada seleksi variabel *backward* pada pemodelan status pemeriksaan kanker payudara terlampir pada Lampiran 4.

**Tabel 4.9** Seleksi Variabel Backward

<b>Step</b>	<b>Variabel Prediktor Pembentuk Model</b>	<b>AIC</b>
1	- Keluar cairan dari puting susu - Cekungan di kulit payudara - Nyeri di bagian tertentu payudara - Penarikan puting susu - Kerutan di kulit payudara - Benjolan di bawah ketiak - Perubahan bentuk-ukuran payudara - Usia	59,2
2	- Keluar cairan dari puting susu - Nyeri di bagian tertentu payudara - Penarikan puting susu - Kerutan di kulit payudara - Benjolan di bawah ketiak - Perubahan bentuk-ukuran payudara - Usia	57,2
3	- Keluar cairan dari puting susu - Nyeri di bagian tertentu payudara - Penarikan puting susu - Kerutan di kulit payudara - Benjolan di bawah ketiak - Perubahan bentuk-ukuran payudara	55,5
4	- Nyeri di bagian tertentu payudara - Penarikan puting susu - Kerutan di kulit payudara - Benjolan di bawah ketiak - Perubahan bentuk-ukuran payudara	54,14
5	- Nyeri di bagian tertentu payudara - Penarikan puting susu - Benjolan di bawah ketiak - Perubahan bentuk-ukuran payudara	52,4
6	- Nyeri di bagian tertentu payudara - Penarikan puting susu - Perubahan bentuk-ukuran payudara	52,3
7	- Nyeri di bagian tertentu payudara - Penarikan puting susu	51,22

Berdasarkan Tabel 4.9, diperoleh sebanyak 7 *step* dalam melakukan pemodelan status pemeriksaan kanker payudara menggunakan seleksi *backward*. Gejala klinis yang pertama kali dikeluarkan dari model yaitu durasi tidur malam, dilanjutkan dengan mengeluarkan cekungan di kulit payudara pada *step* 2, usia pada *step* 3, dan seterusnya hingga *step* 7 perubahan bentuk dan ukuran payudara menjadi variabel yang dikeluarkan terakhir dari model. Model terbaik dengan seleksi *backward* menghasilkan AIC akhir yaitu 51,22. Seleksi variabel berhenti pada *step* 7 dengan keseluruhan variabel prediktor pembentuk model telah signifikan. Variabel prediktor yang terpilih pada pemodelan menggunakan seleksi *backward* yaitu nyeri di bagian tertentu payudara dan penarikan puting susu. Kemudian dilakukan pengujian parameter secara serentak dan parsial dengan nilai pada Tabel 4.10.

**Tabel 4.10** Estimasi Parameter Regresi Logistik Biner dengan Seleksi *Backward*

Variabel	$\hat{\beta}$	SE ( $\hat{\beta}$ )	Wald	P-value
Konstan	-1,7547	0,4524	15,0466	0,0001
Nyeri di bagian tertentu payudara (1)	1,79	0,8860	4,0804	0,0434
Penarikan puting susu (1)	1,8135	1,0066	3,2472	0,0716

Statistik uji *likelihood ratio test* pada pengujian secara serentak yaitu 11,48, jika dibandingkan dengan  $\chi^2_{0,1;2}$  (4,61) dihasilkan keputusan tolak  $H_0$  yang menunjukkan bahwa terdapat minimal satu variabel prediktor yang berpengaruh signifikan di dalam model pada signifikansi  $\alpha$  (0,1). Pada pengujian parameter secara parsial, terlihat nilai *Wald* variabel nyeri di bagian tertentu payudara dan penarikan puting susu yaitu 4,0804 dan 3,2472. Pada taraf signifikansi 5%, diperoleh variabel yang berpengaruh terhadap status pemeriksaan kanker payudara yaitu nyeri di bagian tertentu payudara. Jika dibandingkan dengan taraf signifikansi 10% yaitu  $\chi^2_{0,1;1}$  (2,71), diperoleh keputusan tolak  $H_0$  sehingga disimpulkan bahwa nyeri di bagian tertentu payudara dan penarikan puting susu berpengaruh terhadap status pemeriksaan kanker payudara. Koefisien  $\hat{\beta}$  pada variabel nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam bernilai positif yang

menunjukkan bahwa kedua variabel tersebut memiliki hubungan yang berbanding lurus dengan status pemeriksaan kanker payudara positif. Berikut adalah model logit yang terbentuk.

$$\ln \left( \frac{\hat{\pi}(x_i)}{1 - \hat{\pi}(x_i)} \right) = -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1$$

Berdasarkan model logit yang terbentuk pada seleksi *backward*, diperoleh model probabilitas sebagai berikut.

$$\hat{\pi}(x_i) = \frac{\exp \left( \begin{array}{l} -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 \\ + 1,8135 \text{ Penarikan puting susu}_1 \end{array} \right)}{1 + \exp \left( \begin{array}{l} -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 \\ + 1,8135 \text{ Penarikan puting susu}_1 \end{array} \right)}$$

Model probabilitas tersebut kemudian dihitung dengan mensubstitusikan nilai variabel prediktor ke dalam model. Dimisalkan, apabila seorang wanita memiliki gejala nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam maka probabilitas wanita tersebut mengalami status pemeriksaan kanker payudara positif sebesar 0,864. Lampiran 10 menampilkan probabilitas status pemeriksaan kanker payudara positif pada tiap responden. Kemudian dilakukan pengujian kesesuaian model untuk mengetahui model yang terbentuk telah sesuai. Diperoleh nilai *Chi-squared* sebesar 1,025, dengan menggunakan perbandingan  $\chi^2_{0,1;1}$  (2,71) maka dihasilkan keputusan gagal tolak  $H_0$  sehingga dapat disimpulkan bahwa model telah sesuai dan tidak terdapat perbedaan antara hasil pengamatan dengan kemungkinan hasil prediksi.

Wanita dengan gejala nyeri di bagian tertentu payudara memiliki kecenderungan 5,99 kali untuk mengalami status pemeriksaan kanker payudara positif dibandingkan wanita yang tidak memiliki nyeri di bagian tertentu payudara.

**Tabel 4.11** Odds Ratio Regresi Logistik dengan Seleksi *Backward*

Variabel	$\hat{\beta}$	OR
Nyeri di bagian tertentu payudara	1,79	5,9895
Penarikan puting susu	1,8135	6,1319

Wanita dengan gejala penarikan puting susu ke dalam cenderung mengalami status pemeriksaan kanker payudara positif 6,13 kali

dibandingkan wanita yang tidak memiliki gejala penarikan puting susu.

**Tabel 4.12** *Confusion Matrix* Regresi Logistik dengan Seleksi *Backward*

<i>Actual Membership</i>	<i>Predicted Membership</i>	
	Negatif	Positif
Negatif	31	5
Positif	6	7

Penentuan kelompok status kanker payudara menggunakan seleksi *backward*, menghasilkan *confusion matrix* pada Tabel 4.12 dengan ketepatan klasifikasi AUC yaitu 0,6998.

### 4.3 Regresi Logistik Biner-Seleksi *Forward*

Metode seleksi variabel menggunakan *forward* dilakukan dengan cara memasukkan satu persatu variabel prediktor hingga diperoleh AIC terkecil.

**Tabel 4.13** Seleksi Variabel *Forward*

Step	Variabel Prediktor Pembentuk Model	AIC
1	- Nyeri di bagian tertentu payudara	52,56
2	- Nyeri di bagian tertentu payudara - Penarikan puting susu	51,22

Tahapan seleksi variabel *forward* pada pemodelan status pemeriksaan kanker payudara terlampir pada Lampiran 6. Pada Tabel 4.13 terlihat gejala nyeri di bagian tertentu payudara masuk terlebih dahulu ke dalam model, kemudian terjadi penambahan gejala yaitu penarikan puting susu. Nyeri di bagian tertentu payudara masuk terlebih dahulu karena memiliki hubungan yang lebih kuat, kemudian diikuti penarikan puting susu. Variabel lain yang tidak dimasukkan karena dengan menambahkannya tidak memberikan peningkatan yang berarti. Model terbaik dengan seleksi *forward* menghasilkan AIC 51,22. Tabel 4.14 berikut menunjukkan variabel yang terpilih menggunakan seleksi *forward*.

**Tabel 4.14** Estimasi Parameter Regresi Logistik Biner dengan Seleksi *Forward*

Variabel	$\hat{\beta}$	SE ( $\hat{\beta}$ )	Wald	P-value
Konstan	-1,7547	0,4524	15,0466	0,0001
Nyeri di bagian tertentu payudara (1)	1,79	0,8860	4,0804	0,0434
Penarikan puting susu (1)	1,8135	1,0066	3,2472	0,0716

Statistik uji  $G^2$  pada pengujian parameter secara serentak sebesar 11,48, jika dibandingkan dengan  $\chi_{0,1;2}^2$  (4,61) diperoleh keputusan tolak  $H_0$  sehingga dapat disimpulkan bahwa terdapat minimal satu variabel prediktor yang berpengaruh di dalam model pada signifikansi  $\alpha$  (0,1). Pada pengujian parameter secara parsial menggunakan taraf signifikansi 5%, diperoleh variabel yang signifikan yaitu nyeri di bagian tertentu payudara. Jika digunakan dengan taraf signifikansi 10%, diperoleh keputusan tolak  $H_0$ . Hal itu ditunjukkan oleh nilai *Wald* variabel nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam masing-masing memiliki yang lebih dari  $\chi_{0,1;1}^2$  (2,71) sehingga kedua variabel tersebut berpengaruh terhadap status pemeriksaan kanker payudara secara signifikan.

Disamping itu, diketahui pula variabel nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam memiliki koefisien  $\hat{\beta}$  yang bernilai positif sehingga kedua variabel tersebut memiliki hubungan yang berbanding lurus dengan status pemeriksaan kanker payudara positif. Apabila seorang wanita mengalami gejala nyeri di bagian tertentu payudara atau penarikan puting susu ke dalam maka peluang status pemeriksaan kanker payudara positif semakin besar. Berikut model logit yang terbentuk.

$$\ln \left( \frac{\hat{\pi}(x_i)}{1 - \hat{\pi}(x_i)} \right) = -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1$$

Fungsi logit yang terbentuk berdasarkan seleksi variabel metode *forward* menghasilkan model probabilitas berikut.

$$\hat{\pi}(x_i) = \frac{\exp \left( -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1 \right)}{1 + \exp \left( -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1 \right)}$$

Model probabilitas tersebut mengindikasikan bahwa jika wanita dengan gejala nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam memiliki probabilitas mengalami status pemeriksaan kanker payudara positif sebesar 0,864. Sementara itu, probabilitas wanita dengan gejala nyeri di bagian tertentu payudara

mengalami status pemeriksaan kanker payudara positif sebesar 0,51. Pada pengujian kesesuaian model diperoleh nilai *Chi-squared* sebesar 1,025. Nilai tersebut menghasilkan keputusan gagal tolak  $H_0$  dengan menggunakan perbandingan  $\chi^2_{0,1;1}$  (2,71) sehingga dapat disimpulkan bahwa model telah sesuai.

**Tabel 4.15** *Odds Ratio* Regresi Logistik dengan Seleksi *Forward*

Variabel	$\hat{\beta}$	OR
Nyeri di bagian tertentu payudara	1,79	5,9895
Penarikan puting susu	1,8135	6,1319

Kecenderungan wanita dengan gejala nyeri di bagian tertentu payudara untuk mengalami status pemeriksaan kanker payudara positif 5,99 kali dibandingkan wanita tanpa gejala nyeri di bagian tertentu payudara. Begitu juga pada *odds ratio* penarikan puting susu sebesar 6,13 menunjukkan wanita dengan gejala penarikan puting susu ke dalam cenderung untuk mengalami status pemeriksaan kanker payudara positif 6,13 kali dibandingkan wanita yang tidak memiliki gejala penarikan puting susu.

**Tabel 4.16** *Confusion Matrix* Regresi Logistik dengan Seleksi *Forward*

<i>Actual Membership</i>	<i>Predicted Membership</i>	
	Negatif	Positif
Negatif	31	5
Positif	6	7

Tabel 4.16 menunjukkan adanya kesalahan prediksi yaitu responden dengan status pemeriksaan kanker payudara positif diklasifikasikan pada kelompok responden status negatif. Begitu juga sebaliknya, responden kanker payudara negatif diklasifikasikan pada kelompok responden dengan status pemeriksaan kanker payudara positif. Ketepatan klasifikasi (AUC) pada model regresi dengan seleksi *forward* sebesar 0,6998.

#### 4.4 Regresi Logistik Biner-Seleksi *Stepwise*

Seleksi *stepwise* merupakan kombinasi dari metode *backward* dan *forward*. Pada tiap tahap seleksi *stepwise*, memungkinkan adanya sebuah variabel untuk dimasukkan dalam model kemudian dikombinasikan dengan mengeluarkan variabel prediktor yang tidak signifikan dalam model. Langkah-langkah

seleksi variabel *stepwise* terlampir pada Lampiran 8. Tabel 4.17 menampilkan variabel yang terpilih menggunakan seleksi variabel *stepwise* yaitu nyeri di bagian tertentu payudara dan penarikan puting susu.

**Tabel 4.17** Estimasi Parameter Regresi Logistik Biner dengan Seleksi *Stepwise*

Variabel	$\hat{\beta}$	$SE(\hat{\beta})$	Wald	P-value
Konstan	-1,7547	0,4524	15,0466	0,0001
Nyeri di bagian tertentu payudara (1)	1,79	0,8860	4,0804	0,0434
Penarikan puting susu (1)	1,8135	1,0066	3,2472	0,0716

Nilai estimasi parameter pada variabel tersebut kemudian dilakukan pengujian signifikansi secara serentak dan parsial. Pada pengujian signifikansi parameter secara serentak, menghasilkan *likelihood ratio test* sebesar 11,48 dengan  $\chi^2_{0,1;2}$  yaitu 4,61. Nilai tersebut menunjukkan keputusan tolak  $H_0$  karena  $G^2 > \chi^2_{0,1;2}$  sehingga dapat disimpulkan bahwa minimal terdapat satu variabel prediktor yang berpengaruh signifikan dalam model. Pada pengujian parameter secara parsial, diketahui bahwa nyeri di bagian tertentu payudara memiliki pengaruh terhadap status pemeriksaan kanker payudara pada taraf signifikansi 5%. Apabila taraf signifikansi ditingkatkan, diperoleh gejala yang berpengaruh terhadap status pemeriksaan kanker payudara yaitu nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam dengan taraf signifikansi 10%.

Tabel 4.17 juga menunjukkan variabel nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam memiliki nilai estimasi parameter bernilai positif yang menunjukkan bahwa apabila seorang wanita mengalami gejala-gejala tersebut maka peluang status pemeriksaan kanker payudara positif semakin tinggi. Kemudian, didapatkan model terbaik pada seleksi *stepwise* adalah sebagai berikut.

$$\ln\left(\frac{\hat{\pi}(x_i)}{1 - \hat{\pi}(x_i)}\right) = -1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1$$

dengan model probabilitas sebagai berikut.

$$\hat{\pi}(x_i) = \frac{\exp\left(-1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1\right)}{1 + \exp\left(-1,7547 + 1,79 \text{ Nyeri di bagian tertentu payudara}_1 + 1,8135 \text{ Penarikan puting susu}_1\right)}$$

Model probabilitas tersebut menunjukkan peluang seorang wanita mengalami status pemeriksaan kanker payudara positif. Contoh, peluang wanita dengan gejala nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam untuk mengalami status pemeriksaan kanker payudara positif sebesar 0,864. Dilakukan pengujian kesesuaian model dengan statistik uji *Chi-squared* yaitu 1,025 yang menunjukkan bahwa model telah sesuai.

**Tabel 4.18** Odds Ratio Regresi Logistik dengan Seleksi *Stepwise*

Variabel	$\hat{\beta}$	OR
Nyeri di bagian tertentu payudara	1,79	5,9895
Penarikan puting susu	1,8135	6,1319

Kecenderungan wanita dengan nyeri di bagian payudara terhadap status pemeriksaan kanker payudara positif 5,99 kali dibandingkan wanita tanpa gejala nyeri di bagian tertentu payudara. Disamping itu, wanita dengan gejala penarikan puting susu ke dalam cenderung mengalami status pemeriksaan kanker payudara positif 6,13 kali dibandingkan wanita yang tidak mengalami penarikan puting susu.

**Tabel 4.19** Confusion Matrix Regresi Logistik dengan Seleksi *Stepwise*

Actual Membership	Predicted Membership	
	Negatif	Positif
Negatif	31	5
Positif	6	7

Terlihat pada Tabel 4.19, prediksi dari pemodelan status pemeriksaan kanker payudara menunjukkan adanya kesalahan prediksi. Dengan menggunakan *confusion matrix*, diperoleh AUC pada pemodelan status pemeriksaan kanker payudara menggunakan seleksi *stepwise* yaitu sebesar 0,6998.

#### 4.5 Regresi Logistik Biner-*Genetic Algorithm*

Analisis regresi logistik biner menggunakan seleksi variabel *genetic algorithm* dilakukan dengan memilih kombinasi variabel yang menghasilkan nilai ketepatan klasifikasi tertinggi. Tahap

pertama *genetic algorithm* yaitu membentuk representasi populasi awal yang terdiri dari kromosom-kromosom. Kromosom ini berisi gen yang berperan sebagai variabel prediktor.

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
1	1	1	0	1	1	0	0	1

**Gambar 4.12** Ilustrasi Kromosom Awal

Gambar 4.12 menampilkan ilustrasi kromosom awal yang terbentuk. Pada penelitian ini digunakan tipe representasi genetik yaitu biner, nilai 1 menunjukkan variabel prediktor yang masuk ke dalam model regresi logistik sedangkan nilai 0 menunjukkan variabel yang tidak masuk ke dalam model. Proses *labelling* tersebut dilakukan dengan membangkitkan angka biner secara acak pada setiap variabel. Panjang kromosom yang terbentuk sesuai dengan jumlah variabel prediktor pada penelitian sehingga tiap kromosom akan berisi 9 gen. Langkah awal dalam *genetic algorithm* yaitu inisialisasi populasi awal yang terdiri dari 80 kromosom.

**Tabel 4.20** Ilustrasi Kromosom Pada Populasi Awal

Kromosom	Variabel								
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
1	1	1	0	1	1	1	0	0	0
2	0	1	1	0	0	1	0	0	1
3	1	0	1	1	0	1	1	1	1
$\vdots$									
79	1	1	1	1	1	1	1	0	0
80	0	0	0	1	0	1	1	1	0

Setelah diperoleh populasi awal seperti pada Tabel 4.20. Tahap berikutnya yaitu melakukan evaluasi terhadap nilai *fitness* pada tiap kromosom yang terbentuk pada populasi awal sebagai kriteria solusi terbaik. Fungsi *fitness* yang digunakan dalam penelitian ini adalah AUC. Ilustrasi nilai *fitness* yang diperoleh dari masing-masing kromosom dalam populasi awal dapat dilihat pada Tabel 4.21.

**Tabel 4.21** Ilustrasi Nilai *Fitness* Tiap Kromosom

Kromosom	Variabel									Nilai <i>fitness</i>
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	
1	1	1	0	1	1	1	0	0	0	0,56309
2	0	1	1	0	0	1	0	0	1	0,67559
3	1	0	1	1	0	1	1	1	1	0,63036
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
79	1	1	1	1	1	1	1	0	0	0,61131
80	0	0	0	1	0	1	1	1	0	0,54524

Kemudian, melakukan pembentukan kromosom induk. Pemilihan kromosom pada populasi awal dilakukan menggunakan metode seleksi *roulette wheel* sehingga kromosom yang terpilih akan menjadi orang tua atau induk pada generasi berikutnya. Langkah-langkah metode *roulette wheel* yaitu menentukan proporsi nilai *fitness* dan menentukan nilai kumulatif proporsi nilai *fitness*. Semakin besar nilai *fitness* suatu kromosom, maka peluang kromosom tersebut dapat terpilih menjadi individu baru pada generasi berikutnya akan semakin besar pula. Kromosom terpilih dilihat melalui nilai proporsi *fitness* kumulatif pertama dengan nilai yang lebih tinggi dibandingkan bilangan *random*.

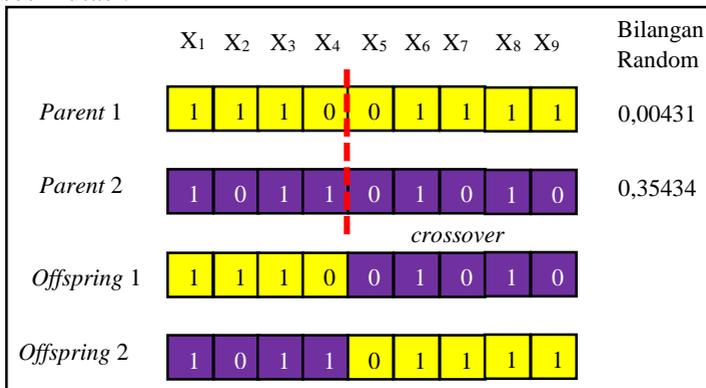
**Tabel 4.22** Ilustrasi Seleksi *Roulette Wheel*

Kromosom	Nilai <i>Fitness</i>	Proporsi Nilai <i>Fitness</i>	Proporsi Kumulatif	Bilangan Random
1	0,56309	0,0090559	0,00906	0,02644
2	0,67559	0,0108651	0,01992	0,68218
3	0,63036	0,0101377	0,03006	0,42678
⋮	⋮	⋮	⋮	⋮
79	0,61131	0,0098313	0,99123	0,10082
80	0,54524	0,0087688	1,00000	0,05539

Terlihat pada Tabel 4.22 kromosom 3 memiliki nilai proporsi kumulatif lebih dari nilai random pertama (0,0264) sehingga kromosom 3 terpilih sebagai kromosom induk. Proses tersebut dilanjutkan hingga diperoleh kromosom induk sejumlah ukuran populasi.

Tahap selanjutnya yaitu melakukan *crossover* dengan tujuan menghasilkan kromosom baru dari persilangan dua kromosom

induk. Pada penelitian ini, digunakan *single point crossover* dengan probabilitas *crossover* adalah 0,8. Dua kromosom induk akan dipotong pada suatu titik random dan menukarkan isi nilai subset kromosom tersebut. Kromosom induk yang mengalami *crossover* yaitu kromosom dengan bilangan random kurang dari 0,8. Ilustrasi pada *single point crossover* ditampilkan pada Gambar 4.13. Berdasarkan Gambar 4.13, terlihat bilangan random pada induk 1 dan induk 2 kurang dari probabilitas *crossover* 0,8, sehingga terjadi pindah silang pada kedua kromosom induk tersebut. Kemudian, menentukan lokasi titik pindah silang dengan membangkitkan bilangan random sejumlah variabel prediktor yaitu antara 1 sampai 9. Pada ilustrasi diperoleh titik potong ke-4 sehingga kromosom induk akan menghasilkan anak (*offspring*) dengan nilai gen merupakan perpaduan antara kedua kromosom induk tersebut. Proses ini dilanjutkan untuk semua pasang kromosom yang memiliki nilai bilangan random kurang dari probabilitas *crossover*. Hasil persilangan akan dilanjutkan pada proses mutasi.



**Gambar 4.13** Ilustrasi Proses *Crossover*

Mutasi dilakukan dengan mengubah nilai suatu gen, yaitu apabila gen bernilai 1 diubah menjadi 0 sedangkan 0 diubah menjadi 1. Proses ini bertujuan mendapatkan solusi lokal optimal dengan membangkitkan bilangan random yang relatif kecil. Ilustrasi mutasi digambarkan pada Gambar 4.14.

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>
Sebelum mutasi	0,03	0,51	0,61	0,69	0,08	0,86	0,4	0,5	0,45
Sesudah mutasi	1	0	0	1	0	1	0	1	0
	0	0	1	0	1	1	1	1	1

**Gambar 4.14** Ilustrasi Proses Mutasi

Langkah awal mutasi yaitu dengan membangkitkan bilangan random pada setiap gen. Kemudian, bilangan random tersebut dibandingkan dengan probabilitas mutasi yang telah ditentukan yaitu 0,1. Apabila bilangan random yang dihasilkan kurang dari 0,1 maka gen tersebut akan mengalami proses mutasi. Terlihat pada ilustrasi Gambar 4.14, gen 1 dan 5 memiliki bilangan random kurang dari 0,1 sehingga terjadi proses mutasi pada kedua gen tersebut. Tahap berikutnya yaitu melakukan proses elitisme agar kromosom terbaik dalam populasi dapat bertahan yaitu dengan melakukan penggandaan individu. Proses *genetic algorithm* dilanjutkan hingga diperoleh nilai *fitness* yang konvergen.

**Tabel 4.23** Estimasi Parameter Regresi Logistik Biner-*Genetic Algorithm*

Variabel	$\hat{\beta}$	SE ( $\hat{\beta}$ )	Wald	P-value
Konstan	-2,0639	0,5502	14,0700	0,0002
Keluar cairan dari puting susu (1)	-0,6660	1,8635	0,1274	0,7208
Nyeri di bagian tertentu payudara (1)	1,6679	0,9732	2,9378	0,0866
Penarikan puting susu (1)	2,3533	1,5335	2,3562	0,1249
Benjolan di bawah ketiak (1)	-2,3584	1,9697	1,4328	0,2312
Perubahan bentuk dan ukuran payudara (1)	2,2478	1,3995	2,5792	0,1082
Usia (1)	0,4411	0,9330	0,2237	0,6364

Seleksi variabel dengan *genetic algorithm* menghasilkan variabel yang terpilih yaitu keluar cairan dari puting susu, nyeri di bagian tertentu payudara, penarikan puting susu ke dalam, benjolan di bawah ketiak, perubahan bentuk dan ukuran payudara, dan usia. Kemudian dilakukan pengujian signifikansi parameter secara

serentak dan parsial dengan hasil pada Tabel 4.23. Pemodelan status pemeriksaan kanker payudara menggunakan seleksi *genetic algorithm* menghasilkan AIC 56,031 dengan nilai estimasi parameter pada Tabel 4.20. Dilakukan pengujian parameter secara serentak dengan *likelihood ratio test* sebesar 14,67, jika dibandingkan dengan  $\chi^2_{0,1;6}$  (10,65) diperoleh kesimpulan bahwa terdapat minimal satu variabel prediktor yang berpengaruh signifikan terhadap model. Pada pengujian parameter secara parsial menggunakan  $\alpha$  (0,1), variabel prediktor yang signifikan terhadap status pemeriksaan kanker payudara yaitu nyeri di bagian tertentu payudara. Hal itu ditunjukkan oleh nilai *Wald* pada variabel tersebut lebih dari  $\chi^2_{0,1;1}$  (2,71). Apabila taraf signifikansi  $\alpha$  (0,1) ditingkatkan maka variabel perubahan bentuk dan ukuran payudara juga memiliki pengaruh terhadap status pemeriksaan kanker payudara.

**Tabel 4.24** Pemodelan Kembali Regresi Logistik Biner-*Genetic Algorithm*

Variabel	$\hat{\beta}$	SE ( $\hat{\beta}$ )	Wald	P-value
Konstan	-1,7528	0,456	14,77634	0,0001
Nyeri di bagian tertentu payudara (1)	1,8996	0,8658	4,813636	0,0282
Perubahan bentuk dan ukuran payudara (1)	1,4333	0,9161	2,446096	0,1177

Tabel 4.24 merupakan hasil estimasi parameter menggunakan variabel nyeri di bagian tertentu payudara dan perubahan bentuk dan ukuran payudara. Variabel nyeri di bagian tertentu payudara dan penarikan puting susu ke dalam diketahui memiliki koefisien  $\hat{\beta}$  yang bernilai positif yang mengartikan bahwa kedua variabel tersebut memiliki hubungan yang berbanding lurus dengan status pemeriksaan kanker payudara positif. Pengujian parameter secara serentak menghasilkan nilai  $G^2$  sebesar 10,55. Nilai tersebut memberikan keputusan tolak  $H_0$  jika dibandingkan dengan  $\chi^2_{0,1;2}$  (4,61) sehingga disimpulkan bahwa terdapat minimal satu variabel prediktor yang berpengaruh di dalam model secara signifikan. Pada pengujian parameter secara parsial, diperoleh variabel yang signifikan yaitu nyeri di bagian

tertentu payudara dengan taraf signifikansi 5%. Apabila taraf signifikansi tersebut dinaikkan hingga 12%, terjadi penambahan gejala klinis yang signifikan yaitu perubahan bentuk dan ukuran payudara sehingga model yang terbentuk adalah sebagai berikut.

$$\ln\left(\frac{\hat{\pi}(x_i)}{1 - \hat{\pi}(x_i)}\right) = -1,7528 + 1,8996 \text{ Nyeri di bagian tertentu payudara}_1 + 1,4333 \text{ Perubahan bentuk dan ukuran payudara}_1$$

dengan model probabilitas adalah sebagai berikut.

$$\hat{\pi}(x_i) = \frac{\exp\left(-1,7528 + 1,8996 \text{ Nyeri di bagian tertentu payudara}_1 + 1,4333 \text{ Perubahan bentuk dan ukuran payudara}_1\right)}{1 + \exp\left(-1,7528 + 1,8996 \text{ Nyeri di bagian tertentu payudara}_1 + 1,4333 \text{ Perubahan bentuk dan ukuran payudara}_1\right)}$$

Kemudian, dilakukan uji kesesuaian model dengan nilai *Chi-squared* yaitu 0,258 yang menunjukkan bahwa model telah sesuai. Contoh, seorang wanita dengan gejala nyeri di bagian tertentu payudara dan perubahan bentuk dan ukuran payudara maka wanita tersebut memiliki probabilitas untuk mengalami status pemeriksaan kanker payudara positif sebesar 0,61.

**Tabel 4.25** *Odds Ratio* Regresi Logistik Biner-Genetic Algorithm

Variabel	$\hat{\beta}$	OR
Nyeri di bagian tertentu payudara	1,8996	6,6832
Perubahan bentuk dan ukuran payudara	1,4333	4,1925

Wanita dengan gejala nyeri di bagian tertentu payudara cenderung mengalami status pemeriksaan kanker payudara positif 6,7 kali dibandingkan wanita yang tidak mengalami nyeri di bagian tertentu payudara. *Odds ratio* variabel perubahan bentuk dan ukuran payudara yaitu 4,1925 yang menjelaskan bahwa wanita dengan perubahan bentuk dan ukuran payudara cenderung mengalami status pemeriksaan kanker payudara positif 4,2 kali dibandingkan wanita tanpa perubahan bentuk dan ukuran payudara.

**Tabel 4.26** *Confusion Matrix* Regresi Logistik Biner-Genetic Algorithm

<i>Actual Membership</i>	<i>Predicted Membership</i>	
	Negatif	Positif
Negatif	33	3
Positif	7	6

Pemodelan status pemeriksaan kanker payudara menggunakan *genetic algorithm* menghasilkan AIC yaitu 52,15 dengan ukuran ketepatan klasifikasi AUC yaitu sebesar 0,6891.

#### 4.6 Evaluasi Metode Terbaik

Setelah melakukan pemodelan menggunakan regresi logistik biner metode *backward*, *forward*, *stepwise*, dan *genetic algorithm*, dilakukan perbandingan dari keseluruhan metode. Hasil ketepatan klasifikasi pada tiap metode seleksi terangkum pada Tabel 4.27.

**Tabel 4.27** Perbandingan Kebaikan Metode Seleksi

<b>Metode Seleksi</b>	<b>Variabel Signifikan</b>	<b>AUC</b>
<i>Backward</i>	-Nyeri di bagian tertentu payudara -Penarikan puting susu	0,6998
<i>Forward</i>	-Nyeri di bagian tertentu payudara -Penarikan puting susu	0,6998
<i>Stepwise</i>	-Nyeri di bagian tertentu payudara -Penarikan puting susu	0,6998
<i>Genetic algorithm</i>	-Nyeri di bagian tertentu payudara -Perubahan bentuk dan ukuran payudara	0,6891

Terlihat ketepatan klasifikasi tertinggi menggunakan AUC yaitu pemodelan menggunakan seleksi variabel *backward*, *forward*, dan *stepwise*. Namun ketepatan klasifikasi yang dihasilkan pada pemodelan regresi logistik dengan *genetic algorithm* juga tidak jauh berbeda sehingga pemodelan status pemeriksaan kanker payudara menggunakan seleksi variabel *genetic algorithm* juga perlu dipertimbangkan.

*(Halaman ini sengaja dikosongkan)*

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Gejala yang paling banyak dialami oleh responden dengan status pemeriksaan kanker payudara positif dan negatif secara urut yaitu nyeri di bagian tertentu payudara, penarikan puting susu ke dalam, kerutan di kulit payudara, serta perubahan bentuk dan ukuran payudara. Responden dengan status positif dan negatif masing-masing memiliki rata-rata usia 45 tahun dan 32 tahun. Pada karakteristik durasi tidur, hampir 50 persen responden dengan status pemeriksaan positif maupun negatif memiliki durasi tidur  $\geq 5-6$  jam.
2. Model terbaik diperoleh menggunakan seleksi *backward*, *forward*, dan *stepwise*. Namun pemodelan menggunakan *genetic algorithm* juga perlu dipertimbangkan karena perbedaan ketepatan klasifikasi yang tidak begitu jauh. Gejala klinis yang perlu diperhatikan yaitu khususnya nyeri di bagian tertentu payudara, penarikan puting susu ke dalam, dan perubahan bentuk dan ukuran payudara. Wanita yang mengalami gejala nyeri di bagian tertentu payudara memiliki kecenderungan status pemeriksaan kanker payudara positif 5,99 kali dibandingkan wanita tanpa nyeri di bagian tertentu payudara. Kecenderungan wanita dengan gejala penarikan puting susu ke dalam untuk mengalami status pemeriksaan kanker payudara positif 6,13 kali dibandingkan wanita tanpa gejala penarikan puting susu. Wanita dengan perubahan bentuk dan ukuran payudara cenderung mengalami status pemeriksaan kanker payudara

positif 4,2 kali dibandingkan wanita tanpa perubahan bentuk dan ukuran payudara.

## **5.2 Saran**

Berdasarkan kesimpulan yang diperoleh, dirumuskan saran sebagai pertimbangan selanjutnya yaitu.

1. Mengumpulkan lebih banyak data, khususnya data status pemeriksaan kanker payudara positif agar diperoleh hasil analisis yang lebih baik.
2. Bagi masyarakat, diharapkan dapat lebih sadar mengenai penyakit kanker payudara dengan mengenali gejala-gejala kanker khususnya jika terjadi nyeri di bagian tertentu payudara, penarikan puting susu ke dalam, dan perubahan bentuk dan ukuran payudara.

## DAFTAR PUSTAKA

- Agresti, A. (2013). *Categorical Data Analysis*. New Jersey: John Wiley & Sons, Inc.
- Barton, M. B., Elmore, J. G., & Fletcher, S. W. (1999). Breast Symptoms Among Women Enrolled in A Health Maintenance Organization: Frequency, Evaluation, and Outcome. *Annals of Internal Medicine*, 130(8), 651-657.
- Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation Measures for Models Assesment Over Imbalanced Datasets. *Journal of Information Engineering and Applications*, 3, 27-38.
- Besedovsky, L., Lange, T., & Born, J. (2012). Sleep and Immune Function. *European Journal of Physiology*, 463(1), 121-137.
- Dumitrescu, R. G., & Cotarla, I. (2005). Understanding Breast Cancer Risk-Where Do We Stand In 2005? *Journal of Cellular and Molecular Medicine*, 9, 208-221.
- Eberl, M. M., Philips, R. L., Lamberts, H., Okkes, I., & Maboney, M. C. (2008). Characterizing Breast Symptoms in Family Practice. *Analysis of Family Medicine*, 6(6), 528-533.
- Fang, H. F., Miao, N. F., Chen, C. D., Sithole, T., & Chung, M. H. (2015). Risk of Cancer in Patients with Insomnia, Parasomnia, and Obstructive Sleep Apnea: A Nationwide Nested Case-Control Study. *Journal Of Cancer*, 6, 1140-1147.
- Fidler, I. J. (1990). Critical Factors in The Biology of Human Cancer Metastasis. *Cancer Research*, 50, 6130-6138.
- Gayou, O., Das, S. K., Zhou, S.-M., Marks, L. B., Parda, D. S., & Miften, M. (2008). A Genetic Algorithm For Variable Selection in Logistic Regression Analysis of Radiotherapy Treatment Outcomes. *Medical Pyhsics*, 35(12), 5426-5433.

- Goldberg, E. D. (1989). *Genetic Algorithm in Search, Optimization, and Machine Learning*. Illinois: Addison Wesley Professional.
- Grunfeld, E. A., Ramirez, A. J., & Richards, M. A. (2002). Women's Knowledge And Beliefs Regarding Breast Cancer. *British Journal of Cancer*, 86, 1373 – 1378.
- Haupt, S. E., & Haupt, R. L. (2004). *Practical Genetic Algorithm*. New Jersey: John Wiley & Sons Inc.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. New York: John Wiley & Sonc, Inc.
- IARC, I. (2018). *Cancer Today*. Diakses pada tanggal 10 Februari 2019 dari International Agency for Research on Cancer: <http://gco.iarc.fr>.
- Jadaan, O. A., Rajamani, L., & Rao, C. R. (2005). Improved Selection Operator for GA. *Journal of Theoretical and Applied Information Technology (JATIT)*, 4(4), 269-277.
- Johnson, P., Vanderwater, L., Wilson, W., Maruff, P., Savage, G., Graham, P., Macaulay, L. S., Ellis, K. A., Szoeki, C., Martins, R. N., Rowe, C. C., Masters, C. L., Ames, D., & Zhang, P. (2014). Genetic Algorithm with Logistic Regression for Prediction of Progression to Alzheimer's Disease. *BMC Bioinformatics*, 15(16), 1-14.
- Johnson, R. A., & Winchern, D. W. (2007). *Applied Multivariate Statistical Analysis (6th ed)*. New Jersey: Pearson Prentice Hall.
- Kaneita, Y., Ohida, T., Osaki, Y., MD, Tanihata, T., Minowa, M., Suzuki K., Wada K., Kanda H., & Hayashi, K. (2006). Insomnia Among Japanese Adolescents: A Nationwide Representative Survey. *Sleep*, 29, 1543-1550.
- KEMENKES. (2017). *Direktorat Pencegahan dan Pengendalian Penyakit Tidak Menular*. Diakses pada tanggal 10 Februari 2019 dari Kementerian Kesehatan Republik Indonesia: <http://www.p2ptm.kemkes.go.id>

- Komite Penanggulangan Kanker Nasional. (2017). *Panduan Penatalaksanaan Kanker Payudara*. Jakarta: Kementerian Kesehatan Republik Indonesia.
- Lumachi, F., Ermani, M., Brandes, A. A., Boccagni, P., Polistina, F., Basso, S. M., Favia, G., & D'Amico, D. F. (2002). Breast Complaints and Risk of Breast Cancer. *Biomedicine & Pharmacotherapy*, *56*, 88-92.
- Lumbantobing, S. M. (2008). *Gangguan Tidur*. Jakarta: Fakultas Kedokteran Universitas Indonesia.
- McElroy, J. A., Newcomb, P. A., Titus-Ernstoff, L., Trentham-Dietz, A., Hampton, J. M., & Egan, K. M. (2006). Duration of Sleep and Breast Cancer Risk in A Large Population Based Case-Control Study. *Journal of Sleep Research*, *15*, 241-249.
- Nishiura, M., Tamura, A., Nagai, H., & Matsushima, E. (2015). Assessment of Sleep Disturbance in Lung Cancer Patients: Relationship Between Sleep Disturbance and Pain, Fatigue, Quality of Life, and Psychological Distress. *Palliative and Supportive Care*, *13*(3), 575-581.
- Palesh, O., Aldridge, A., Zeitzer, J. M., Koopman, C., Neri, E., Giese, J., Jo, B., Kraemer, H., Nouriani, B., & Spiegel, D. (2014). Measured Sleep Disruption as A Predictor of Survival Among Women with Advanced Breast Cancer. *Sleep*, *37*, 837-842.
- Purwantaka, R. I. (2010). *Klasifikasi Penderita Kanker Payudara dengan Pendekatan Metode Regresi Logistik*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Rasjidi, I. (2009). *Deteksi Dini dan Pencegahan Kanker Pada Wanita*. Jakarta: Sagung Seto.
- Ryerson, A. B., Miller, J., & Ehemann, C. R. (2015). Reported Breast Symptoms in The National Breast and Cervical Cancer Early Detection Program. *Cancer Causes Control*, *26*(5), 733-740.
- Sabiston, D. C. (1995). *Buku Ajar Bedah*. Jakarta: EGC.

- Seltzer, M. H. (2004). Breast Complaints, Biopsies, and Cancer Correlated with Age in 10,000 Consecutive New Surgical Referrals. *The Breast Journal*, 10(2), 111-117.
- Shaharudin, S. H., Sulaiman, S., Emran, N. A., Shahril, M. R., & Hussain, S. N. (2011). The Use of Complementary and Alternative Medicine Among Malay Breast Cancer Survivors. *Alternative Therapies*, 17(1), 50-56.
- Suyanto. (2005). *Algoritma Genetika dalam Matlab*. Yogyakarta: Andi Offset.
- Suzanna, E., Sirait, T., Rahayu, P., Shalmon, G., Anwar, E., Andalusia, R., Harjati, & Panigoro, S. (2012). SUPLEMEN: Registrasi Kanker Berbasis Rumah Sakit di Rumah Sakit Kanker Dharmais-Pusat Kanker Nasional 1993-2007. *Indonesian Journal of Cancer*, 6(4), 1-12.
- Walker, S., Hyde, C., & Hamilton, W. (2014). Risk of Breast Cancer in Symptomatic Women In Primary Care. *British Journal of General Practice*, 64(629), 788-793.
- Wu, A. H., Wang R., Koh, W. P., Stanczyk, F. Z., Lee, H. P., & Yu, M. C. (2008). Sleep Duration, Melatonin and Breast Cancer Among Chinese Women in Singapore. *Carcinogenesis*, 29, 1244-1248.
- Yadav, P., & Jaroli, D. P. (2010). Breast Cancer: Awareness and Risk Factors in College-Going Younger Age Group Women in Rajasthan. *Asian Pacific Journal of Cancer Prevention*, 11, 319-322.

## LAMPIRAN

**Lampiran 1.** Data Penelitian

Status	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>
1	0	0	0	0	0	0	1	2	0
1	0	0	0	0	0	0	0	2	0
1	0	0	0	0	0	0	0	2	0
1	0	0	0	0	1	0	0	1	0
1	1	0	1	1	1	0	0	2	0
0	0	0	0	0	0	0	0	2	1
0	0	0	0	0	0	0	0	2	0
1	1	1	1	1	1	1	1	3	1
0	1	0	0	1	1	1	1	0	0
1	0	1	1	1	1	1	1	2	0
1	0	1	1	0	0	0	1	3	1
1	0	0	0	0	0	0	0	3	0
0	0	0	0	0	0	0	0	2	0
1	0	0	0	0	0	0	0	1	1
1	0	0	0	1	0	1	1	0	0
1	0	0	1	0	0	0	0	2	1
0	0	0	0	0	0	0	0	2	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	2	0
0	0	0	0	0	0	0	0	1	0

**Lampiran 2.** *Syntax* R Regresi Logistik Biner

```

data<-read.csv("E:/TA DINDA GALUH/runTA/dataTA4.csv",
header=TRUE, sep=",")

for(k in 1:9) {
  data[, k] <- as.factor(data[, k])
}

fullmodel = glm(Status~.,data=data, family = binomial(link='logit'))

#backward
bkwrд_model = step(fullmodel, direction = "backward")
summary(bkwrд_model)
logitgof(data$Status, fitted(bkwrд_model))

#forward
nothing<-glm(Status~1, family=binomial(link='logit'), data=data)
frwrд_model<-step(nothing,scope=list(lower=formula(nothing),
upper=formula(fullmodel)),
direction="forward")
summary(frwrд_model)
logitgof(data$Status, fitted(frwrд_model))

#stepwise
stpwise_model<-step(nothing,scope=list(lower=formula(nothing),
upper=formula(fullmodel)),
direction="both")
summary(stpwise_model)
logitgof(data$Status, fitted(stpwise_model))

fitted.results <- predict(fullmodel,newdata=data,type='response')
fitted.results <- ifelse(fitted.results > 0.5,1,0)
cm <- confusionMatrix(as.factor(fitted.results), data$Status)
sensi <- cm$byClass['Sensitivity']
speci <- cm$byClass['Specificity']
AUC <- 0.5*(sensi+speci)

```

### Lampiran 3. *Syntax R Genetic Algorithm*

```

xdata <- subset(data, select=-Status)
ydata <- data$Status
library(ROCR) # Fungsi fitness
fitness <- function(string){
  inc <- which(string==1)
  xdata <- xdata[,inc]
  b <- data.frame(cbind(ydata,xdata))
  model <- glm(as.factor(ydata)~.,data=b,
family=binomial(link='logit'))
  fitted.results <- predict(model,b,type='response')
  fitted.results <- ifelse(fitted.results > 0.5,1,0)
  pr <- prediction(fitted.results, ydata)
  prf <- performance(pr, measure = "tpr", x.measure = "fpr")
  auc <- performance(pr, measure = "auc")
  auc <- auc@y.values[[1]]
  return(auc)
}
library(GA)
gaControl("binary" = list(selection = "ga_rwSelection"))
GA_REGLOG <- ga("binary", fitness = fitness, nBits = ncol(xdata),
  names = colnames(xdata),
  selection = gaControl("binary")$selection,
  popSize = 80, pcrossover = 0.8, pmutation = 0.1, run = 5,
  elitism = 5, maxiter = 50, keepBest = TRUE,
  parallel = T, seed = 1112)
summary(GA_REGLOG)
f <- GA_REGLOG@solution # Variabel terpilih
f <- f[1,]
ff <- c(1:9)
fff <- f*ff
f1 <- sort(fff)
f2 <- f1[(10-sum(f)):9]
xt1 <- xdata[,f2]
newdata <- cbind(ydata,xt1)

```

## Lampiran 4. Tahapan Seleksi *Backward*

Start: AIC=64.54

Status ~ Keluar.cairan.dari.puting.susu + Cekungan.di.kulit.payudara + Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam + Kerutan.di.sekitar.payudara + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukuran.payudara + Durasi.tidur.malam + Usia

	Df	Deviance	AIC
- Durasi.tidur.malam	3	41.200	59.200
- Cekungan.di.kulit.payudara	1	40.545	62.545
- Usia	1	40.649	62.649
- Benjolan.di.bawah.ketiak	1	40.992	62.992
- Keluar.cairan.dari.puting.susu	1	41.354	63.354
- Perubahan.bentuk.dan.ukuran.payudara	1	41.444	63.444
- Kerutan.di.sekitar.payudara	1	41.473	63.473
- Penarikan.puting.susu.ke.dalam	1	42.172	64.172
- Nyeri.di.bagian.tertentu.payudara	1	42.387	64.387
<none>		40.537	64.537

Step: AIC=59.2

Status ~ Keluar.cairan.dari.puting.susu + Cekungan.di.kulit.payudara + Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam + Kerutan.di.sekitar.payudara + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukuran.payudara + Usia

	Df	Deviance	AIC
- Cekungan.di.kulit.payudara	1	41.202	57.202
- Usia	1	41.492	57.492
- Keluar.cairan.dari.puting.susu	1	41.740	57.740
- Benjolan.di.bawah.ketiak	1	41.916	57.916
- Kerutan.di.sekitar.payudara	1	41.985	57.985
- Perubahan.bentuk.dan.ukuran.payudara	1	42.387	58.387
- Nyeri.di.bagian.tertentu.payudara	1	42.725	58.725
- Penarikan.puting.susu.ke.dalam	1	42.806	58.806
<none>		41.200	59.200

Step: AIC=57.2

Status ~ Keluar.cairan.dari.puting.susu + Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam + Kerutan.di.sekitar.payudara + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukuran.payudara + Usia

	Df	Deviance	AIC
- Usia	1	41.550	55.550
- Keluar.cairan.dari.puting.susu	1	41.791	55.791
- Benjolan.di.bawah.ketiak	1	41.978	55.978
- Kerutan.di.sekitar.payudara	1	42.031	56.031
- Perubahan.bentuk.dan.ukuran.payudara	1	42.580	56.580
- Penarikan.puting.susu.ke.dalam	1	42.835	56.835
<none>		41.202	57.202
- Nyeri.di.bagian.tertentu.payudara	1	43.749	57.749

Step: AIC=55.55

Status ~ Keluar.cairan.dari.puting.susu + Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam + Kerutan.di.sekitar.payudara + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukuran.payudara

#### Lampiran 4. Tahapan Seleksi *Backward* (Lanjutan)

Step: AIC=55.55  
 Status ~ Keluar.cairan.dari.puting.susu + Nyeri.di.bagian.tertent  
 u.payudara + Penarikan.puting.susu.ke.dalam + Kerutan.di.sekitar.  
 payudara + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukuran  
 .payudara

	Df	Deviance	AIC
- Keluar.cairan.dari.puting.susu	1	42.138	54.138
- Kerutan.di.sekitar.payudara	1	42.248	54.248
- Benjolan.di.bawah.ketiak	1	42.615	54.615
- Perubahan.bentuk.dan.ukuran.payudara	1	43.069	55.069
<none>		41.550	55.550
- Penarikan.puting.susu.ke.dalam	1	43.748	55.748
- Nyeri.di.bagian.tertentu.payudara	1	44.668	56.668

Step: AIC=54.14  
 Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.sus  
 u.ke.dalam + Kerutan.di.sekitar.payudara + Benjolan.di.bawah.keti  
 ak + Perubahan.bentuk.dan.ukuran.payudara

	Df	Deviance	AIC
- Kerutan.di.sekitar.payudara	1	42.404	52.404
- Benjolan.di.bawah.ketiak	1	43.683	53.683
- Penarikan.puting.susu.ke.dalam	1	43.858	53.858
- Perubahan.bentuk.dan.ukuran.payudara	1	44.093	54.093
<none>		42.138	54.138
- Nyeri.di.bagian.tertentu.payudara	1	45.063	55.063

Step: AIC=52.4  
 Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.sus  
 u.ke.dalam + Benjolan.di.bawah.ketiak + Perubahan.bentuk.dan.ukur  
 an.payudara

	Df	Deviance	AIC
- Benjolan.di.bawah.ketiak	1	44.297	52.297
<none>		42.404	52.404
- Perubahan.bentuk.dan.ukuran.payudara	1	45.079	53.079
- Nyeri.di.bagian.tertentu.payudara	1	45.682	53.682
- Penarikan.puting.susu.ke.dalam	1	45.853	53.853

Step: AIC=52.3  
 Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.sus  
 u.ke.dalam + Perubahan.bentuk.dan.ukuran.payudara

	Df	Deviance	AIC
- Perubahan.bentuk.dan.ukuran.payudara	1	45.220	51.220
- Penarikan.puting.susu.ke.dalam	1	46.151	52.151
<none>		44.297	52.297
- Nyeri.di.bagian.tertentu.payudara	1	47.405	53.405

Step: AIC=51.22  
 Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.sus  
 u.ke.dalam

	Df	Deviance	AIC
<none>		45.220	51.220
- Penarikan.puting.susu.ke.dalam	1	48.555	52.555
- Nyeri.di.bagian.tertentu.payudara	1	49.276	53.276

## Lampiran 5. Tahapan Pemodelan dengan Seleksi *Backward*

```

Call:
glm(formula = Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.pu
ting.susu.ke.dalam, family = binomial(link = "logit"), data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2025 -0.5649 -0.5649  0.5407  1.9566

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)        -1.7547    0.4524  -3.879 0.000105
Nyeri.di.bagian.tertentu.payudara1  1.7900    0.8860   2.020 0.043359
Penarikan.puting.susu.ke.dalam1    1.8135    1.0066   1.802 0.071606
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 56.696  on 48  degrees of freedom
Residual deviance: 45.220  on 46  degrees of freedom
AIC: 51.22

Number of Fisher Scoring iterations: 4

                Hosmer and Lemeshow test (binary model)
data:  data$Status, fitted(bkwrdr_model)
X-squared = 1.0248, df = 1, p-value = 0.3114

```

## Lampiran 6. Tahapan Seleksi *Forward*

Start: AIC=58.7  
Status ~ 1

	Df	Deviance	AIC
+ Nyeri.di.bagian.tertentu.payudara	1	48.555	52.555
+ Penarikan.puting.susu.ke.dalam	1	49.276	53.276
+ Kerutan.di.sekitar.payudara	1	49.276	53.276
+ Cekungan.di.kulit.payudara	1	49.588	53.588
+ Perubahan.bentuk.dan.ukuran.payudara	1	51.057	55.057
+ Benjolan.di.bawah.ketiak	1	53.895	57.895
+ Keluar.cairan.dari.puting.susu	1	54.426	58.426
<none>		56.696	58.696
+ Usia	1	54.934	58.934
+ Durasi.tidur.malam	3	55.225	63.225

Step: AIC=52.56  
Status ~ Nyeri.di.bagian.tertentu.payudara

	Df	Deviance	AIC
+ Penarikan.puting.susu.ke.dalam	1	45.220	51.220
+ Kerutan.di.sekitar.payudara	1	45.220	51.220
+ Perubahan.bentuk.dan.ukuran.payudara	1	46.151	52.151
<none>		48.555	52.555
+ Cekungan.di.kulit.payudara	1	47.647	53.647
+ Usia	1	47.967	53.967
+ Benjolan.di.bawah.ketiak	1	47.989	53.989
+ Keluar.cairan.dari.puting.susu	1	48.060	54.060
+ Durasi.tidur.malam	3	47.280	57.280

Step: AIC=51.22  
Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam

	Df	Deviance	AIC
<none>		45.220	51.220
+ Kerutan.di.sekitar.payudara	1	44.291	52.291
+ Perubahan.bentuk.dan.ukuran.payudara	1	44.297	52.297
+ Usia	1	44.848	52.848
+ Cekungan.di.kulit.payudara	1	44.849	52.849
+ Keluar.cairan.dari.puting.susu	1	45.019	53.019
+ Benjolan.di.bawah.ketiak	1	45.079	53.079
+ Durasi.tidur.malam	3	44.279	56.279

## Lampiran 7. Pemodelan dengan Seleksi *Forward*

```

Call:
glm(formula = Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting
.susu.ke.dalam,
     family = binomial(link = "logit"), data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2025 -0.5649 -0.5649  0.5407  1.9566

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -1.7547     0.4524  -3.879 0.000105
Nyeri.di.bagian.tertentu.payudara1  1.7900     0.8860   2.020 0.043359
Penarikan.puting.susu.ke.dalam1     1.8135     1.0066   1.802 0.071606
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 56.696  on 48  degrees of freedom
Residual deviance: 45.220  on 46  degrees of freedom
AIC: 51.22

Number of Fisher Scoring iterations: 4

      Hosmer and Lemeshow test (binary model)

data: data$Status, fitted(frwr_model)
X-squared = 1.0248, df = 1, p-value = 0.3114

```

## Lampiran 8. Tahapan Seleksi *Stepwise*

Start: AIC=58.7  
Status ~ 1

	Df	Deviance	AIC
+ Nyeri.di.bagian.tertentu.payudara	1	48.555	52.555
+ Penarikan.puting.susu.ke.dalam	1	49.276	53.276
+ Kerutan.di.sekitar.payudara	1	49.276	53.276
+ Cekungan.di.kulit.payudara	1	49.588	53.588
+ Perubahan.bentuk.dan.ukuran.payudara	1	51.057	55.057
+ Benjolan.di.bawah.ketiak	1	53.895	57.895
+ Keluar.cairan.dari.puting.susu	1	54.426	58.426
<none>		56.696	58.696
+ Usia	1	54.934	58.934
+ Durasi.tidur.malam	3	55.225	63.225

Step: AIC=52.56  
Status ~ Nyeri.di.bagian.tertentu.payudara

	Df	Deviance	AIC
+ Penarikan.puting.susu.ke.dalam	1	45.220	51.220
+ Kerutan.di.sekitar.payudara	1	45.220	51.220
+ Perubahan.bentuk.dan.ukuran.payudara	1	46.151	52.151
<none>		48.555	52.555
+ Cekungan.di.kulit.payudara	1	47.647	53.647
+ Usia	1	47.967	53.967
+ Benjolan.di.bawah.ketiak	1	47.989	53.989
+ Keluar.cairan.dari.puting.susu	1	48.060	54.060
+ Durasi.tidur.malam	3	47.280	57.280
- Nyeri.di.bagian.tertentu.payudara	1	56.696	58.696

Step: AIC=51.22  
Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.puting.susu.ke.dalam

	Df	Deviance	AIC
<none>		45.220	51.220
+ Kerutan.di.sekitar.payudara	1	44.291	52.291
+ Perubahan.bentuk.dan.ukuran.payudara	1	44.297	52.297
- Penarikan.puting.susu.ke.dalam	1	48.555	52.555
+ Usia	1	44.848	52.848
+ Cekungan.di.kulit.payudara	1	44.849	52.849
+ Keluar.cairan.dari.puting.susu	1	45.019	53.019
+ Benjolan.di.bawah.ketiak	1	45.079	53.079
- Nyeri.di.bagian.tertentu.payudara	1	49.276	53.276
+ Durasi.tidur.malam	3	44.279	56.279

## Lampiran 9. Pemodelan dengan Seleksi *Stepwise*

```

Call:
glm(formula = Status ~ Nyeri.di.bagian.tertentu.payudara + Penarikan.pu
ting.susu.ke.dalam,
     family = binomial(link = "logit"), data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2025 -0.5649 -0.5649  0.5407  1.9566

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -1.7547    0.4524  -3.879 0.000105
Nyeri.di.bagian.tertentu.payudara1  1.7900    0.8860   2.020 0.043359
Penarikan.puting.susu.ke.dalam1     1.8135    1.0066   1.802 0.071606
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 56.696  on 48  degrees of freedom
Residual deviance: 45.220  on 46  degrees of freedom
AIC: 51.22

Number of Fisher Scoring iterations: 4

                Hosmer and Lemeshow test (binary model)

data: data$Status, fitted(stpwise_model)
X-squared = 1.0248, df = 1, p-value = 0.3114

```

### Lampiran 10. Perhitungan Probabilitas Tiap Responden

Responden	Status	Probabilitas	Prediksi
1	1	0,1475	0
2	1	0,1475	0
3	1	0,1475	0
4	1	0,1475	0
5	1	0,8640	1
6	0	0,1475	0
7	0	0,1475	0
8	1	0,8640	1
9	0	0,5147	1
10	1	0,8640	1
11	1	0,5088	1
12	1	0,1475	0
13	0	0,1475	0
14	1	0,1475	0
15	1	0,5147	1
16	1	0,5088	1
17	0	0,1475	0
18	1	0,8640	1
19	0	0,1475	0
20	0	0,1475	0
21	0	0,1475	0
22	0	0,1475	0
23	0	0,1475	0
24	0	0,5088	1
25	0	0,1475	0
26	0	0,1475	0
27	0	0,1475	0
28	0	0,1475	0
29	0	0,1475	0
30	0	0,1475	0
31	0	0,5088	1
32	0	0,1475	0
33	0	0,5088	1
34	0	0,1475	0
35	0	0,5147	1
36	0	0,1475	0
37	0	0,1475	0
38	0	0,1475	0
39	0	0,1475	0
40	0	0,1475	0
41	0	0,1475	0
42	0	0,1475	0
43	0	0,1475	0
44	0	0,1475	0
45	0	0,1475	0
46	0	0,1475	0
47	0	0,1475	0
48	0	0,1475	0
49	0	0,1475	0

## Lampiran 11. Hasil *Genetic Algorithm*

```

-- Genetic Algorithm -----
GA settings:
Type           = binary
Population size = 80
Number of generations = 50
Elitism        = 5
Crossover probability = 0.8
Mutation probability = 0.1

GA results:
Iterations           = 10
Fitness function value = 0.7799145
Solution =
    Keluar.cairan.dari.puting.susu Cekungan.di.kulit.payudara Ny
eri.di.bagian.tertentu.payudara      1                0
[1,]
1
    Penarikan.puting.susu.ke.dalam Kerutan.di.sekitar.payudara B
enjolan.di.bawah.ketiak              1                0
[1,]
1
    Perubahan.bentuk.dan.ukuran.payudara Durasi.tidur.malam Usia
[1,]                                1                0      1

```

## Lampiran 12. Pemodelan dengan *Genetic Algorithm*

```

Call:
glm(formula = ydata ~ ., family = binomial(link = "logit"), data = newdata)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4991 -0.4889 -0.4889  0.4170  2.0897

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -2.0639     0.5502  -3.751 0.000176
Keluar.cairan.dari.puting.susu1 -0.6660     1.8635  -0.357 0.720793
Nyeri.di.bagian.tertentu.payudara1  1.6679     0.9732  1.714 0.086567
Penarikan.puting.susu.ke.dalam1    2.3533     1.5335  1.535 0.124873
Benjolan.di.bawah.ketiak1         -2.3584     1.9697  -1.197 0.231194
Perubahan.bentuk.ukuran.payudara1  2.2478     1.3995  1.606 0.108237
Usia1                               0.4411     0.9330  0.473 0.636388
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 56.696  on 48  degrees of freedom
Residual deviance: 42.031  on 42  degrees of freedom
AIC: 56.031

Number of Fisher Scoring iterations: 4

```

### Lampiran 13. Pemodelan Kembali dengan *Genetic Algorithm*

```

Call:
glm(formula = Status ~ Nyeri.di.bagian.tertentu.payudara + Perubahan.bentuk
.dan.ukuran.payudara,
     family = binomial(link = "logit"), data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8801  -0.5653  -0.5653   0.6120   1.9558

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -1.7528    0.4560  -3.844 0.000121
Nyeri.di.bagian.tertentu.payudara1  1.8996    0.8658   2.194 0.028239
Perubahan.bentuk.dan.ukuran.payudara1  1.4333    0.9161   1.564 0.117703
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 56.696  on 48  degrees of freedom
Residual deviance: 46.151  on 46  degrees of freedom
AIC: 52.151

Number of Fisher Scoring iterations: 4

```

**Lampiran 14.** Perhitungan Probabilitas pada Model dengan *Genetic Algorithm*

Responden	Status	Probabilitas	Prediksi
1	1	0,420786	0
2	1	0,14769	0
3	1	0,14769	0
4	1	0,14769	0
5	1	0,536629	1
6	0	0,14769	0
7	0	0,14769	0
8	1	0,829214	1
9	0	0,420786	0
10	1	0,829214	1
11	1	0,829214	1
12	1	0,14769	0
13	0	0,14769	0
14	1	0,14769	0
15	1	0,420786	0
16	1	0,536629	1
17	0	0,14769	0
18	1	0,536629	1
19	0	0,420786	0
20	0	0,14769	0
21	0	0,14769	0
22	0	0,14769	0
23	0	0,14769	0
24	0	0,536629	1
25	0	0,14769	0
26	0	0,14769	0
27	0	0,14769	0
28	0	0,14769	0
29	0	0,14769	0
30	0	0,14769	0
31	0	0,536629	1
32	0	0,14769	0
33	0	0,829214	1
34	0	0,14769	0
35	0	0,14769	0
36	0	0,14769	0
37	0	0,14769	0
38	0	0,14769	0
39	0	0,14769	0
40	0	0,14769	0
41	0	0,14769	0
42	0	0,14769	0
43	0	0,14769	0
44	0	0,14769	0
45	0	0,14769	0
46	0	0,14769	0
47	0	0,14769	0
48	0	0,14769	0
49	0	0,14769	0

## Lampiran 15. Contoh Hasil Tes Pemeriksaan Responden

	<b>LABORATORIUM RUMAH SAKIT PANTI RAPIH</b> <b>JL. TEUKU CIK DITIRO 30 - YOGYAKARTA</b> <b>TEL.P.0274-563333, 562233, 514845</b>		Penanggung Jawab Laboratorium: Dr. Tri Djoko Endro Susilo, Sp.PK Dicetak : 10.03.2018 10:43:42
---	--	---	--

---

**JAWABAN PEMERIKSAAN IMMUNOHISTOKIMIA**

NO. PA	: IM 18 - 143	NO.LAB	: 1803020229
NO. BLOK PARAFIN	: PJ 18 - 427	NO.MED.REC	: 1072071
DOKTER PENGIRIM	: Suwardjo, Sp.B(K)Onk	NAMA	: BEYNA HANDAYANI NY
RUMAH SAKIT	: Panti Rapih	UMUR	: 37 tahun
TGL TERIMA	: 05 Maret 2018	JENIS KELAMIN	: P
TGL JAWAB	: 12 Maret 2018	PEKERJAAN	:
DIAGNOSA PA	: Invasive ductal carcinoma	ALAMAT	: BOLU RT 02 RW 16 MARGOKATON SEYEGAN SLEMAN YOGYAKARTA

---

**Jawaban Pemeriksaan :**

ER	: Tidak didapatkan pulasan inti positif pada seluruh sel tumor.
PR	: Tidak didapatkan pulasan inti positif pada seluruh sel tumor.
HER2	: Tidak didapatkan pulasan membran positif pada seluruh sel tumor.
Ki67	: Didapatkan pulasan inti positif pada +/- 70% populasi sel tumor.

**Kesimpulan :** ER : Terpulas negatif/(-).  
PR : Terpulas negatif/(-).  
Her2 : Terpulas negatif/0).  
Ki67 : Index proliferasi 70%.

---

Yogyakarta, 10 Mar 2018  
Dokter Pemeriksa

Kalau ada keraguan hasil,  
Harap segera menghubungi instalasi laboratorium

Dr FX. EDIATI T Sp.PA

---

Pengesahan hasil ini dilakukan secara elektronik

## Lampiran 16. Surat Izin Pengambilan Data



Profesional + Bersahabat + Islami

Nomor : 97/RSNH/KET/DIR/IV/2019  
 Perihal : Balasan Ijin Pengambilan Data Tugas Akhir  
 Lampiran : -

**Kepada Yth.  
 Dinda Galuh Guminta  
 Di tempat**

*Assalamu'alaikum Warohmatullahi Wabarokatuh*

Puji syukur kehadirat Allah SWT atas limpahan rahmat dan karuniaNya. Shalawat dan salam semoga tercurah kepada Nabi Besar Muhammad SAW, keluarga, sahabat dan para pengikutnya. Semoga kemudahan dan kebaikan senantiasa diberikan pada usaha kita. Menindaklanjuti surat dari Institut Teknologi Sepuluh Nopember Fakultas Matematika Komputasi dan Sains Data Departemen Statistika dengan 25164/IT2.VI.9.2/TU.00.09/2019 perihal permohonan Ijin Penelitian:

Nama : Dinda Galuh Guminta  
 No. Mhs : 0621154000093 (1315 100 093)  
 Prodi : Sarjana (S1)  
 Perguruan Tinggi : Institut Teknologi Sepuluh Nopember

Perihal kegiatan : Ijin Pengambilan Data Tugas Akhir  
 Jenis kegiatan : RS Nur Hidayah  
 Lokasi Kegiatan : "Analisis Gejala Klinis Kanker Payudara Menggunakan Regresi Logistik Biner - Genetic Algorithm di RS Nur Hidayah Bantul  
 Mata Kuliah :  
 Waktu pelaksanaan : 06 - 07 April 2019  
 Pembimbing : Dr. Estianna Khoirunnisa (+62 822-2027-5909)

Maka dengan ini kami memberikan Ijin kepada mahasiswa tersebut untuk melakukan Pengambilan Data Tugas Akhir di RS Nur Hidayah dengan ketentuan sebagai berikut:

1. Wajib menjaga tata tertib dan mentaati ketentuan-ketentuan yang berlaku selama Penelitian di lingkungan rumah sakit
2. Wajib memberikan laporan hasil penelitian berupa Hard Copy dan Soft Copy kepada Direktur c/q Penanggungjawab Diklat RS Nur Hidayah Bantul
3. Surat izin ini hanya diperlukan untuk kegiatan ilmiah
4. Surat izin ini dapat dibatalkan sewaktu-waktu apabila tidak dipenuhi ketentuan-ketentuan yang sudah disampaikan

Demikian surat ini kami sampaikan. Atas perhatian dan kerjasamanya kami ucapkan terima kasih.

*Wassalamu'alaikum Warohmatullohi Wabarokatuh*

Bantul, April 2019  
 Menyetujui  
 Direktur RS Nur Hidayah  
  
 dr. Arusi Ferry MPH.  
 RS Nur Hidayah Bantul  
 Rumah Sakit Peduli Masyarakat

- Tembusan:
1. Pembimbing Lapangan/ CI
  2. Bagian Diklat
  3. Yang bersangkutan

## Lampiran 17. Surat Pernyataan Data

### SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Dinda Galuh Guminta

NRP : 0621154000093

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini benar-benar merupakan hasil survey pada responden sebagaimana terlampir.

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data, maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui  
Pembimbing Tugas Akhir



Dr. Santi Wulan Purnami, S.Si., M.Si.  
NIP. 19720923 199803 2 001

Surabaya, 3 Juli 2019



Dinda Galuh Guminta  
NRP. 0621154000093

\*(coret yang tidak perlu)

## Lampiran 18. Surat Keterangan

### SURAT KETERANGAN

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Dinda Galuh Guminta

NRP : 0621154000093

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : RS Nur Hidayah

sejak tanggal 7 April 2019 sampai dengan 28 April 2019 untuk keperluan Tugas Akhir/ Thesis Semester Gasal/Genap\* 2018/ 2019.

2. Tidak Keberatan/Keberatan\* nama perusahaan dicantumkan dalam Tugas Akhir/ Thesis mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak Keberatan/Keberatan\* bahwa hasil analisis data dari perusahaan dipublikasikan dalam E journal ITS yaitu Jurnal Sains dan Seni ITS.

Bantul, 1 Juli 2019  
Direktur RS Nur Hidayah

  
Dr. Arris Ferry, MPH

*Ramah Sakit Peduli Ibadah*

\*(coret yang tidak perlu)

## BIODATA PENULIS



Penulis yang memiliki nama lengkap Dinda Galuh Guminta ini lahir di Surabaya pada tanggal 11 Desember 1996. Penulis telah menempuh pendidikan formal di SDN Rangkah VII Surabaya (2003-2009), SMPN 1 Surabaya (2009-2012), dan SMAN 5 Surabaya (2012-2015). Kemudian, penulis melanjutkan pendidikan Program Studi Sarjana di Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data (FMKSD), Institut Teknologi Sepuluh Nopember Surabaya melalui jalur SBMPTN. Selama masa perkuliahan, penulis aktif dalam organisasi Himpunan Mahasiswa Statistika (HIMASTA) ITS sebagai staff departemen dalam negeri pada periode 2017/2018. Penulis turut berpartisipasi dalam beberapa kepanitiaan seperti Ini Lho ITS, Kesma Expo, Pekan Raya Statistika (PRS), GERIGI, dan ITS Expo. Selain itu, penulis berkesempatan melakukan kerja praktik di PT Pertamina (Persero) MOR V. Apabila pembaca ingin berdiskusi lebih lanjut, penulis dapat dihubungi melalui email [dindagaluhg@gmail.com](mailto:dindagaluhg@gmail.com).

*(Halaman ini sengaja dikosongkan)*