



TUGAS AKHIR - SS 141501

**KLASIFIKASI RISIKO INFEKSI PADA BAYI BARU LAHIR  
DI RUMAH SAKIT UMUM DAERAH SIDOARJO  
MENGUNAKAN METODE *CLASSIFICATION TREES***

**AULIA RAHMA SAFITRI**  
NRP 1312 100 066

Dosen Pembimbing  
Ir. Sri Pingit Wulandari, M.Si.

Program Studi S1 Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember  
Surabaya 2016



**TUGAS AKHIR - SS 141501**

**KLASIFIKASI RISIKO INFEKSI PADA BAYI BARU LAHIR  
DI RUMAH SAKIT UMUM DAERAH SIDOARJO  
MENGUNAKAN METODE *CLASSIFICATION TREES***

**AULIA RAHMA SAFITRI**  
NRP 1312100 066

Dosen Pembimbing  
Ir. Sri Pingit Wulandari, M.Si.

Program Studi S1 Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember  
Surabaya 2016



**FINAL PROJECT - SS 141501**

**CLASSIFICATION OF INFECTION RISK OF NEWBORN  
BABIES IN SIDOARJO REGIONAL PUBLIC HOSPITAL  
USING CLASSIFICATION TREES METHOD**

**AULIA RAHMA SAFITRI**  
NRP 1312 100 066

Supervisor  
Ir. Sri Pingit Wulandari, M.Si.

Undergraduate Programme of Statistics  
Faculty of Mathematics and Natural Sciences  
Institut Teknologi Sepuluh Nopember  
Surabaya 2016

**LEMBAR PENGESAHAN**

**KLASIFIKASI RISIKO INFEKSI PADA BAYI BARU  
LAHIR DI RSUD SIDOARJO MENGGUNAKAN METODE  
CLASSIFICATION TREES**

**TUGAS AKHIR**

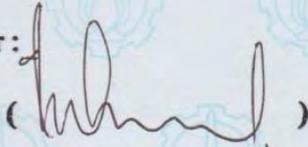
Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada  
Program Studi S-1 Jurusan Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember

Oleh :

**AULIA RAHMA SAFITRI  
NRP 1312 100 066**

Disetujui oleh Pembimbing Tugas Akhir :

**Ir. Sri Pingit Wulandari, M.Si.**  
**NIP. 19620603 198701 2 001**



Mengetahui  
**Ketua Jurusan Statistika FMIPA-ITS**

**Dr. Suhartono**

**NIP. 19710929 199512 1 001**

**SURABAYA, JANUARI 2016**



# KLASIFIKASI RISIKO INFEKSI PADA BAYI BARU LAHIR DI RUMAH SAKIT UMUM DAERAH SIDOARJO MENGGUNAKAN METODE *CLASSIFICATION TREES*

Nama Mahasiswa : Aulia Rahma Safitri  
NRP : 1312 100 066  
Jurusan : Statistika FMIPA-ITS  
Dosen Pembimbing : Ir. Sri Pingit Wulandari, M.Si.

## Abstrak

Angka kematian bayi (AKB) merupakan salah satu indikator derajat kesehatan masyarakat yang dapat mengukur tingkat kesejahteraan masyarakat ataupun tingkat kemiskinan di Indonesia. AKB pada tahun 2012 sebesar 32 per 1.000 kelahiran hidup dan proporsi kematian neonatal (bayi baru lahir) terhadap kematian bayi meningkat menjadi 59,4%. Infeksi pada bayi baru lahir gejalanya kurang jelas dan seringkali tidak diketahui sampai keadaannya sudah sangat terlambat. Karakteristik bayi baru lahir perlu diketahui untuk membantu menetapkan status risiko infeksi, sehingga bayi dengan status positif risiko infeksi dapat diberikan perawatan intensif dan beberapa antibiotik untuk melawan bakteri atau kuman yang menyebabkan infeksi. Penelitian ini berguna untuk mengetahui hasil ketepatan klasifikasi yang sudah dilakukan oleh RSUD Sidoarjo dan juga untuk mengetahui variabel terpenting yang berguna untuk pengklasifikasian. Berdasarkan hasil analisis *Classification Trees* setelah dilakukan *pre-processing* dengan menghapus data yang mengandung *missing* diperoleh akurasi untuk data testing dengan aturan pemilihan pemilah indeks gini dan 10-fold cross validation estimate untuk menentukan pohon optimal sebesar 93,5%, dengan persentase *sensitivity* dan *specificity* masing-masing sebesar 93,1% dan 93,9%. Faktor terpenting dalam klasifikasi risiko infeksi pada bayi baru lahir di RSUD Sidoarjo adalah kondisi sisa air ketuban.

**Kata Kunci :** Akurasi, Angka Kematian Bayi, *Classification Trees*, Indeks Gini, Risiko Infeksi



*“Halaman ini sengaja dikosongkan”*

# CLASSIFICATION OF INFECTION RISK OF NEWBORN BABIES IN SIDOARJO REGIONAL PUBLIC HOSPITAL USING CLASSIFICATION TREES METHOD

**Name** : Aulia Rahma Safitri  
**NRP** : 1312 100 066  
**Department** : Statistika FMIPA-ITS  
**Supervisor** : Ir. Sri Pingit Wulandari, M.Si.

## Abstract

*The infant mortality rate (IMR) is one indicator of the degree of public health that can measure the level of welfare or poverty in Indonesia. IMR in 2012 amounts to 32 per 1.000 live births and the proportion of neonatal deaths (newborn) to infant mortality increased to 59.4%. Symptoms of infection in newborn are less obvious and often unknown until the situation is already very late. Characteristics of newborn need to know to help define the status of the infection risk, so newborn with a positive status of infection risk can be given intensive care and some antibiotics to fight bacteria or germs that cause infection. This research is useful to know the accuracy of the classification has been done by the Sidoarjo Regional Public Hospital and also to determine the most important variable that are useful for the classification. Based on the analysis Classification Trees, after pre-processing to remove the data that contain missing obtained accuracy for testing data using the method of selecting a classifier, gini index and 10-fold cross validation estimate to determine optimum trees is 93.5% with the percentage of sensitivity and specificity respectively by 93.1% and 93.9%. The most important factor in the classification of infection risk of newborns babies in Sidoarjo Regional Public Hospital is a condition of the rest of the amniotic fluid.*

**Keywords** : Accuracy, Classification Trees, Gini Index, Infant Mortality Rate, Infection Risk



## KATA PENGANTAR

Puji syukur kehadirat Allah SWT, atas segala rahmat dan karunia-Nya yang tak pernah henti diberikan, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul

**“Klasifikasi Risiko Infeksi Pada Bayi Baru Lahir di Rumah Sakit Umum Daerah Sidoarjo Menggunakan Metode *Classification Trees*”**

dengan baik dan tepat pada waktunya. Penyusunan Tugas Akhir ini tak luput dari bantuan serta dukungan dari berbagai pihak. Pada kesempatan ini, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Ir. Sri Pingit Wulandari, M.Si selaku dosen pembimbing yang telah sabar dan meluangkan waktunya untuk membimbing dan memberikan masukan dalam penyelesaian Tugas Akhir ini.
2. Bapak Dr. Suhartono selaku Ketua Jurusan Statistika yang telah menyediakan fasilitas guna kelancaran pengerjaan Tugas Akhir ini.
3. Bapak Dr. Sutikno, M.Si dan Dr. Bambang W. Otok, S.Si, M.Si selaku dosen penguji atas segala kritikan dan saran yang sangat membangun.
4. Ibu Dra. Lucia Aridinanti, M.T selaku Kaprodi S1 Jurusan Statistika ITS.
5. Ibu Erma Oktania Permatasari, S.Si., M.Si selaku dosen wali atas segala nasehat dan bimbingan yang berguna selama pada saat perwalian.
6. Bapak drg. Syaf Satriawarman, Sp. Pros sebagai Wakil Perencanaan dan Pendidikan Rumah Sakit Umum Daerah Sidoarjo yang telah memberikan ijin kepada penulis untuk melakukan penelitian mengenai risiko infeksi pada bayi baru lahir di RSUD Sidoarjo.
7. Ayahanda tercinta Sjaiful Muttaqien dan Ibunda tersayang Nur Fitri atas doa, nasehat, dan kasih sayang yang sangat

besar yang telah diberikan untuk penulis sehingga dapat menjadi motivasi dan penyemangat bagi penulis disaat menghadapi kesulitan.

8. Kakak dan adikku, Affan Safrian dan Alif Haidar Safrian yang selalu memberikan hiburan tersendiri kepada penulis agar tetap semangat dalam mengerjakan tugas akhir ini.
9. Sahabat penulis yang tersayang, Puspita Permatasari dan Yollafie Asmara yang selalu ada di saat suka dan duka serta memberikan dukungan secara penuh kepada penulis selama menjalani kuliah 3,5 tahun di Statistika ITS.
10. Teman satu atap penulis, yakni kontrakan “Barakallah” yang sudah seperti keluarga yang selalu kompak. Noorahma, Nova, Winda, Devi, Faras, Firda, Rukmi, Vonesa, Dini, Eno, Fauzah dan Yolla merupakan duabelas orang yang selalu memberikan dukungan, semangat dan menghibur penulis untuk melewati permasalahan selama hidup di Surabaya.
11. Teman-teman seperjuangan PW 113, serta teman-teman Statistika ITS  $\Sigma 23$  yang memberikan hangatnya sebuah kebersamaan dan rasa kekeluargaan selama 3,5 tahun ini.
12. Pengurus HIMASTA-ITS periode 2013/2014 dan 2014/2015 yang telah memberikan dukungan kepada penulis untuk selalu berkarya.
13. Serta semua pihak yang telah memberikan dukungan baik moril maupun materiil yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari masih banyaknya kekurangan dalam pembuatan laporan Tugas Akhir ini, besar harapan bagi penulis untuk dapat menerima saran dan kritik yang bersifat membangun guna perbaikan di masa mendatang. Semoga laporan Tugas Akhir ini dapat bermanfaat bagi penelitian selanjutnya.

Surabaya, Januari 2016

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> .....	xvii
<b>DAFTAR GAMBAR</b> .....	xix
<b>DAFTAR LAMPIRAN</b> .....	xxi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Statistik Deskriptif.....	7
2.2 <i>Cross Tabulation</i> .....	7
2.3 <i>Missing Value</i> .....	8
2.4 <i>Classification and Regression Trees (CART)</i> .....	9
2.4.1 Pembentukan Pohon Klasifikasi Maksimal.....	12
2.4.1.1 Pemilihan Pemilah.....	12
2.4.1.2 Penentuan Simpul Terminal.....	14
2.4.1.3 Penandaan Label Kelas.....	15
2.4.2 Pemangkasan Pohon Klasifikasi.....	15
2.4.3 Penentuan Pohon Klasifikasi Optimal.....	16
2.4.4 Ukuran Ketepatan Klasifikasi.....	18
2.4.5 <i>Odds Ratio</i> .....	20
2.4.6 Ilustrasi CART.....	20
2.5 Periode Perinatal dan Neonatal Dini.....	29
2.6 <i>Infection Specific to the Perinatal Period</i> .....	30

2.7 Perawatan Neonatal Esensial Pada Saat Lahir.....	30
--	----

### **BAB III METODE PENELITIAN**

3.1 Sumber Data .....	33
3.2 Variabel Penelitian .....	33
3.3 Langkah Analisis Penelitian.....	36
3.4 Diagram Alir Penelitian.....	38

### **BAB IV ANALISIS DAN PEMBAHASAN**

4.1 Karakteristik Bayi Baru Lahir di Rumah Sakit Umum Daerah Sidoarjo .....	41
4.1.1 Karakteristik Bayi Baru Lahir Berdasarkan Berat Badan Lahir .....	43
4.1.2 Karakteristik Bayi Baru Lahir Berdasarkan Status Lahir Prematur.....	45
4.1.3 Karakteristik Bayi Baru Lahir Berdasarkan Kondisi Air Ketuban .....	46
4.1.4 Karakteristik Bayi Baru Lahir Berdasarkan Status Ketuban Pecah Dini.....	47
4.1.5 Karakteristik Bayi Baru Lahir Berdasarkan Suhu Tubuh Bayi.....	47
4.1.6 Karakteristik Bayi Baru Lahir Berdasarkan Warna Kulit.....	48
4.1.7 Karakteristik Bayi Baru Lahir Berdasarkan Denyut Jantung.....	49
4.1.8 Karakteristik Bayi Baru Lahir Berdasarkan Refleks Rangsangan.....	50
4.1.9 Karakteristik Bayi Baru Lahir Berdasarkan Kecepatan Bernafas.....	51
4.1.10 Karakteristik Bayi Baru Lahir Berdasarkan Kondisi Tali Puser.....	52
4.2 Analisis <i>Classification Trees</i> untuk Klasifikasi Risiko Infeksi pada Bayi Baru Lahir di RSUD Sidoarjo .....	52
4.2.1 Kemungkinan Pemilah untuk Setiap Variabel Prediktor .....	53
4.2.2 Pembentukan Pohon Klasifikasi Maksimal.....	54
4.2.3 Pemangkasan Pohon Klasifikasi Maksimal.....	56

4.2.4	Penentuan Pohon Klasifikasi Optimal.....	58
4.2.5	Hasil Ketepatan Klasifikasi <i>Classification Trees</i> .....	64
4.2.6	<i>Odds Ratio</i> Bayi Baru Lahir Positif Risiko Infeksi ..	66
4.2.7	Ilustrasi Klasifikasi Risiko Infeksi .....	68
<b>BAB V KESIMPULAN DAN SARAN</b>		
5.1	Kesimpulan.....	73
5.2	Saran.....	74
<b>DAFTAR PUSTAKA</b> .....		75
<b>LAMPIRAN</b> .....		77
<b>BIODATA PENULIS</b> .....		113



*“Halaman ini sengaja dikosongkan”*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran A</b> Data Pengamatan Bayi Baru Lahir di RSUD Sidoarjo .....	77
<b>Lampiran B</b> Kemungkinan Pemilah pada Variabel Prediktor dengan Skala Kontinyu .....	79
<b>Lampiran C</b> Statistik Deskriptif dan Tabulasi Silang antara Variabel Respon dan Prediktor .....	80
<b>Lampiran D</b> Output Pohon Klasifikasi <i>10-Fold Cross Validation Estimate</i> .....	85
<b>Lampiran E</b> Output Kesalahan Klasifikasi .....	91
<b>Lampiran F</b> Hasil Ketepatan Klasifikasi ( <i>Accuracy</i> ) .....	92
<b>Lampiran G</b> Informasi Simpul Terminal Pohon Klasifikasi .....	95
<b>Lampiran H</b> Pohon Klasifikasi Maksimal .....	99
<b>Lampiran I</b> Pohon Klasifikasi Optimal .....	100
<b>Lampiran J</b> Ilustrasi Perhitungan Skor Variabel Terpenting ..	101
<b>Lampiran K</b> Ilustrasi <i>10-fold Cross Validation Estimate</i> .....	102
<b>Lampiran L</b> Hasil Klasifikasi Data Rekam Medis Bayi Baru Lahir .....	109
<b>Lampiran M</b> Surat Keterangan Penelitian di RSUD Sidoarjo .	111



*“Halaman ini sengaja dikosongkan”*

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b>	Ilustrasi Struktur Pohon Klasifikasi..... 11
<b>Gambar 2.2</b>	Ilustrasi Prosedur <i>10-fold Cross Validation</i> ..... 18
<b>Gambar 2.3</b>	Pohon Klasifikasi Ilustrasi Hasil Pemilahan..... 27
<b>Gambar 3.1</b>	Diagram Alir Penelitian..... 39
<b>Gambar 4.1</b>	Deskripsi Status Risiko Infeksi pada Bayi..... 42
<b>Gambar 4.2</b>	Deskripsi Bayi Baru Lahir Berdasarkan Jenis Kelamin..... 42
<b>Gambar 4.3</b>	Deskripsi BBL Berdasarkan Status Risiko Infeksi ..... 44
<b>Gambar 4.4</b>	Deskripsi Status Lahir Prematur Berdasarkan Status Risiko Infeksi ..... 45
<b>Gambar 4.5</b>	Deskripsi Kondisi Air Ketuban Berdasarkan Status Risiko Infeksi ..... 46
<b>Gambar 4.6</b>	Deskripsi Status KPD Berdasarkan Status Risiko Infeksi..... 47
<b>Gambar 4.7</b>	Deskripsi Warna Kulit Berdasarkan Status Risiko Infeksi..... 49
<b>Gambar 4.8</b>	Deskripsi Refleks Rangsangan Berdasarkan Status Risiko Infeksi ..... 50
<b>Gambar 4.9</b>	Deskripsi Kecepatan Bernafas Berdasarkan Status Risiko Infeksi ..... 51
<b>Gambar 4.10</b>	Deskripsi Kondisi Tali Puser Berdasarkan Status Risiko Infeksi ..... 52
<b>Gambar 4.11</b>	Topologi Pohon Klasifikasi Maksimal ..... 56
<b>Gambar 4.12</b>	Nilai <i>Relative Cost</i> Hasil Pemangkasian Pohon Klasifikasi Maksimal ..... 57
<b>Gambar 4.13</b>	Topologi Pohon Klasifikasi Optimal ..... 60



*“Halaman ini sengaja dikosongkan”*

## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> Tabel Kontingensi $rx_c$ .....	8
<b>Tabel 2.2</b> Tabulasi Silang Ketepatan Klasifikasi .....	19
<b>Tabel 2.3</b> Variabel Penelitian Ilustrasi CART .....	20
<b>Tabel 2.4</b> Data Sampel Ilustrasi CART .....	21
<b>Tabel 2.5</b> Proporsi Pengamatan Variabel $Y$ .....	22
<b>Tabel 2.6</b> Perhitungan <i>Goodness of Split</i> untuk <i>Threshold</i> $X_1=1400$ .....	23
<b>Tabel 2.7</b> Perhitungan <i>Goodness of Split</i> untuk <i>Threshold</i> sebesar $X_1=2450$ .....	24
<b>Tabel 2.8</b> Perhitungan <i>Goodness of Split</i> untuk <i>Threshold</i> sebesar $X_1=2950$ .....	24
<b>Tabel 2.9</b> Perhitungan <i>Goodness of Split</i> untuk <i>Threshold</i> sebesar $X_1=3400$ .....	25
<b>Tabel 2.10</b> Perhitungan <i>Goodness of Split</i> untuk <i>Threshold</i> sebesar $X_2$ .....	26
<b>Tabel 2.11</b> Perbandingan <i>Goodness of Split</i> dari Pemilah Awal .....	26
<b>Tabel 2.12</b> Penandaan Label Kelas pada Simpul Terminal .....	28
<b>Tabel 2.13</b> Data Ilustrasi Bayi Baru Lahir .....	29
<b>Tabel 2.14</b> Hasil Klasifikasi Ilustrasi Bayi Baru Lahir .....	29
<b>Tabel 3.1</b> Struktur Data .....	33
<b>Tabel 3.2</b> Variabel Penelitian .....	34
<b>Tabel 4.1</b> Deskripsi Jenis Kelamin Berdasarkan Status Risiko Infeksi .....	43
<b>Tabel 4.2</b> Deskripsi Bayi Baru Lahir Berdasarkan BBL .....	43
<b>Tabel 4.3</b> Deskripsi Bayi Baru Lahir Berdasarkan Suhu Tubuh .....	48
<b>Tabel 4.4</b> Deskripsi Bayi Baru Lahir Berdasarkan Denyut Jantung .....	49
<b>Tabel 4.5</b> Kemungkinan Pemilah Setiap Variabel Prediktor .....	54
<b>Tabel 4.6</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal .....	55
<b>Tabel 4.7</b> Urutan Pohon Klasifikasi .....	58

<b>Tabel 4.8</b>	Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal .....	59
<b>Tabel 4.9</b>	Pelabelan Kelas Simpul Terminal .....	61
<b>Tabel 4.10</b>	Karakteristik Bayi Baru Lahir Berdasarkan Simpul Terminal .....	64
<b>Tabel 4.11</b>	Evaluasi Performansi Hasil Klasifikasi pada Data <i>Learning</i> .....	65
<b>Tabel 4.12</b>	Ketepatan Klasifikasi Risiko Infeksi pada Bayi Baru Lahir untuk Data <i>Testing</i> .....	66
<b>Tabel 4.13</b>	Ilustrasi Data Rekam Medis Bayi Baru Lahir .....	68
<b>Tabel 4.14</b>	Hasil Klasifikasi Ilustrasi Bayi Baru Lahir .....	70
<b>Tabel 4.15</b>	Tabulasi Silang Hasil Ketepatan Klasifikasi Bayi Baru Lahir .....	71

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Angka kematian bayi (AKB) merupakan salah satu indikator derajat kesehatan masyarakat yang termasuk dalam salah satu tujuan Renstra (Rencana Strategis) Kementerian Kesehatan Republik Indonesia pada tahun 2015-2019, yaitu menurunkan AKB dari 32 menjadi 24 per 1.000 kelahiran hidup pada tahun 2019. Menurut Survei Demografi dan Kesehatan Indonesia (2012) AKB berhasil diturunkan secara tajam dari 68 per 1.000 kelahiran hidup pada tahun 1990an menjadi 32 per 1.000 kelahiran hidup pada tahun 2012. Kontribusi terbesar terhadap kematian bayi diberikan oleh kematian neonatal. Kematian neonatal adalah peluang bayi meninggal dalam bulan pertama setelah lahir. Angka kematian neonatal (AKN) tetap sejak tahun 2003 yaitu sebesar 19 per 1.000 kelahiran hidup. Proporsi kematian neonatal terhadap kematian bayi meningkat dari 47,1 persen pada tahun 2009 menjadi 59,4 persen pada tahun 2012. Berdasarkan hasil Susenas Jawa Timur pada tahun 2013 menunjukkan bahwa AKB Kabupaten Sidoarjo sebesar 22,11 per 1.000 kelahiran hidup (BPS, 2014). Menurut data Dinas Kesehatan Kabupaten Sidoarjo penyebab utama kematian bayi adalah bayi berat lahir rendah, prematuritas dan infeksi. Penyebab kematian neonatal usia 0-6 hari adalah gangguan pernapasan, prematuritas, infeksi dan hipotermia. Infeksi dan hipotermia dapat diakibatkan oleh berat badan lahir rendah dan sebagian prematuritas memiliki berat lahir yang rendah (Depkes, 2008).

Masalah utama bayi baru lahir pada periode neonatal dini (0-6 hari) dapat mengakibatkan kematian, kesakitan dan kecacatan. Kasus infeksi yang dialami pada bayi baru lahir merupakan salah satunya. Umumnya, kasus infeksi pada bayi baru lahir menunjukkan gejala yang kurang jelas dan seringkali tidak diketahui sampai keadaannya sudah sangat terlambat. Oleh sebab itu, tenaga kesehatan perlu melakukan pencegahan infeksi melalui

perawatan intensif terhadap bayi baru lahir yang memiliki risiko infeksi. Risiko infeksi sendiri tidak bisa dihilangkan secara total, tetapi dapat dikurangi hingga sekecil mungkin dengan menerapkan tindakan-tindakan pencegahan infeksi yang benar dan konsisten. Diagnosa yang tepat dari tenaga kesehatan terhadap bayi baru lahir yang berisiko infeksi merupakan hal yang sangat penting. Melalui diagnosa tersebut bayi baru lahir yang memiliki risiko infeksi dapat diberikan perawatan dan beberapa antibiotik untuk melawan bakteri atau kuman yang menyebabkan infeksi. Karakteristik bayi baru lahir perlu diketahui melalui pemeriksaan esensial yang dilakukan tenaga kesehatan guna menetapkan diagnosa terhadap bayi tersebut apakah berisiko infeksi atau tidak. Hal tersebut dapat diwujudkan dengan mengklasifikasi bayi baru lahir berisiko infeksi dan tidak berisiko infeksi. Metode pengklasifikasian yang tepat yaitu metode *Classification Trees*. Metode *Classification Trees* merupakan bagian dari metode CART (*Classification and Regression Trees*). Metode CART adalah metode nonparametrik yang dapat memilih variabel dan interaksi yang paling mempengaruhi variabel respon. Jika variabel respon yang dipelajari kontinu, maka CART akan menghasilkan pohon regresi, namun jika variabel respon yang dipelajari memiliki skala kategorik, maka CART akan menghasilkan pohon klasifikasi (Breiman, Friedman, Olshen, & Stone, 1993).

Penelitian sebelumnya mengenai risiko infeksi pada bayi baru lahir dilakukan oleh Rucy (2014) yang meneliti tentang faktor-faktor yang mempengaruhi risiko infeksi perinatal menggunakan Regresi Logistik Biner. Hasil penelitian tersebut menyimpulkan bahwa faktor-faktor yang mempengaruhi risiko infeksi perinatal yaitu kondisi air ketuban dan ketuban pecah dini. Sinseng (2008) dalam penelitiannya menghasilkan bahwa nilai APGAR pada persalinan dengan riwayat ketuban pecah dini lebih rendah daripada nilai APGAR pada persalinan normal. Hal tersebut disebabkan oleh selaput ketuban yang pecah sebelum terjadi persalinan dapat menjadi jalan masuk bagi kuman,

sehingga meningkatkan potensi infeksi. Sedangkan penelitian penggunaan metode CART yang dilakukan oleh Khasanah (2015) mengenai klasifikasi penyakit Retardasi Mental (RM) menghasilkan akurasi penderita RM menggunakan pemilah indeks gini berdasarkan pohon klasifikasi optimal dengan variabel terpenting gejala psikiatrik sebesar 72,7% dan simpul terminal dihasilkan sebanyak 15 simpul. Akurasi penderita RM menggunakan pemilah indeks twoing pada pohon klasifikasi optimal dengan variabel terpenting adalah gejala psikiatrik sebesar 71,4% dengan 11 simpul terminal. Penelitian juga dilakukan oleh Margasari (2014) yang menerapkan metode CART dan Regresi Logistik Biner pada klasifikasi profil mahasiswa FMIPA Universitas Brawijaya. Dalam penelitian tersebut menghasilkan ketepatan klasifikasi data *testing* pada pohon optimal sebesar 94,2%, sementara pada analisis regresi logistik biner nilai ketepatan klasifikasi yang dihasilkan sebesar 86,7%. Oleh sebab itu metode CART lebih baik dalam memprediksi variabel respon dilihat dari besarnya hasil klasifikasi dalam kasus masa studi mahasiswa FMIPA Universitas Brawijaya tahun lulus 2011-2013. Menelaah dari penelitian sebelumnya terkait risiko infeksi pada bayi baru lahir dan metode CART, maka dalam penelitian ini menggunakan metode CART, khususnya *Classification Trees* untuk mendapatkan klasifikasi risiko infeksi pada bayi baru lahir.

Studi kasus yang digunakan dalam penelitian ini merupakan studi kasus dari Rumah Sakit Umum Daerah (RSUD) Sidoarjo. Karena RSUD Sidoarjo merupakan rumah sakit yang menjadi tujuan utama bagi ibu hamil untuk melakukan persalinan di Kabupaten Sidoarjo. Kasus infeksi sendiri merupakan salah satu faktor penyebab tingginya angka kematian bayi di Sidoarjo. Berdasarkan data rekam medis RSUD Sidoarjo pada minggu awal Januari 2015, jumlah bayi baru lahir yang memiliki risiko infeksi (*infection specific to the perinatal period unspecified*) di RSUD Sidoarjo menunjukkan angka yang cukup tinggi, yaitu sebanyak 17 bayi positif risiko infeksi dari total 35 bayi yang lahir. Oleh

sebab itu, untuk menurunkan kasus infeksi di Kabupaten Sidoarjo perlu diketahui faktor yang paling berpengaruh terhadap kasus infeksi pada bayi tersebut. Selain itu, dilakukan pula pengklasifikasian berisiko dan tidak berisiko infeksi pada bayi baru lahir menggunakan metode *Classification Trees* untuk mengetahui perbedaan karakteristik diantara dua kategori bayi tersebut berdasarkan kriteria kondisi bayi saat lahir. Tujuan utama *Classification Trees* adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian (Lewis, 2000).

## 1.2 Rumusan Masalah

Menurunkan Angka Kematian Bayi (AKB) merupakan salah satu tujuan pembangunan kesehatan Indonesia untuk meningkatkan derajat kesehatan dan kualitas sumber daya manusia. Sejak tahun 2003 angka kematian neonatal tetap yaitu sebesar 19 per 1.000 kelahiran hidup dengan proporsi terhadap angka kematian bayi sebesar 59,4% pada tahun 2012. Penyebab kematian neonatal salah satunya adalah karena infeksi berat. Oleh sebab itu, perlu adanya penanganan lebih lanjut untuk menurunkan kasus infeksi pada bayi baru lahir melalui pelayanan kesehatan pada bayi baru lahir dan tindakan medis yang tepat pada bayi berisiko infeksi. Upaya yang dapat dilakukan untuk menurunkan kasus infeksi yaitu dengan mengetahui faktor paling dominan yang menyebabkan kasus infeksi pada bayi serta melakukan pengklasifikasian bayi baru lahir berdasarkan status risiko infeksi untuk mengetahui perbedaan karakteristik bayi dari kedua kategori status risiko infeksi tersebut. Hal tersebut dapat diketahui menggunakan metode *Classification Trees* yang menghasilkan pohon klasifikasi. Pohon klasifikasi tersebut memilah bayi baru lahir yang berisiko dan tidak berisiko infeksi berdasarkan faktor-faktor yang mempengaruhi. Berdasarkan uraian tersebut, permasalahan yang dibahas dalam penelitian ini adalah mengenai karakteristik pada bayi baru lahir di RSUD Sidoarjo yang berisiko dan tidak berisiko infeksi.

### **1.3 Tujuan Penelitian**

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendapatkan faktor yang paling berpengaruh terhadap penentuan status risiko infeksi pada bayi baru lahir berdasarkan hasil analisis *Classification Trees*.
2. Menghitung akurasi hasil klasifikasi risiko infeksi pada bayi baru lahir di Rumah Sakit Umum Daerah Sidoarjo menggunakan metode *Classification Trees* berdasarkan faktor-faktor yang mempengaruhi.
3. Membandingkan karakteristik bayi baru lahir di Rumah Sakit Umum Daerah Sidoarjo yang berisiko infeksi dan bayi yang tidak memiliki risiko infeksi dengan kriteria kondisi bayi saat baru lahir berdasarkan hasil analisis *Classification Trees*.

### **1.4 Manfaat Penelitian**

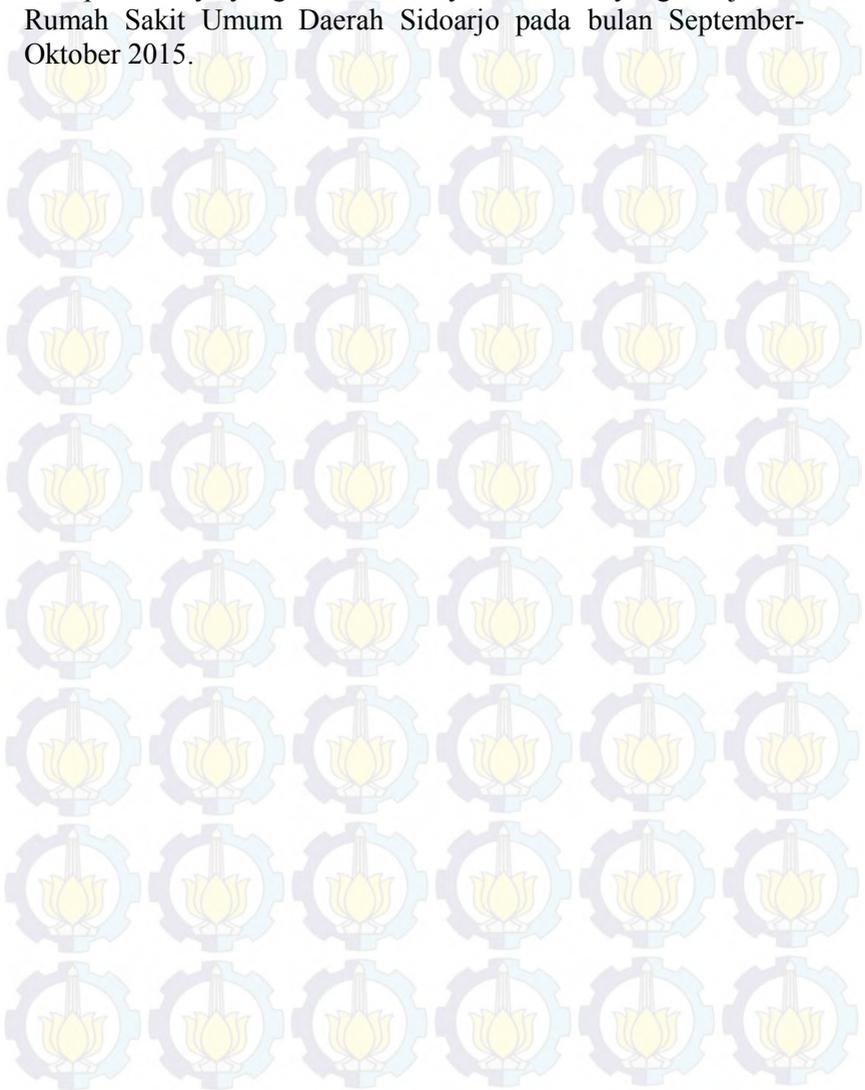
Adapun manfaat yang diharapkan dari pelaksanaan penelitian ini adalah sebagai berikut.

1. Bagi bidang kesehatan, penelitian ini dapat memberikan informasi mengenai karakteristik bayi baru lahir yang memiliki risiko infeksi atau tidak menggunakan metode *Classification Trees*. Kemudian ketepatan klasifikasi yang dihasilkan dapat dijadikan informasi tambahan bagi Dinas Kesehatan Kabupaten Sidoarjo dan Rumah Sakit Umum Daerah Sidoarjo dalam membuat kebijakan pelayanan kesehatan dan tindakan medis bagi bayi baru lahir berisiko infeksi guna menurunkan kasus infeksi bayi baru lahir.
2. Bagi bidang pendidikan, penelitian ini diharapkan dapat berguna untuk menambah wawasan dan pengetahuan pada penelitian selanjutnya tentang klasifikasi dalam bidang kesehatan.

### **1.5 Batasan Masalah**

Kategori yang digunakan dalam penelitian ini hanya ada 2 kategori sebagai variabel respon yakni bayi positif berisiko infeksi dan negatif berisiko infeksi. Selain itu, batasan masalah

yang digunakan dalam penelitian ini adalah pasien yang diteliti merupakan bayi yang lahir dan bayi baru lahir yang dirujuk di Rumah Sakit Umum Daerah Sidoarjo pada bulan September-Oktober 2015.



## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Statistika Deskriptif**

Statistika deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna, namun teknik statistika ini sama sekali tidak menarik kesimpulan dari data yang akan diteliti (Walpole, 1995). Statistika deskriptif lebih berkenaan dengan pengumpulan dan peringkasan data, serta penyajian hasil ringkasan tersebut. Data-data yang bisa diperoleh merupakan hasil sensus, survei, jajak pendapat atau pengamatan lainnya yang secara umum masih bersifat acak dan belum terorganisir dengan baik, atau biasa disebut *raw data*. Data-data tersebut harus diringkas dengan baik dan teratur, baik dalam bentuk tabel atau presentasi grafis agar lebih mudah dipahami oleh pengguna data. Contoh statistika deskriptif yang sering digunakan yaitu penyajian dalam bentuk tabel, diagram, dan grafik.

#### **2.2 Cross Tabulation**

Tabel kontingensi atau yang sering disebut tabulasi silang (*cross tabulation* atau *cross classification*) adalah tabel yang berisi data jumlah atau frekuensi atau beberapa kategori. *Cross tabulation* yaitu suatu metode statistik yang menggambarkan dua atau lebih variabel secara simultan dan hasil yang ditampilkan berupa tabel yang merefleksikan distribusi bersama dua atau lebih variabel dengan jumlah kategori yang terbatas (Agresti, 2007). Tabel kontingensi dapat digunakan untuk mengetahui asosiasi antara dua atau lebih variabel tetapi bukan hubungan sebab akibat. Semakin bertambah jumlah variabel yang ditabulasikan, maka semakin kompleks interpretasinya. Tabel kontingensi akan ditampilkan pada Tabel 2.1 berikut.

**Tabel 2.1** Tabulasi Silang  $rx c$ 

Baris	Kolom				Total
	1	2	...	c	
1	$n_{11}$	$n_{12}$	...	$n_{1c}$	$n_{1.}$
2	$n_{21}$	$n_{22}$	...	$n_{2c}$	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r	$n_{r1}$	$n_{r2}$	...	$n_{rc}$	$n_{r.}$
<b>Total</b>	$n_{.1}$	$n_{.2}$	...	$n_{.c}$	$n_{..}$

keterangan:

$n_{rc}$  : frekuensi/banyaknya individu yang termasuk dalam sel ke- $i, j$ , dengan  $i = 1, 2, \dots, r$  dan  $j = 1, 2, \dots, c$

$n_{r.} = \sum_{j=1}^c n_{rj}$  : frekuensi pengamatan pada baris kategori ke- $r$

$n_{.c} = \sum_{i=1}^r n_{ic}$  : frekuensi pengamatan pada kolom kategori ke- $c$

$n_{..} = n = \sum_{i=1}^r \sum_{j=1}^c n_{rc}$  : jumlah seluruh pengamatan

### 2.3 Missing Value

*Missing value* merupakan kasus yang terjadi ketika ada nilai dalam satu atau lebih variabel tidak tersedia untuk analisis. Dalam kehidupan nyata, peneliti sering berhadapan dengan beberapa kasus *missing value*. Hal yang diperhatikan oleh peneliti adalah bagaimana mengidentifikasi pola dan hubungan yang mendasari kasus *missing value* dalam suatu data agar peneliti dapat mempertahankan sedekat mungkin distribusi pada nilai asli ketika harus dilakukan perbaikan pada data tersebut. *Missing value* dapat menyebabkan hasil analisis yang diperoleh tidak akurat.

Aturan yang tersedia untuk menentukan apakah *missing value* dapat langsung dihilangkan atau harus diperbaiki antara lain untuk variabel yang memiliki kurang dari 15% data yang mengandung *missing*, maka variabel tersebut dapat dihilangkan,

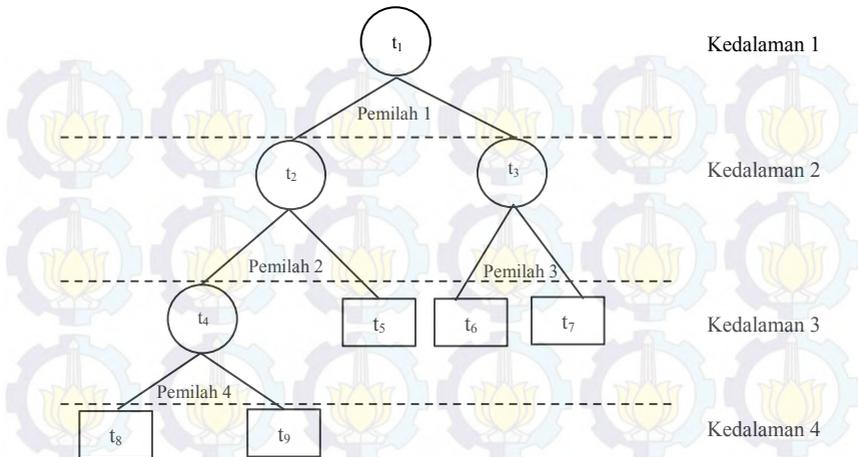
tetapi jika level dari *missing value* sebesar 20% hingga 30%, maka variabel tersebut perlu diperbaiki. Jika kasus *missing value* yang terjadi pada variabel respon dapat dihilangkan, untuk menghindari beberapa kepalsuan yang meningkat dalam hubungan variabel respon dengan variabel prediktor. Ketika menghilangkan variabel dalam analisis, pastikan bahwa variabel-variabel pengganti mempunyai korelasi yang tinggi untuk menggantikan variabel yang asli. Peneliti harus selalu mempertimbangkan performansi yang dihasilkan dari kedua analisis, baik dengan menghilangkan maupun tanpa menghilangkan data atau variabel, untuk mengidentifikasi beberapa perbedaan yang nyata (Hair, Black, Babin, & Anderson, 2010).

#### **2.4 Classification and Regression Trees (CART)**

CART merupakan salah satu metode nonparametrik yang digunakan untuk teknik pohon keputusan. Metode ini digunakan untuk menggambarkan hubungan antara variabel respon dengan satu atau lebih variabel prediktor. Apabila variabel responnya berbentuk kontinu, maka CART yang dihasilkan adalah regresi pohon, namun apabila variabel responnya berbentuk kategorik, maka CART akan menghasilkan klasifikasi pohon (Breiman *et al.*, 1993). Penggunaan variabel dalam suatu pengklasifikasian ada 4 komponen, yaitu variabel respon, variabel prediktor, data *learning*, data *testing*. Variabel respon merupakan karakteristik yang diharapkan dapat diprediksi dengan menggunakan variabel prediktor. Secara umum, ada banyak variabel prediktor yang mungkin mempengaruhi variabel respon. Himpunan data *learning* adalah himpunan data yang terdiri dari nilai-nilai variabel prediktor dan variabel respon yang berasal dari sekumpulan permasalahan. Permasalahan yang terdapat pada data *learning* mempunyai kemiripan dengan permasalahan yang akan diprediksi hasilnya untuk masa depan. Sedangkan himpunan data *testing* adalah himpunan data dari permasalahan yang akan diprediksi hasilnya secara tepat. Data tersebut digunakan untuk mengetahui seberapa tepat model yang sudah dibentuk atau klasifikasi yang telah dihasilkan oleh data *learning* (Lewis, 2000).

Metode pengklasifikasian CART memiliki beberapa kelebihan, diantaranya yang pertama adalah metode ini bersifat non parametrik sehingga tidak memerlukan asumsi yang mengikat seperti asumsi distribusi normal untuk variabel prediktor. Kedua, metode CART mampu mengeksplorasi data berdimensi tinggi dengan komputasi yang efisien. Ketiga, metode klasifikasi ini memperhitungkan interaksi antar variabel prediktor yang berpengaruh, karena diterapkan pengambilan keputusan secara bertahap dalam himpunan data yang kompleks. Keempat, kombinasi data kontinu maupun kategorik dapat digunakan pada metode ini. Kelima, CART tidak hanya memberikan klasifikasi, tetapi juga memberikan estimasi probabilitas kesalahan pengklasifikasian. Selain itu, hasil klasifikasi berbentuk sederhana dan mengklasifikasikan data baru secara efisien serta mudah diinterpretasikan (Lewis, 2000). Sedangkan kelemahan yang dimiliki oleh metode CART terletak pada hasil prediksi pohon klasifikasi yang dapat mengalami perubahan besar, yang disebabkan oleh perubahan data *learning* yang kurang stabil.

Analisis CART memiliki sifat "*binary recursive partitioning*". Istilah "*binary*" mengimplikasikan bahwa sekelompok data yang terkumpul dalam suatu ruang direpresentasikan sebagai simpul atau *node* pada pohon keputusan, dimana *node* tersebut hanya dapat diklasifikasikan menjadi dua grup. Dengan kata lain, setiap *node* dapat diklasifikasikan menjadi dua simpul anak (*child nodes*), dan dalam hal ini simpul utama disebut sebagai *parent node*. Istilah "*recursive*" dihubungkan pada kenyataan bahwa proses penyekatan secara biner dilakukan secara berulang-ulang. Artinya, setiap simpul utama dapat dipilah menjadi dua simpul anak, kemudian setiap simpul anak dapat memilah dirinya membentuk simpul anak yang lain, begitu seterusnya hingga memenuhi kriteria tertentu. Sedangkan istilah "*partitioning*" memiliki arti bahwa proses klasifikasi dapat dilakukan dengan cara memilah kumpulan data menjadi beberapa bagian atau partisi (Lewis, 2000).



**Gambar 2.1** Ilustrasi Struktur Pohon Klasifikasi

(Sumber: Breiman, Friedman, Olshen, & Stone, 1993)

Ilustrasi struktur pohon klasifikasi yang ditunjukkan pada Gambar 2.1. Simpul awal yang mengandung seluruh data dengan notasi  $t_1$ . Pada Gambar 2.1 simpul dalam (*internal node*) dinotasikan dengan  $t_2$ ,  $t_3$ , dan  $t_4$ , sedangkan simpul akhir (*terminal node*) dinotasikan dengan  $t_5$ ,  $t_6$ ,  $t_7$ ,  $t_8$ , dan  $t_9$  dimana setelahnya tidak ada lagi pemilahan, artinya simpul anak yang dihasilkan telah homogen. Setiap simpul berada pada kedalaman (*depth*) tertentu, dimulai dari simpul awal  $t_1$  yang berada pada kedalaman 1,  $t_2$  dan  $t_3$  berada pada kedalaman 2, dan begitu seterusnya hingga didapatkan simpul  $t_4$ ,  $t_5$ ,  $t_6$ ,  $t_8$ , dan  $t_9$  yang berada pada kedalaman 4.

Analisis CART memiliki tiga tahap dasar. Pertama, metode ini membentuk pohon klasifikasi. Prosedur pembentukan menggunakan pemilahan simpul secara berulang (*recursive*). Kemudian, setiap simpul akan ditentukan apakah simpul tersebut dapat dipilah menjadi simpul anak lagi atau tidak, dan kelas yang telah diprediksi akan diberikan tanda. Tahap kedua adalah *pruning* atau pemangkasan pohon klasifikasi yang menghasilkan rangkaian pohon klasifikasi yang lebih sederhana dengan

memangkas simpul yang kepentingannya meningkat. Tahap terakhir, penentuan pohon klasifikasi optimal, dimana pohon klasifikasi tersebut dapat merepresentasikan informasi dari himpunan data *learning*, tetapi tidak berlebihan (*overfit*).

#### 2.4.1 Pembentukan Pohon Klasifikasi Maksimal

Pada pembentukan pohon klasifikasi diawali dengan menentukan variabel apa dan nilai mana dari variabel tersebut yang layak dijadikan pemilah bagi setiap simpul. Dalam pembentukan pohon klasifikasi dibutuhkan data *learning*  $L$  yang terdiri atas pengamatan berukuran  $N$ . Ukuran pohon akan menentukan kekuatan prediksi atau akurasi pada pohon klasifikasi yang terbentuk. Proses pembentukan pohon klasifikasi terdiri dari tiga tahap, yaitu pemilihan pemilah, penentuan simpul terminal, dan penandaan label kelas.

##### 2.4.1.1 Pemilihan Pemilah

Pada tahap pemilihan pemilah dilakukan pemilahan pada sampel data *learning* ( $L$ ) berdasarkan aturan pemilahan dan kriteria *goodness of split*, dimana sampel data *learning* yang digunakan masih bersifat heterogen. Pemilihan pemilah tergantung pada jenis pohon atau pada jenis variabel respon. Himpunan bagian yang dihasilkan dari proses pemilahan harus lebih homogen dibandingkan dengan simpul induknya. Tingkat keheterogenan simpul tersebut dapat diukur menggunakan nilai *impurity* atau  $i(t)$ . Aturan pemilahan simpul induk menjadi dua simpul anak bergantung pada nilai yang berasal dari satu variabel prediktor. Setiap pemilahan hanya bergantung pada satu variabel prediktor saja. Apabila variabel prediktornya merupakan variabel kontinyu, maka pemilahan yang diperbolehkan adalah  $x_j \leq c_i$  dan  $x_j > c_i$  dengan  $i=1,2,3,\dots,n-1$  dengan  $c_i$  adalah nilai tengah atau median dari dua nilai amatan sampel yang berbeda dan berurutan. Sehingga jika terdapat sejumlah  $n$  sampel yang memiliki nilai berbeda pada variabel  $x_j$ , maka terdapat  $n-1$  kemungkinan pemilahan yang berbeda. Namun jika variabel

prediktornya merupakan variabel kategorik, maka pemilahan berasal dari semua kemungkinan pemilahan berdasarkan terbentuknya dua simpul yang saling lepas (*disjoint*). Bila kategori berskala nominal bertaraf  $L$ , maka akan diperoleh sebanyak  $2^{L-1}-1$  pemilahan yang mungkin terjadi. Sedangkan, jika kategori berskala ordinal bertaraf  $L$ , maka akan diperoleh sebanyak  $L-1$  pemilahan. Fungsi heterogenitas yang sering digunakan adalah indeks gini. Penggunaan indeks gini dalam pemilahan memiliki kelebihan, yaitu proses perhitungannya sederhana dan relatif cepat, serta mudah dan sesuai untuk diterapkan dalam berbagai kasus (Breiman *et al.*, 1993). Pemilah terbaik dipilih berdasarkan nilai penurunan tingkat keheterogenan yang paling tinggi dari semua kemungkinan pemilahan yang dilakukan pada setiap variabel prediktor. Fungsi Indeks Gini dituliskan dalam persamaan berikut.

$$i(t) = \sum_{j=1} p(i|t)p(j|t) \quad (2.1)$$

dengan,

$i(t)$  : fungsi keheterogenan indeks gini pada simpul- $t$

$p(i|t)$  : proporsi kelas- $i$  pada simpul- $t$

$p(j|t)$  : proporsi kelas- $j$  pada simpul- $t$

Langkah selanjutnya yaitu menentukan pemilah terbaik dari setiap variabel prediktor. Pemilah terbaik adalah pemilah yang memaksimalkan ukuran kehomogenan setiap simpul anak relatif terhadap simpul induknya dan memaksimalkan ukuran pemisahan antara dua simpul anak tersebut. Pemilah yang terpilih akan membentuk himpunan kelas yang disebut simpul satu. Setiap pemilahan akan dilakukan pada setiap simpul sampai diperoleh simpul akhir dan menghasilkan dua simpul anak. Jika terdapat data *missing value* pada variabel pemilah, maka pemilahan akan digantikan dengan variabel lain yang disebut *surrogate splitter* (pemilah pengganti), dimana nilai asosiasinya terbesar dibandingkan pemilah pengganti lainnya.

Kemudian menentukan kriteria *goodness of split* yang merupakan suatu evaluasi bagi pemilahan yang dilakukan oleh

pemilah  $s$  pada suatu simpul  $t$ . *Goodness of split* menunjukkan ukuran penurunan keheterogenan suatu kelas dengan formula sebagai berikut.

$$\phi(s, t) = \Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R) \quad (2.2)$$

dengan,

$\phi(s, t)$  : nilai *goodness of split*

$i(t)$  : fungsi keheterogenan pada simpul- $t$

$p_L$  : proporsi pengamatan simpul kiri

$p_R$  : proporsi pengamatan simpul kanan

$i(t_L)$  : fungsi keheterogenan pada simpul anak kiri

$i(t_R)$  : fungsi keheterogenan pada simpul anak kanan

Pemilah yang menghasilkan nilai *goodness of split* tertinggi merupakan pemilah terbaik (*variable importance*), karena dapat mereduksi heterogenitas lebih tinggi. Setiap variabel akan menghasilkan skor untuk menunjukkan seberapa besar variabel tersebut memberikan kontribusi dalam proses pembentukan pohon. Berikut ini merupakan persamaan untuk menentukan besarnya skor pada tiap variabel.

$$\text{skor} = \sum_{i=1}^n \phi(s, t_i) \quad (2.3)$$

Dimana  $\phi(s, t_i)$  merupakan nilai *goodness of split* pada setiap simpul. Nilai skor diperoleh dengan menjumlah nilai *goodness of split (improvement)* dari masing-masing variabel yang berperan sebagai *surrogate* untuk setiap simpul (CART *References Guide*, 2000). Ilustrasi dari perhitungan skor *variable importance* ditampilkan dalam Lampiran J.

#### 2.4.1.2 Penentuan Simpul Terminal

Suatu simpul  $t$  dikatakan sebagai simpul terminal ketika tidak terdapat penurunan heterogenitas yang signifikan, atau hanya terdapat satu pengamatan pada tiap simpul anak atau terdapat batasan minimum  $n$  pengamatan pada tiap simpul anak yang dihasilkan. Menurut Breiman, Friedman, Olshen, dan Stone (1993) pengembangan pohon akan berhenti jika dalam simpul hanya terdapat kurang dari 5 pengamatan.

### 2.4.1.3 Penandaan Label Kelas

Penandaan label kelas pada simpul terminal berdasarkan aturan jumlah terbanyak seperti yang ditunjukkan dalam persamaan berikut.

$$p(j_0|t) = \max p(j|t) = \max \frac{N_j(t)}{N(t)} \quad (2.4)$$

dengan,

$p(j|t)$  : proporsi kelas- $j$  pada simpul- $t$

$N_j(t)$  : jumlah pengamatan kelas- $j$  pada simpul- $t$

$N(t)$  : jumlah total pengamatan pada simpul- $t$

Label kelas untuk simpul terminal  $t$  adalah  $j_0$  yang memberikan nilai dugaan kesalahan pengklasifikasian pada simpul  $t$  yang paling kecil, yaitu sebesar  $r(t) = 1 - \max p_j(j|t)$ .

### 2.4.2 Pemangkasan Pohon Klasifikasi

Pemangkasan pohon klasifikasi atau yang biasa disebut *pruning* perlu dilakukan karena semakin banyak pemilahan yang dilakukan mengakibatkan makin kecilnya tingkat kesalahan prediksi (*overfitting*), artinya nilai prediksi melebihi nilai yang sebenarnya. Selain itu, bila dalam proses pemilahan diberikan batasan, padahal pemilahan masih layak untuk dilakukan maka akan terjadi kasus *underfitting*. *Pruning* merupakan suatu penilaian ukuran pohon tanpa mengorbankan ketepatan atau kebaikannya melalui pengurangan simpul pohon sehingga dicapai ukuran pohon yang layak. Caranya, pemangkasan pohon dilakukan dengan *cost complexity minimum* (Breiman *et al.*, 1993). Untuk  $\alpha \geq 0$  maka ukuran *cost complexity* adalah sebagai berikut.

$$R_\alpha(T) = R(T) + \alpha |\tilde{T}| \quad (2.5)$$

dimana,

$R_\alpha(T)$  : ukuran kompleksitas suatu pohon T pada kompleksitas  $\alpha$

$R(T)$  : penduga pengganti (*resubstitution estimate*) pohon atau ukuran kesalahan klasifikasi pohon T

$\alpha$  : parameter *cost complexity* bagi penambahan satu simpul terminal pada pohon T

$|\tilde{T}|$  : banyaknya simpul terminal pada pohon T

*Cost complexity pruning* digunakan untuk menentukan pohon bagian  $T(\alpha)$  yang dapat meminimumkan  $R_\alpha(T)$  pada seluruh pohon bagian atau untuk setiap nilai  $\alpha$ . Nilai parameter kompleksitas ( $\alpha$ ) akan secara perlahan meningkat selama proses pemangkasan. Selanjutnya, pencarian pohon bagian  $T(\alpha) < T_{maks}$  yang dapat meminimumkan  $R_\alpha(T)$ . Pemangkasan pohon dimulai dengan mengambil  $t_R$  dan  $t_L$  dari  $T_{maks}$  yang dihasilkan dari simpul induk  $t$ . Jika diperoleh dua simpul anak dari proses pemilahan yang dilakukan pada simpul induk yang memenuhi persamaan  $R(t) = R(t_R) + R(t_L)$ , maka dua simpul anak akan dipangkas. Sehingga diperoleh pohon  $T_1$  yang memenuhi kriteria  $R(T_1) = R(T_{maks})$ . Proses ini terus dilakukan secara berulang hingga tidak mungkin lagi dilakukan pemangkasan. Jika  $R(T)$  digunakan sebagai kriteria penentuan pohon klasifikasi optimal, maka nilai penduga pengganti tersebut akan cenderung memilih pohon besar  $T_j$ . Karena semakin besar pohon, semakin kecil nilai penduga penggantian.

Hasil yang diperoleh dari tahap pemangkasan akan berupa urutan pohon yaitu  $T_{maks} > T_1 > T_2 > \dots > \{t_1\}$ . Urutan pohon tersebut memiliki nilai  $\alpha$  yang semakin menurun, yaitu  $\alpha_k < \alpha_{k+1}$  dimana  $\alpha_1 = 0$  untuk  $k \geq 1$  dan  $T(\alpha) = T(\alpha_k) = T_k$ .

### 2.4.3 Penentuan Pohon Klasifikasi Optimal

Ukuran pohon klasifikasi yang sangat besar memberikan nilai penduga yang sangat kecil, sehingga pohon tersebut lebih dipilih untuk menduga nilai respon. Namun, ukuran pohon yang besar bersifat *overfitting* sehingga menyebabkan nilai kompleksitas yang tinggi. Karena struktur data yang digambarkan cenderung kompleks, maka perlu pemilihan pohon yang optimum dengan ukuran sederhana dan memberikan nilai penduga pengganti cukup kecil.

Penduga pengganti yang sering digunakan jika ukuran pengamatan yang cukup besar adalah *Test Sample Estimate*. Prosedur ini diterapkan dengan membagi sampel L menjadi dua himpunan, yaitu  $L_1$  dan  $L_2$ . Pengamatan  $L_1$  digunakan untuk

membentuk pohon  $T$ . Sedangkan pengamatan  $L_2$  digunakan untuk menduga  $R(T)$ . Jika  $N_2$  merupakan jumlah pengamatan  $L_2$ , dan  $X(\cdot)$  bernilai 0 jika pernyataan dalam kurung salah dan bernilai 1 jika pernyataan dalam tanda kurung benar. Penduga sampel uji dapat ditunjukkan dalam persamaan berikut.

$$R^{ts}(T_t) = \frac{1}{N_2} \sum (x_n \cdot j_n) \in L_2 X(d(x_n) \neq j_n) \quad (2.6)$$

dengan,

$R^{ts}(T_t)$  : total proporsi kesalahan *test sample estimate*

$N_2$  : jumlah pengamatan dari data *learning* ( $L_2$ )

Karena dalam hal ini ingin menduga proporsi kesalahan yang dihasilkan dari proses pembentukan pohon klasifikasi, sehingga pohon klasifikasi optimal yang dipilih adalah pohon  $T_t$  yang memiliki nilai penduga sampel uji minimum atau  $R^{ts}(T_t) = \min_t R^{ts}(T_t)$ .

Apabila ukuran pengamatan tidak cukup besar, maka metode yang digunakan untuk mencari nilai penduga pengganti adalah *v-fold cross validation estimate*. Dalam *v-fold cross validation*, data sampel dibagi secara random menjadi  $v$  bagian dengan jumlah kasus pada setiap bagian sama (sedekat mungkin jumlahnya pada tiap bagian) dan dilakukan pengulangan sebanyak  $v$  kali. Nilai  $v$  yang sering digunakan adalah 10 sehingga menjadi *10-fold cross validation estimate*. Nilai 10 sering digunakan karena menghasilkan estimasi *error* yang paling baik dan membagi data menjadi proporsi yang seimbang. Data akan dibagi menjadi 10 bagian, dimana 9 bagian sebagai data *learning* dan 1 bagian sebagai data *testing*. Kemudian dilakukan pengulangan hingga 10 kali. Sehingga setiap data memiliki peluang menjadi data *learning* atau data *testing* (Witten, Frank, & Hall, 2011).

Dalam *v-fold cross validation estimate*, pohon akan terbentuk menggunakan data *learning* ( $(v-1)/v$  bagian), yaitu  $T_{maks}^{(v)}$  dengan  $v = 1, \dots, v$  dengan kriteria pemilahan sampai simpul terminal memiliki jumlah pengamatan paling minimum.

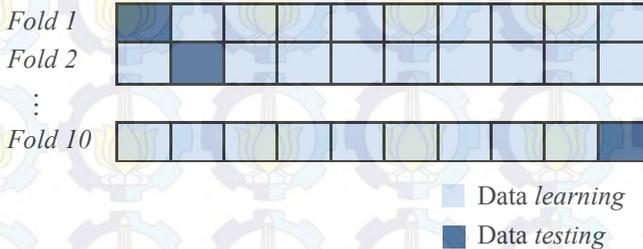
Apabila  $d^{(v)}(x)$  adalah hasil pengklasifikasian dari pohon bentukun tiap *fold*, maka penduga sampel uji untuk  $R^{cv}(T_t^{(v)})$  adalah sebagai berikut.

$$R^{cv}(T_t^{(v)}) = \frac{1}{N_v} \sum_{(x_n, j_n) \in L_v} X(d^{(v)}(x_n) \neq j_n) \quad (2.7)$$

dimana  $N \cong N/V$  adalah jumlah pengamatan dalam  $L_v$ . Tahap selanjutnya adalah dilakuka prosedur yang sama menggunakan semua pengamatan dalam  $L$  untuk membentuk deret pohon  $T_t$ . Penduga *v-fold cross validation* untuk  $T_t^{(v)}$  adalah sebagai berikut.

$$R^{cv}(T_t) = \frac{1}{V} \sum_{v=1}^V R^{cv}(T_t^{(v)}) \quad (2.8)$$

Pohon klasifikasi optimal merupakan pohon klasifikasi ( $T^*$ ) dengan  $R^{cv}(T^*) = \min_t R^{cv}(T_t)$ . Berikut ini merupakan ilustrasi pembagian data pada metode *10-fold cross validation estimate*. Sedangkan ilustrasi untuk menjalankan metode *10-fold cross validation* pada data selengkapnya ditampilkan pada Lampiran K.



**Gambar 2.2** Ilustrasi Prosedur *10-fold Cross Validation*

#### 2.4.4 Ukuran Ketepatan Klasifikasi

Untuk mengevaluasi hasil ketepatan klasifikasi yaitu dengan cara menghitung akurasi klasifikasi. Pada penelitian ini evaluasi akurasi klasifikasi ialah dengan menghitung nilai APER (*apparent error rate*) dan 1-APER (*total accuracy rate*). APER merupakan proporsi observasi yang diprediksi secara tidak benar.

Sedangkan tingkat akurasi total merupakan proporsi observasi yang diprediksi secara benar oleh fungsi klasifikasi. Nilai total tingkat kesalahan merupakan proporsi observasi yang diprediksi secara tidak benar oleh fungsi klasifikasi (Johnson dan Winchern, 2007). *Sensitivity* menggambarkan akurasi pada sampel kelas  $i$ , sedangkan *specificity* menggambarkan akurasi pada kelas  $j$ . Metode klasifikasi yang baik seharusnya mampu mengukur *sensitivity* dan *specificity* sama baiknya. Berikut disajikan *crosstab* untuk menghitung ketepatan klasifikasi yang ditunjukkan dalam Tabel 2.2.

**Tabel 2.2** Tabulasi Silang Ketepatan Klasifikasi

Kelas Pengamatan Y	Kelas Prediksi Y		Total
	1	2	
1	$n_{11}$	$n_{12}$	$N_{1.}$
2	$n_{21}$	$n_{22}$	$N_{2.}$
Total	$N_{.1}$	$N_{.2}$	$N$

keterangan:

$n_{11}$  : Jumlah pengamatan dari variabel Y kelas 1 yang tepat diprediksi sebagai variabel Y kelas 1

$n_{12}$  : Jumlah pengamatan dari variabel Y kelas 1 yang salah diprediksi sebagai variabel Y kelas 2

$n_{21}$  : Jumlah pengamatan dari variabel Y kelas 2 yang salah diprediksi sebagai variabel Y kelas 1

$n_{22}$  : Jumlah pengamatan dari variabel Y kelas 2 yang tepat diprediksi sebagai variabel Y kelas 2

$N_{1.}$  : Jumlah pengamatan dari variabel Y kelas 1

$N_{2.}$  : Jumlah pengamatan dari variabel Y kelas 2

$N_{.1}$  : Jumlah prediksi dari variabel Y kelas 1

$N_{.2}$  : Jumlah prediksi dari variabel Y kelas 2

$N$  : Jumlah total pengamatan/prediksi

Berikut ini adalah formula untuk menghitung total akurasi, *sensitivity* dan *specificity*.

$$\text{Apparent Error Rate (APER)} = \frac{\text{Jumlah prediksi salah}}{\text{Total prediksi}} = \frac{n_{21} + n_{12}}{N} \quad (2.9)$$

$$\text{Total Accuracy Rate (1 - APER)} = \frac{\text{Jumlah prediksi benar}}{\text{Total prediksi}} = \frac{n_{11} + n_{22}}{N} \quad (2.10)$$

$$\text{Sensitivity} = \frac{n_{11}}{N_1} \quad (2.11)$$

$$\text{Specificity} = \frac{n_{22}}{N_2} \quad (2.12)$$

### 2.4.5 Odds Ratio

*Odds ratio* merupakan nilai yang menunjukkan besarnya pengaruh antara kategori satu dengan kategori dua (kategori dua terhadap variabel respon dengan kategori pembanding) dalam satu variabel prediktor tersebut. Berikut ini merupakan persamaan untuk menghitung *odds ratio* (Agresti, 2007).

$$\theta = \frac{n_{11}n_{22}}{n_{12}n_{21}} \quad (2.13)$$

### 2.4.6 Ilustrasi CART

Algoritma CART dengan pemilihan pemilah menggunakan indeksi gini dalam pembentukan pohon klasifikasi dapat diilustrasikan dengan 2 variabel prediktor yaitu berat badan lahir bayi dan ketuban pecah dini serta satu variabel respon yang terdiri dari dua kelas yakni kelas 0 adalah bayi negatif berisiko infeksi dan kelas 1 adalah bayi positif berisiko infeksi. Variabel yang digunakan dalam ilustrasi CART dijelaskan dalam Tabel 2.3 sebagai berikut.

**Tabel 2.3** Variabel Penelitian Ilustrasi CART

Variabel	Nama Variabel	Kategori	Skala Pengukuran
Y	Status Risiko	0: negatif	Nominal
	Infeksi	1: positif	
X <sub>1</sub>	Berat Badan Lahir Bayi	-	Rasio
X <sub>2</sub>	KPD (Ketuban Pecah Dini)	0: tidak	Nominal
		1: ya	

Pada ilustrasi CART menggunakan data bangkitan sebanyak 10 sampel pengamatan. Data sampel yang digunakan

untuk ilustrasi pembentukan pohon klasifikasi dengan metode CART ditampilkan dalam Tabel 2.4 sebagai berikut.

Tabel 2.4 Data Sampel Ilustrasi CART

Data Sampel ke-	Y	X <sub>1</sub>	X <sub>2</sub>
1	0	3500	0
2	1	3300	1
3	0	3500	0
4	1	2300	1
5	1	500	1
6	1	2600	0
7	0	3300	0
8	0	2600	0
9	1	2300	1
10	1	2600	1

Data sampel hasil bangkitan tersebut akan mementuk simpul induk (*parent node*) yang kemudian akan memilah menjadi simpul anak (*child nodes*). Algoritma CART yang pertama adalah pembentukan pohon klasifikasi maksimal dengan pemilihan pemilah menggunakan indeks gini. Sebelum melakukan proses pemilihan pemilah dalam analisis CART, terlebih dahulu menghitung nilai keheterogenan simpul induk berdasarkan fungsi keheterogenan indeks gini yang telah didefinisikan pada persamaan (2.1). Penentuan pemilah dari semua kemungkinan variabel dan nilai variabel (*threshold*) dilakukan untuk memilah simpul induk menjadi dua simpul anak. Sebelum menentukan pemilah perlu dilakukan perhitungan untuk mencari nilai *goodness of split* menggunakan persamaan (2.2) dari kemungkinan *threshold* pada tiap variabel yang dihasilkan. Setelah itu dilakukan perbandingan terhadap nilai *goodness of split* dari tiap variabel. Variabel yang terpilih merupakan variabel yang memiliki nilai *goodness of split* yang paling tinggi. Selanjutnya *threshold* variabel terpilih tersebut dijadikan sebagai pemilah

terbaik untuk simul awal. Berikut ini merupakan contoh perhitungan manual dari proses pemilihan pemilah terbaik.

- a. Perhitungan Indeks Gini untuk Variabel  $Y$

**Tabel 2.5** Proporsi Pengamatan Variabel  $Y$

Variabel $Y$	$N_i(t)$	$p(i t)$
Risiko Negatif (0)	4	$4/10 = 0,4$
Risiko Positif (1)	6	$6/10 = 0,6$
Total	10	1

Kemudian perhitungan Indeks Gini dicari berdasarkan persamaan (2.1) sebagai berikut.

$$i(t) = \sum_{j=1} p(i|t)p(j|t), i \neq j \text{ untuk } i, j = 0, 1$$

$$i(t) = 2p(0|t)p(1|t) = 2*(0,4)(0,6) = 0,48$$

Indeks Gini yang diperoleh untuk variabel respon  $Y$  (status risiko infeksi pada bayi baru lahir) yaitu sebesar 0,48.

- b. Penentuan *Threshold* dan Perhitungan Nilai *Goodness of Split* Variabel Berat Badan Lahir Bayi ( $X_1$ )

Variabel  $X_1$  memiliki skala pengukuran kontinyu, sehingga *threshold* pada variabel ini diperoleh melalui nilai tengah (median) dari dua nilai amatan sampel berurutan yang berbeda. Berikut adalah kemungkinan nilai median yang diperoleh.

Variabel $X_1$	Median
500	1400
2300	
2300	2450
2600	
2600	2950
2600	
3300	

3300	}	3400
3500		
3500		

Berdasarkan perhitungan diperoleh kemungkinan nilai median sebanyak 4 nilai. *Threshold* untuk variabel  $X_1$  dengan nilai median 1400 maka untuk data  $X_1 \leq 1400$  akan dipilah ke simpul kiri, sedangkan untuk data  $X_1 > 1400$  akan dipilah ke dalam simpul kanan. Selanjutnya perhitungan nilai *goodness of split* ditunjukkan dalam Tabel 2.6 sebagai berikut.

**Tabel 2.6** Perhitungan *Goodness of Split* untuk *Threshold*  $X_1=1400$

Simpul Kiri		Simpul Kanan	
$N_L = 1$	$p_L = 0,1$	$N_R = 9$	$p_R = 0,9$
$N(0   t_L) = 0$	$p(0   t_L) = 0$	$N(0   t_R) = 4$	$p(0   t_R) = 4/9$
$N(1   t_L) = 1$	$p(1   t_L) = 1$	$N(1   t_R) = 5$	$p(1   t_R) = 5/9$

Ket: L=left (kiri); R=right (kanan)

Perhitungan Indeks Gini untuk simpul kanan dan simpul kiri adalah sebagai berikut.

$$\text{Simpul kiri : } i(t_L) = 2 * (0)(1) = 0$$

$$\text{Simpul kanan : } i(t_R) = 2 * (4/9)(5/9) = 40/81$$

Sehingga perhitungan *goodness of split* untuk *threshold*  $X_1=1400$  adalah sebagai berikut.

$$\phi(s, t) = 0,48 - 0,1(0) - 0,9(40/81) = 0,035$$

Pada *threshold* dengan nilai median sebesar 2450, data  $X_1 \leq 2450$  akan dipilah ke simpul kiri, sedangkan untuk data  $X_1 > 2450$  akan dipilah ke dalam simpul kanan. Selanjutnya perhitungan nilai *goodness of split* untuk *threshold*  $X_1=2450$  ditunjukkan dalam Tabel 2.7 sebagai berikut.

**Tabel 2.7** Perhitungan *Goodness of Split* untuk *Threshold*  $X_1=2450$ 

Simpul Kiri		Simpul Kanan	
$N_L = 3$	$p_L = 0,3$	$N_R = 7$	$p_R = 0,7$
$N(0 t_L) = 0$	$p(0 t_L) = 0$	$N(0 t_R) = 4$	$p(0 t_R) = 4/7$
$N(1 t_L) = 3$	$p(1 t_L) = 1$	$N(1 t_R) = 3$	$p(1 t_R) = 3/7$

Perhitungan Indeks Gini untuk simpul kanan dan simpul kiri adalah sebagai berikut.

$$\text{Simpul kiri : } i(t_L) = 2 * (0)(1) = 0$$

$$\text{Simpul kanan : } i(t_R) = 2 * (4/7)(3/7) = 24/49$$

Sehingga perhitungan *goodness of split* untuk *threshold*  $X_1=2450$  adalah sebagai berikut.

$$\phi(s, t) = 0,48 - 0,3(0) - 0,7(24/49) = 0,137$$

Pada *threshold* dengan nilai median sebesar 2950, data  $X_1 \leq 2950$  akan dipilah ke simpul kiri, sedangkan untuk data  $X_1 > 2950$  akan dipilah ke dalam simpul kanan. Selanjutnya perhitungan nilai *goodness of split* untuk *threshold*  $X_1=2950$  ditunjukkan dalam Tabel 2.8 sebagai berikut.

**Tabel 2.8** Perhitungan *Goodness of Split* untuk *Threshold*  $X_1=2950$ 

Simpul Kiri		Simpul Kanan	
$N_L = 6$	$p_L = 0,6$	$N_R = 4$	$p_R = 0,4$
$N(0 t_L) = 1$	$p(0 t_L) = 1/6$	$N(0 t_R) = 3$	$p(0 t_R) = 3/4$
$N(1 t_L) = 5$	$p(1 t_L) = 5/6$	$N(1 t_R) = 1$	$p(1 t_R) = 1/4$

Perhitungan Indeks Gini untuk simpul kanan dan simpul kiri adalah sebagai berikut.

$$\text{Simpul kiri : } i(t_L) = 2 * (1/6)(5/6) = 5/18$$

$$\text{Simpul kanan : } i(t_R) = 2 * (3/4)(1/4) = 3/8$$

Sehingga perhitungan *goodness of split* untuk *threshold*  $X_1=2950$  adalah sebagai berikut.

$$\phi(s, t) = 0,48 - 0,6(5/18) - 0,4(3/8) = 0,163$$

Pada *threshold* dengan nilai median sebesar 3400, data  $X_1 \leq 3400$  akan dipilah ke simpul kiri, sedangkan untuk data  $X_1 > 3400$  akan dipilah ke dalam simpul kanan. Selanjutnya perhitungan nilai *goodness of split* untuk *threshold*  $X_1=3400$  ditunjukkan dalam Tabel 2.9 sebagai berikut.

**Tabel 2.9** Perhitungan *Goodness of Split* untuk *Threshold*  $X_1=3400$

Simpul Kiri		Simpul Kanan	
$N_L = 8$	$p_L = 0,8$	$N_R = 2$	$p_R = 0,2$
$N(0 t_L) = 2$	$p(0 t_L) = 2/8$	$N(0 t_R) = 2$	$p(0 t_R) = 1$
$N(1 t_L) = 6$	$p(1 t_L) = 6/8$	$N(1 t_R) = 0$	$p(1 t_R) = 0$

Perhitungan Indeks Gini untuk simpul kanan dan simpul kiri adalah sebagai berikut.

$$\text{Simpul kiri : } i(t_L) = 2 * (2/8)(6/8) = 3/8$$

$$\text{Simpul kanan : } i(t_R) = 2 * (1)(0) = 0$$

Sehingga perhitungan *goodness of split* untuk *threshold*  $X_1=3400$  adalah sebagai berikut.

$$\phi(s, t) = 0,48 - 0,8(3/8) - 0,2(0) = 0,18$$

c. Penentuan *Threshold* dan Perhitungan Nilai *Goodness of Split* Variabel KPD ( $X_2$ )

Variabel  $X_2$  merupakan variabel yang memiliki skala pengukuran kategorik (nominal) dengan jumlah kategori sebanyak dua kelas, sehingga banyaknya kemungkinan *threshold* adalah sebanyak satu kemungkinan pemilah. Hasil tersebut diperoleh melalui perhitungan sebagai berikut.

$$2^{L-1} - 1 = 2^{2-1} - 1 = 1$$

Sehingga data pada variabel  $X_2$  yang termasuk dalam kelas 0 dipilah sebagai simpul kiri, sedangkan data yang termasuk dalam kelas 1 dipilih sebagai simpul kanan. Perhitungan nilai *goodness of split* ditampilkan dalam Tabel 2.10 sebagai berikut.

**Tabel 2.10** Perhitungan *Goodness of Split* untuk *Threshold* Variabel  $X_2=0:1$ 

Simpul Kiri		Simpul Kanan	
$N_L = 5$	$p_L = 0,5$	$N_R = 5$	$p_R = 0,5$
$N(0 t_L) = 4$	$p(0 t_L) = 4/5$	$N(0 t_R) = 0$	$p(0 t_R) = 0$
$N(1 t_L) = 1$	$p(1 t_L) = 1/5$	$N(1 t_R) = 5$	$p(1 t_R) = 1$

Perhitungan Indeks Gini untuk simpul kanan dan simpul kiri adalah sebagai berikut.

$$\text{Simpul kiri : } i(t_L) = 2 * (4/5)(1/5) = 8/25$$

$$\text{Simpul kanan : } i(t_R) = 2 * (0)(1) = 0$$

Sehingga perhitungan *goodness of split* untuk *threshold*  $X_2=0:1$  adalah sebagai berikut.

$$\phi(s,t) = 0,48 - 0,5(8/25) - 0,5(0) = 0,32$$

Selanjutnya dilakukan pemilihan pemilah terbaik dengan memilih *threshold* sebagai pemilah awal yang memiliki nilai *goodness of split* paling maksimum.

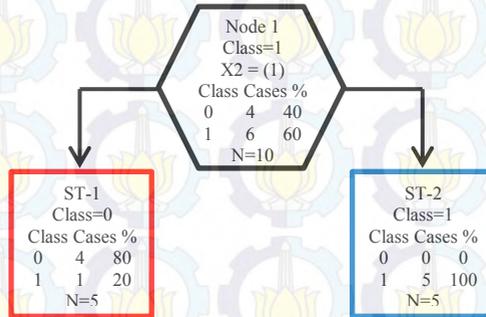
Berdasarkan Tabel 2.11 dapat diketahui bahwa pemilah terbaik merupakan variabel  $X_2$  (Ketuban Pecah Dini) dengan *threshold* 1:2, dimana variabel  $X_2$  memiliki *goodness of split* paling maksimum yaitu sebesar 0,32. Setelah itu dilakukan pemilahan pada data sampel berdasarkan pemilah awal tersebut. Sehingga data sampel pada variabel  $X_2$  yang memiliki kelas 0 masuk ke dalam simpul kiri, sedangkan variabel  $X_2$  yang memiliki kelas 1 masuk ke dalam simpul kanan.

**Tabel 2.11** Perbandingan *Goodness of Split* dari Pemilah Awal

Variabel	Threshold	Goodness of Split
$X_1$	1400	0,035
	2450	0,137
	2950	0,163
	3400	0,18
$X_2$	0:1	0,32*

\**goodness of split* paling maksimum

Berikut ini merupakan hasil pembentukan pohon klasifikasi berdasarkan pemilahan awal yang ditunjukkan dalam Gambar 2.3 sebagai berikut.



**Gambar 2.3** Pohon Klasifikasi Ilustrasi Hasil Pemilahan

Berdasarkan Breiman, Friedman, Olshen dan Stone (1993) pembentukan pohon akan berhenti ketika jumlah anggota pada simpul terminal berjumlah kurang dari sama dengan 5 pengamatan. Sehingga seperti yang ditunjukkan pada Gambar 2.2 dapat diketahui bahwa simpul terminal 1 dan 2 memiliki pengamatan masing-masing sebanyak lima pengamatan. Sehingga pohon tersebut dapat dikatakan sebagai pohon klasifikasi maksimal dan simpul terminal yang terbentuk sebanyak dua.

Kemudian dilakukan pelabelan kelas pada tiap simpul terminal untuk mengetahui karakteristik masing-masing kelas pada variabel respon. Penandaan label kelas yaitu label kelas 0 untuk bayi baru lahir yang negatif berisiko infeksi, sedangkan label kelas 1 untuk bayi baru lahir yang memiliki risiko infeksi yang ditentukan sesuai dengan persamaan (2.4) dengan mencari nilai peluang yang tertinggi. Penandaan label pada simpul terminal ditunjukkan dalam Tabel 2.12 yang menunjukkan bahwa label kelas untuk masing-masing simpul terminal berdasarkan proporsi tertinggi. Pada simpul terminal 1 dapat diketahui bahwa terdapat 4 pengamatan termasuk kelas 0 dan 1 pengamatan termasuk dalam kelas 1. Karena simpul terminal 1 didominasi oleh data yang termasuk dalam kelas 0, maka simpul terminal 2

diberi label kelas 0 yang ditunjukkan oleh kotak berwarna merah. Sedangkan simpul terminal 2 memiliki 5 pengamatan dan seluruh data sampel pengamatan tersebut masuk ke dalam kelas 1. Sehingga simpul terminal 2 diberi label kelas 1 yang ditunjukkan dengan kotak berwarna biru pada Gambar 2.2. berdasarkan 10 data sampel pengamatan dapat diketahui bahwa 9 pengamatan tepat diklasifikasikan, sedangkan terdapat 1 pengamatan yang mengalami salah pengklasifikasian. Proses tersebut dapat dilanjutkan dengan melakukan proses *pruning* atau pemangkasan pohon klasifikasi, apabila pohon klasifikasi yang terbentuk terlalu kompleks. Pemangkasan pohon klasifikasi dapat dilakukan sesuai dengan persamaan (2.5). Proses ini dilakukan untuk memperoleh pohon klasifikasi yang optimal.

**Tabel 2.12** Penandaan Label Kelas pada Simpul Terminal

Simpul Terminal	Kelas	$N(j/t)$	$p(j/t)$	Label Kelas
ST-1	0	4	80%	Bayi Negatif Berisiko Infeksi
	1	1	20%	
Total		5	100%	
ST-2	0	0	0	Bayi Positif Berisiko Infeksi
	1	5	100%	
Total		5	100%	

Selanjutnya pohon klasifikasi optimal yang terbentuk dapat digunakan untuk mengklasifikasi bayi yang baru lahir berdasarkan berat badan lahir bayi ( $X_1$ ) dan kejadian ketuban pecah dini ( $X_2$ ). Ilustrasi data bayi baru lahir ditunjukkan dalam Tabel 2.13.

Aturan yang digunakan untuk mengklasifikasikan bayi baru lahir positif memiliki risiko infeksi atau tidak adalah dengan melihat kondisi yang dimiliki oleh bayi tersebut saat lahir, dalam hal ini adalah melihat kejadian ketuban pecah dini ( $X_2$ ). Jika bayi lahir dimana ibu bayi mengalami ketuban pecah dini sebelum

persalinan (kategori 1), maka bayi tersebut memiliki risiko infeksi. Sebaliknya, jika bayi baru lahir dari ibu yang tidak mengalami ketuban pecah dini sebelum persalinan (kategori 0), maka bayi tersebut negatif risiko infeksi.

**Tabel 2.13** Data Ilustrasi Bayi Baru Lahir

Bayi Baru	$X_1$	$X_2$
A	2600	0
B	2300	0
C	1700	0
D	3900	1
E	2700	1

Sehingga hasil pengklasifikasian bayi baru lahir berdasarkan kondisi berat badan bayi sewaktu lahir ( $X_1$ ) dan status kejadian ketuban pecah dini sebelum proses persalinan ( $X_2$ ) yang ditunjukkan dalam Tabel 2.14 sebagai berikut

**Tabel 2.14** Hasil Klasifikasi Ilustrasi Bayi Baru Lahir

Bayi Baru Lahir	$X_1$	$X_2$	Status Risiko Infeksi (Y)
A	2600	0	Negatif
B	2300	0	Negatif
C	1700	0	Negatif
D	3900	1	Positif
E	2700	1	Positif

Berdasarkan 5 data bayi baru lahir seperti yang ditunjukkan dalam Tabel 2.13 dapat diketahui status risiko infeksi pada bayi tersebut. Sebanyak 3 bayi baru lahir tidak memiliki risiko infeksi (normal), sedangkan terdapat 2 bayi baru lahir yang memiliki risiko infeksi.

## 2.5 Periode Perinatal dan Neonatal Dini

Bayi yang termasuk dalam periode perinatal dimulai dari saat bayi masih dalam kandungan dengan usia kehamilan 22 minggu

(sekitar 154 hari) atau ketika bayi lahir dengan berat lahir 500 gram, sampai pada bayi berusia 7 hari setelah lahir. Sedangkan periode neonatal adalah periode yang dimulai dari bayi lahir usia 0 hari sampai usia 28 hari setelah kelahiran. Periode neonatal dapat dibagi menjadi 2, yakni *early neonatal period* (periode neonatal dini) dan *late neonatal period*. Bayi yang berusia kurang dari sama dengan 7 hari termasuk periode neonatal dini. Sementara itu, bayi yang berusia lebih dari 7 hari hingga 28 hari setelah kelahiran termasuk *late neonatal period* (WHO, 2002).

### **2.6 Infection Specific to the Perinatal Period**

Umumnya, infeksi khusus yang terjadi pada periode perinatal dianggap kongenital dan dapat diklasifikasikan dalam diagnosa pada perinatal dengan aturan kode internasional ICD-10 *medical condition* (*International Classification of Disease*). Dalam ICD-10 *infection specific to the perinatal period* memiliki kode diagnosa P35-P39. Bayi diberikan diagnosa tambahan dengan kode P35-P39 dapat diperoleh sebelum kelahiran melalui umbilikus atau setelah proses kelahiran. Infeksi yang terjadi setelah lahir, tetapi muncul selama periode neonatal dini dilihat apakah ada indikasi terkena paparan infeksi bukan dari bawaan lahir. Oleh sebab itu, dokter harus memberikan klarifikasi mengenai diagnosa infeksi pada bayi, ketika catatan rekam medis tidak sepenuhnya jelas. Untuk kode P39.9 merupakan kode diagnosa yang diberikan jika bayi belum diketahui infeksi spesifiknya atau dalam dunia kesehatan biasa disebut *infection unspecified to the perinatal period* (WHO, 2010).

### **2.7 Perawatan Neonatal Esensial Pada Saat Lahir**

Bayi baru lahir sangat rentan terhadap infeksi yang disebabkan oleh paparan atau kontaminasi mikroorganisme selama proses persalinan berlangsung maupun beberapa saat setelah lahir. Beberapa mikroorganisme harus diwaspadai karena dapat ditularkan lewat percikan darah dan cairan tubuh, misalnya virus HIV, Hepatitis B dan Hepatitis C. Penilaian awal untuk semua bayi baru lahir mengenai usia kehamilan dan warna air

ketuban jernih, serta memastikan air ketuban tidak bercampur dengan mekonium. Setelah bayi lahir, bayi diletakkan di atas kain bersih dan kering yang telah disiapkan pada perut bawah ibu. Kemudian dilakukan penilaian lagi mengenai tangisan bayi, pernafasan bayi, warna kulit, denyut jantung, tonus otot bayi baik atau bayi bergerak aktif, pengukuran suhu ketiak, meraba bagian kepala, mata, mulut, perut, tali pusat, punggung atau tulang belakang, lubang anus, alat kelamin luar, ekstremitas (jumlah jari tangan dan kaki), timbang bayi, mengukur panjang dan lingkaran kepala bayi dan menilai cara menyusui. Jika bayi baru lahir cukup bulan dengan air ketuban jernih dan langsung menangis atau bernapas spontan serta bergerak aktif cukup dilakukan perawatan bayi baru lahir normal. Kementerian Kesehatan RI (2010) membuat pedoman asuhan perawatan bayi baru lahir seperti berikut.

1. Jaga bayi tetap hangat
  2. Isap lendir dari mulut dan hidung (hanya jika perlu)
  3. Keringkan
  4. Pemantauan tanda bahaya
  5. Klem, potong dan ikat tali pusat tanpa membubuhi apapun, kira-kira 2 menit setelah lahir (atau setelah bidan menyuntikkan oksitosin kepada ibu) untuk memberi waktu tali pusat mengalir darah (dengan demikian juga zat besi) kepada bayi
  6. Lakukan inisiasi menyusui dini
  7. Beri suntikan vitamin K1 mg intramuskular, di paha kiri anterolateral setelah inisiasi menyusui dini
  8. Beri salep mata antibiotika pada kedua mata
  9. Pemeriksaan fisis
  10. Beri imunisasi Hepatitis B 0,5 ml intramuskular, di paha kanan anteroleteral, kira-kira 1-2 jam setelah pemberian vitamin K1
- Jika bayi kurang bulan (< 37 minggu atau 259 hari) atau lebih bulan ( $\geq$  42 minggu atau 283 hari) dan atau air ketuban bercampur mekonium dan atau tidak bernapas secara spontan

(megap-megap) dan atau tonus otot tidak baik, maka perlu dilakukan perawatan bayi baru lahir dengan asfiksia.

Tanda atau gejala bayi termasuk klasifikasi penyakit sangat berat atau infeksi bakteri berat antara lain tidak mau minum, ri a at kejang, bergerak han a jika dirangsang, napas cepat ( $\geq 60$  kali menit), napas lambat ( $< 30$  kali menit), tarikan dinding dada ke dalam ang kuat, merintih, demam ( $> 37,5$  C), hipotermia ( $< 35,5$  C), nanah di mata atau pusar kemerahan meluas sampai dinding perut. Jika bayi termasuk pada klasifikasi penyakit sangat berat atau infeksi berat, maka bayi perlu diberikan antibiotik intramuskular. Antibiotik pilihan pertama adalah ampicilin dan gentamisin. Sedangkan antibiotik pilihan kedua adalah penisilin prokain dan gentamisin.



## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian tugas akhir ini merupakan data sekunder yang diambil dari data rekam medis bayi baru lahir di Rumah Sakit Umum Daerah (RSUD) Sidoarjo pada bulan September–Oktober 2015. Bayi yang lahir di RSUD Sidoarjo pada periode waktu tersebut sebanyak 222 bayi dan bayi baru lahir yang dirujuk ke RSUD Sidoarjo dalam periode waktu yang sama sebanyak 25 bayi. Sehingga total data rekam medis yang dikumpulkan dalam penelitian ini sebanyak 247 bayi. Data rekam medis yang dikumpulkan merupakan pemeriksaan pertama saat bayi baru lahir atau masih berusia 0 hari untuk bayi yang lahir di RSUD Sidoarjo. Sedangkan pemeriksaan pada bayi yang dirujuk ke RSUD Sidoarjo merupakan pemeriksaan pertama kali saat bayi baru datang di RSUD Sidoarjo. Struktur data yang digunakan dalam penelitian ini ditampilkan dalam Tabel 3.1 berikut.

**Tabel 3.1** Struktur Data

Data ke-	Y	X <sub>1</sub>	X <sub>2</sub>	...	X <sub>10</sub>
1	y <sub>1</sub>	x <sub>1,1</sub>	x <sub>1,2</sub>	...	x <sub>1,10</sub>
2	y <sub>2</sub>	x <sub>2,1</sub>	x <sub>2,2</sub>	...	x <sub>2,10</sub>
⋮	⋮	⋮	⋮	⋮	⋮
247	y <sub>247</sub>	x <sub>247,1</sub>	x <sub>247,2</sub>	...	x <sub>247,10</sub>

### 3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah variabel dependen (respon) dan variabel independen (prediktor). Variabel dependen yang digunakan adalah status risiko infeksi pada bayi dengan kategori risiko infeksi negatif dan positif. Sedangkan variabel prediktor yang digunakan meliputi 10 variabel, yang ditunjukkan pada Tabel 3.2 sebagai berikut.

**Tabel 3.2** Variabel Penelitian

Variabel	Nama Variabel	Kategori dan Koding	Skala Pengukuran
Y	Status Risiko Infeksi	0 : Negatif 1 : Positif	Nominal
X <sub>1</sub>	Berat Badan Lahir (gram)	-	Rasio
X <sub>2</sub>	Bayi Lahir Prematur	0 : Tidak 1 : Ya	Nominal
X <sub>3</sub>	Kondisi Air Ketuban	0 : Normal 1 : Keruh atau Hijau	Nominal
X <sub>4</sub>	Ketuban Pecah Dini (KPD)	0 : Tidak 1 : Ya	Nominal
X <sub>5</sub>		-	Interval
X <sub>6</sub>	Warna Kulit	0 : <i>Blue</i> 1 : <i>All Pink</i>	Nominal
X <sub>7</sub>	Denyut Jantung (kali per menit)	-	Rasio
X <sub>8</sub>	Refleks Rangsangan	0 : Menangis keras 1 : Merintih	Nominal
X <sub>9</sub>	Kecepatan Bernafas	0 : Teratur 1 : Tidak teratur	Nominal
X <sub>10</sub>	Kondisi Tali Pusar	0 : Segar 1 : Layu	Nominal

Klasifikasi variabel respon terdiri dari 2 kategori yaitu bayi berisiko infeksi dan tidak berisiko infeksi. Status risiko infeksi terhadap bayi baru lahir merupakan status dimana bayi menunjukkan gejala-gejala terinfeksi. Bayi yang memiliki risiko infeksi akan diberikan beberapa antibiotik. Pemberian antibiotik bertujuan agar kuman atau bakteri yang menyebabkan infeksi dapat dilawan. Hal tersebut mencegah agar bayi tidak terkena infeksi.

Penentuan variabel-variabel prediktor dilakukan berdasarkan penelitian yang telah dilakukan sebelumnya mengenai risiko infeksi pada bayi baru lahir sebagai berikut.

1. Berat Badan Lahir Bayi

Berat badan lahir bayi yang rendah memiliki kepekaan rangsang yang lebih rendah dibandingkan dengan berat badan bayi normal. Berat badan bayi lahir rendah lebih berisiko terkena infeksi, karena daya tahan tubuh yang masih lemah, kemampuan leukosit masih kurang dan pembentukan antibodi belum sempurna. Berat badan bayi tergolong rendah jika berat badan bayi kurang dari 2500 gram (Carolus, Rompis, & Wilar, 2013).

2. Bayi Lahir Prematur

Bayi lahir prematur (kurang bulan) merupakan bayi yang lahir dengan usia kehamilannya kurang dari 35 minggu. Bayi prematur cenderung memiliki berat badan lahir yang rendah, sehingga bayi lebih berisiko terkena infeksi (Carolus, Rompis, & Wilar, 2013).

3. Kondisi Air Ketuban

Kondisi cairan ketuban keruh atau berwarna hijau disebabkan karena bayi mengeluarkan mekonium (feses janin) sehingga berisiko infeksi pada bayi (Rucy, 2014).

4. Ketuban Pecah Dini (KPD)

Ibu mengalami ketuban pecah dini berpengaruh terhadap janin. Apabila air ketuban berkurang atau mengering, maka hal tersebut dapat menyebabkan bayi berisiko infeksi (Rucy, 2014).

5. Suhu Tubuh Bayi

Jika suhu tubuh

C (hipotermia) merupakan gejala infeksi pada bayi baru lahir (Kementrian Kesehatan RI, 2010).

6. Warna Kulit

Warna kulit merupakan salah satu pemeriksaan untuk menentukan nilai APGAR. Jika seluruh badan biru atau pucat, maka nilai APGAR akan rendah. Nilai APGAR

berpengaruh terhadap risiko infeksi pada bayi. Namun, jika warna kulit keseluruhan merah muda, maka bayi normal (Marmi & Rahardjo, 2012).

7. Denyut Jantung  
Pemeriksaan denyut jantung merupakan salah satu pemeriksaan untuk menentukan nilai APGAR. Denyut jantung bayi normal, jika berdetak lebih dari 100 kali per menit. Namun, jika denyut jantung kurang dari 100 kali per menit, maka akan memberikan nilai yang rendah terhadap APGAR (Marmi & Rahardjo, 2012).
8. Refleks Rangsangan  
Menilai refleks jika bayi diberi rangsangan merupakan salah satu pemeriksaan untuk menentukan nilai APGAR. Jika respon bayi menangis keras terhadap stimulasi, maka bayi mempunyai refleks rangsangan yang normal. Namun jika tidak ada respon atau respon bayi merintih terhadap stimulasi, maka bayi akan memiliki nilai APGAR yang rendah (Marmi & Rahardjo, 2012).
9. Kecepatan Bernafas  
Bayi baru lahir bernafas dengan kecepatan yang berbeda. Jika kecepatan bernafas bayi tidak teratur atau bayi mengalami sesak nafas, maka bayi berisiko infeksi, seperti infeksi sepsis (Marmi & Rahardjo, 2012).
10. Kondisi Tali Puser  
Apabila kondisi tali puser layu, maka tali puser dapat menyebabkan risiko infeksi terhadap bayi. Karena tali puser juga merupakan tempat yang berpotensi untuk pertumbuhan kuman penyebab infeksi (Yefri, Mayetti, & Machmud, 2010).

### 3.3 Langkah Analisis Data

Langkah analisis yang dilakukan dalam penelitian klasifikasi risiko infeksi pada bayi perinatal di RSUD Sidoarjo menggunakan *Classification Trees* adalah sebagai berikut.

1. Melakukan *pre-processing* terhadap data rekam medis bayi yang dikumpulkan untuk meningkatkan akurasi pada hasil

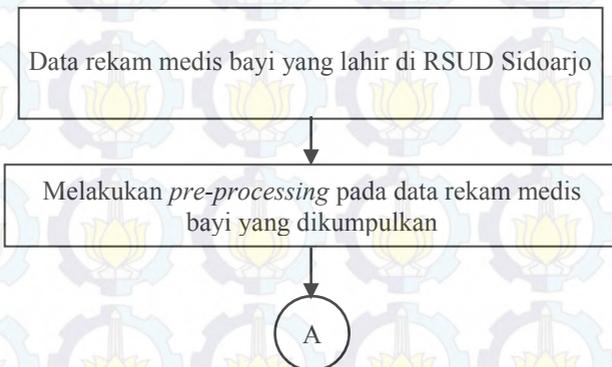
klasifikasi dengan mengatasi *missing value*. Setelah dilakukan penanganan kasus *missing value* dengan aturan apabila jumlah kasus *missing value* kurang dari 30%, maka data yang mengandung *missing* tidak diikuti dalam analisis. Apabila persentase kasus *missing value* lebih dari 30%, maka digunakan metode inputasi pada data yang mengandung *missing*.

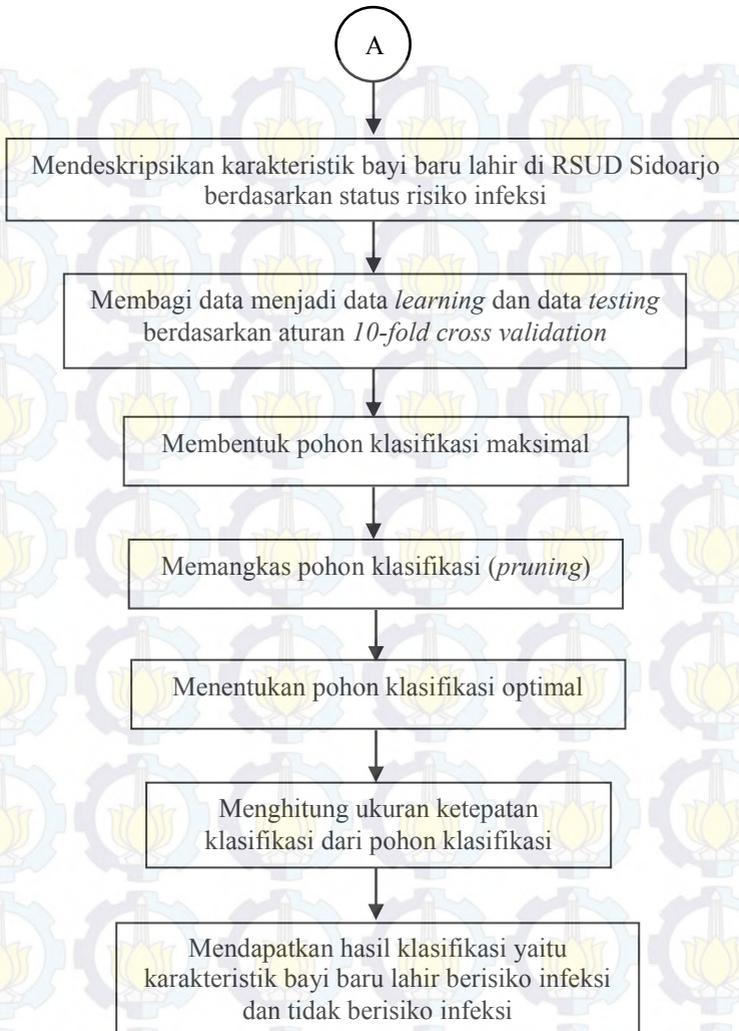
2. Mendeskripsikan karakteristik bayi yang lahir di RSUD Sidoarjo berdasarkan dua tipe variabel respon. Apabila variabel prediktor berskala kategorik, maka dideskripsikan menggunakan *pie-chart* dan *bar-chart*. Sedangkan untuk variabel prediktor yang memiliki skala kontinu dilakukan deskripsi menggunakan rata-rata, deviasi standar, nilai maksimum dan nilai minimum.
3. Membagi data menjadi dua bagian, yaitu data *learning* dan data *testing*. Data dibagi sesuai dengan aturan *10-fold cross validation estimate*, dimana data secara random dibagi menjadi 10 bagian, dengan proporsi 9 bagian dijadikan sebagai data *learning* dan 1 bagian sebagai data *testing*. Karena dalam menentukan pohon optimal menggunakan metode *10-fold cross validation estimate*.
4. Melakukan analisis klasifikasi dengan menggunakan metode *classification trees* menggunakan *software* CART melalui tahapan sebagai berikut.
  - a. Membentuk pohon klasifikasi maksimal dengan menggunakan data *learning* melalui tahapan sebagai berikut.
    1. Melakukan pemilihan pemilah berdasarkan variabel prediktor menurut aturan pemilahan indeks gini yang kemudian hasil pemilahan dievaluasi dengan menggunakan kriteria *goodness of split*.
    2. Menentukan jumlah simpul terminal dengan cara menghentikan pembentukan pohon hingga dicapai batasan minimum pengamatan dan diperoleh simpul terminal yang homogen.

3. Melakukan penandaan label kelas pada simpul terminal berdasarkan aturan jumlah terbanyak dari tiap kelas yang ada pada variabel respon.
  - b. Memangkas pohon klasifikasi (*pruning*) berdasarkan kriteria *cost complexity minimum*.
  - c. Menentukan pohon klasifikasi optimal dengan *10-Fold Cross Validation Estimate*.
  - d. Menghitung ketepatan klasifikasi pohon hasil bentukan dengan menggunakan data *learning* dan melakukan validasi dengan menggunakan data *testing* pada pohon klasifikasi yang terbentuk. Ukuran ketepatan klasifikasi yang digunakan adalah *total accuracy rate* (1-APER), *sensitivity* dan *specificity*.
5. Menarik kesimpulan mengenai perbandingan karakteristik bayi baru lahir berisiko infeksi dan bayi baru lahir tidak berisiko infeksi.

### 3.4 Diagram Alir Analisis Data

Tahapan dalam penelitian berdasarkan langkah metode analisis data akan digambarkan melalui diagram alir penelitian yang disajikan pada Gambar 3.1 berikut.





**Gambar 3.1** Diagram Alir Penelitian



*“Halaman ini sengaja dikosongkan”*

## **BAB IV**

### **ANALISIS DAN PEMBAHASAN**

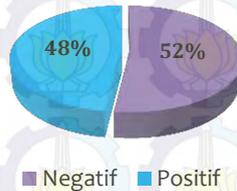
Dalam bab ini membahas tentang analisis data faktor-faktor yang mempengaruhi bayi baru lahir berisiko infeksi di RSUD Sidoarjo untuk menjawab permasalahan serta tujuan dari penelitian yang telah dijelaskan pada bab sebelumnya.

#### **4.1 Karakteristik Bayi Baru Lahir di Rumah Sakit Umum Daerah Sidoarjo**

Bayi yang lahir di RSUD Sidoarjo merupakan hal yang akan diukur dalam penelitian ini, dimana karakteristik-karakteristik dari bayi yang lahir dikumpulkan sebagai data. Kemudian hal yang diperhatikan adalah diagnosa kedua pada bayi tersebut, yakni positif atau negatif risiko infeksi. Status risiko infeksi merupakan variabel respon dalam penelitian ini. Indikator yang mempengaruhi penentuan status risiko infeksi adalah berat badan lahir, lahir prematur, kondisi air ketuban, ketuban pecah dini, suhu tubuh bayi, warna kulit, denyut jantung, refleks rangsangan, kecepatan bernafas dan kondisi tali pusar.

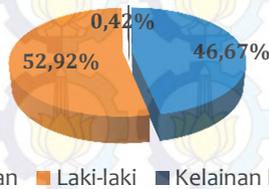
Jumlah bayi yang lahir di RSUD Sidoarjo selama bulan September-Oktober 2015 sebanyak 222 bayi dan jumlah bayi yang dirujuk ke RSUD Sidoarjo pada periode waktu yang sama sebanyak 25 bayi. Sehingga total bayi yang menjadi sampel pengamatan pada penelitian ini adalah 247 bayi. Tampilan beberapa data pengamatan bayi baru lahir disajikan dalam Lampiran A. Berdasarkan jumlah tersebut sebanyak 130 bayi lahir normal atau memiliki status negatif risiko infeksi, sedangkan 117 bayi lahir dengan status positif risiko infeksi. Berdasarkan indikator yang diperoleh terdapat beberapa kasus *missing value* pada dua indikator, yakni suhu tubuh bayi ( $X_5$ ) dan denyut jantung ( $X_7$ ) dengan jumlah kasus masing-masing sebanyak 1 dan 6 kasus. Karena jumlah *missing value* yang terdapat dalam indikator sebanyak 2,43% kurang dari 30%, maka diputuskan untuk menghapus data yang mengandung *missing value* pada indikator suhu tubuh bayi dan denyut jantung dihilangkan dalam

analisis guna memberikan informasi yang lebih akurat dan dapat menghasilkan performansi yang lebih baik. Data yang dihapuskan sebanyak 7 data, yakni 5 data termasuk dalam kategori status negatif dan 2 data termasuk dalam kategori status positif berisiko infeksi. Sehingga jumlah data yang tersisa sebanyak 240 bayi, dimana 125 bayi negatif risiko infeksi dan 115 bayi positif risiko infeksi. Berdasarkan data pada Lampiran A, maka dilakukan rangkuman deskripsi status risiko infeksi pada bayi baru lahir di RSUD Sidoarjo yang disajikan dalam Gambar 4.1 sebagai berikut.



**Gambar 4.1** Deskripsi Status Risiko Infeksi pada Bayi

Berdasarkan Gambar 4.1 dapat diperoleh informasi bahwa bayi yang lahir di RSUD Sidoarjo pada bulan September-Oktober 2015 lebih banyak yang tidak memiliki risiko infeksi daripada bayi yang berstatus positif risiko infeksi, yakni sebanyak 52%. Bayi baru lahir berdasarkan jenis kelamin ditampilkan dalam Gambar 4.2.



**Gambar 4.2** Deskripsi Bayi Baru Lahir Berdasarkan Jenis Kelamin

Gambar 4.2 menampilkan bahwa persentase bayi yang berjenis kelamin perempuan sebesar 46,67%, sedangkan bayi berjenis kelamin laki-laki memiliki persentase sebesar 52,92%.

Dan terdapat 1 bayi (atau sebesar 0,42%) yang memiliki kelainan hemaprodit. Jika jenis kelamin bayi baru lahir yang diamati ditampilkan berdasarkan status risiko infeksi dapat dilihat dalam Tabel 4.1 sebagai berikut.

**Tabel 4.1** Dekripsi Jenis Kelamin Berdasarkan Status Risiko Infeksi

Jenis Kelamin	Status Risiko Infeksi		Total
	Negatif	Positif	
Laki-laki	62	65	127
Perempuan	63	49	112
Hemaprodit	0	1	1
<b>Total</b>	125	115	240

Tabel 4.1 memberikan informasi bahwa sebanyak 125 bayi negatif risiko infeksi, 62 bayi diantaranya berjenis kelamin laki-laki dan 63 bayi berjenis kelamin perempuan. Sedangkan jumlah bayi yang positif risiko infeksi sebanyak 115 bayi, dimana 65 bayi diantaranya berjenis kelamin laki-laki, 49 bayi berjenis kelamin perempuan dan 1 bayi memiliki kelainan hemaprodit.

Selain itu, informasi mengenai karakteristik bayi baru lahir berdasarkan masing-masing indikator yang mempengaruhi dalam pemeriksaan dan diagnosa bayi positif risiko infeksi disajikan secara lengkap pada Lampiran C dan ditampilkan dengan gambaran deskriptif sebagai berikut.

#### 4.1.1 Karakteristik Bayi Baru Lahir Berdasarkan Berat Badan Lahir

Karakteristik bayi baru lahir di RSUD Sidoarjo berdasarkan berat badan sewaktu lahir dapat dijelaskan dengan tampilan Tabel 4.2 sebagai berikut.

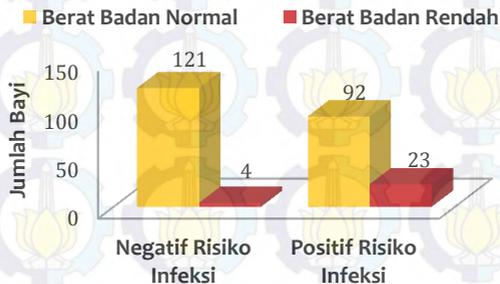
**Tabel 4.2** Deskripsi Bayi Baru Lahir Berdasarkan BBL

Status Risiko Infeksi	Rata- Rata (gram)	Deviasi Standar	Maksimum (gram)	Minimum (gram)
Negatif	3186,56	432,76	4100	1800
Positif	3019,83	641,36	4900	1400

Tabel 4.2 memberikan informasi mengenai bayi baru lahir berdasarkan indikator berat badan lahir. Rata-rata bayi baru lahir yang berstatus negatif risiko infeksi memiliki berat badan saat lahir sebesar 3186,56 gram. Sedangkan bayi baru lahir dengan risiko infeksi memiliki rata-rata berat badan sewaktu lahir sebesar 3019,83 gram. Artinya, bayi dengan risiko infeksi memiliki rata-rata berat badan lahir dibandingkan dengan bayi yang negatif risiko infeksi. Sementara itu deviasi standar untuk masing-masing status bayi terhadap risiko infeksi sebesar 432,76 dan 641,36. Deviasi standar dari tiap status bayi terhadap risiko infeksi cukup besar, artinya variasi berat badan lahir untuk masing-masing status risiko infeksi cukup beragam.

Bayi baru lahir yang tidak memiliki risiko infeksi paling besar mempunyai berat badan lahir sebesar 4100 gram, sedangkan paling minimum memiliki berat badan lahir sebesar 1800 gram. Artinya, ada bayi dengan berat badan lahir rendah atau BBLR, karena berat badannya kurang dari 2500 gram. Untuk bayi baru lahir yang berstatus positif risiko infeksi mempunyai berat badan lahir maksimum sebesar 4900 gram dan paling minimum sebesar 1400 gram. Artinya, bayi yang memiliki risiko infeksi ada yang termasuk bayi besar dan bayi dengan BBLR.

Berikut akan ditampilkan deskriptif bayi baru lahir berdasarkan status berat badan lahir, yakni berat normal atau BBLR. Bayi yang memiliki berat badan lahir rendah merupakan bayi yang memiliki berat badan lahir kurang dari 2500 gram.

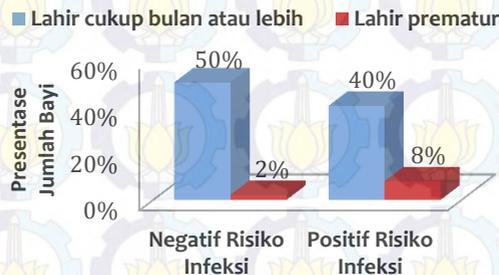


**Gambar 4.3** Deskripsi BBL Berdasarkan Status Risiko Infeksi

Gambar 4.3 menunjukkan informasi bahwa dari 125 bayi yang tidak memiliki risiko infeksi, 121 bayi diantaranya memiliki berat badan lahir normal, sedangkan 4 bayi sisanya memiliki berat badan lahir rendah. Sementara itu, jumlah bayi yang positif risiko infeksi sebanyak 115 bayi, dimana 92 bayi diantaranya memiliki berat badan lahir normal dan 23 bayi memiliki berat badan lahir rendah. Berdasarkan pemaparan tersebut menunjukkan informasi bahwa sebagian besar bayi lahir dengan berat badan normal baik yang tidak memiliki risiko infeksi maupun bayi yang positif risiko infeksi.

#### 4.1.2 Karakteristik Bayi Baru Lahir Berdasarkan Status Lahir Prematur

Pada bayi baru lahir juga dicatat usia kehamilan dari ibu. apabila ibu melalui proses persalinan pada usia kehamilan kurang dari sama dengan 35 minggu, maka bayi yang lahir termasuk bayi yang lahir prematur (belum cukup bulan). Kemudian berdasarkan indikator lahir prematur dapat diketahui pula deskripsi bayi baru lahir untuk kedua status risiko infeksi yang disajikan pada Gambar 4.4 sebagai berikut.



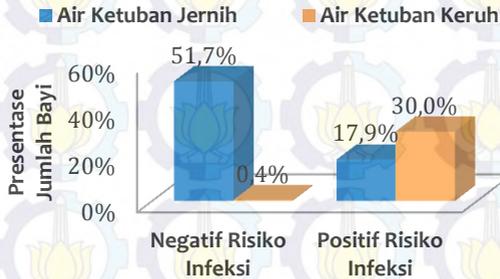
**Gambar 4.4** Dekripsi Status Lahir Prematur Berdasarkan Status Risiko Infeksi

Berdasarkan Gambar 4.4 dapat diketahui bahwa dari total 125 bayi yang tidak memiliki risiko infeksi sebesar 50% diantaranya bayi lahir cukup bulan dan 2% bayi lahir prematur. Sedangkan dari 115 bayi yang positif risiko infeksi sebesar 40% diantaranya bayi lahir cukup bulan dan 8% bayi lahir prematur.

Sehingga dapat disimpulkan bahwa sebagian besar bayi yang lahir di RSUD Sidoarjo pada bulan September -Oktober 2015 lahir cukup bulan, baik pada bayi baru lahir dengan status negatif maupun status positif terhadap penyakit risiko infeksi.

#### 4.1.3 Karakteristik Bayi Baru Lahir Berdasarkan Kondisi Air Ketuban

Kondisi dari sisa air ketuban setelah bayi dilahirkan menjadi perhatian bagi perawat, dokter atau bidan, apakah sisa air ketuban jernih atau keruh. Untuk mengetahui perbandingan kondisi sisa air ketuban dan status risiko infeksi dari bayi baru lahir di RSUD Sidoarjo disajikan pada Gambar 4.5 sebagai berikut.

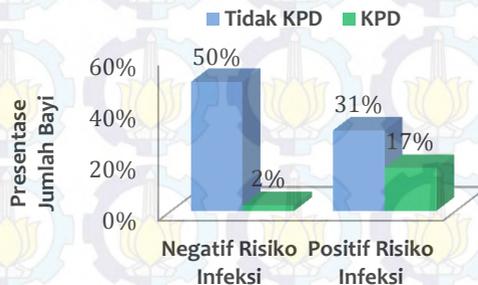


**Gambar 4.5** Dekripsi Kondisi Air Ketuban Berdasarkan Status Risiko Infeksi

Gambar 4.5 memberikan informasi bahwa dari total 125 bayi berstatus negatif risiko infeksi, sebesar 51,7% diantaranya bayi memiliki sisa air ketuban dalam kondisi jernih dan 0,4% bayi memiliki sisa air ketuban dalam kondisi keruh. Sedangkan sari 115 bayi berstatus positif risiko infeksi, 17,9% diantaranya bayi memiliki kondisi sisa air ketuban jernih dan 30% bayi memiliki kondisi sisa air ketuban keruh. Sehingga dapat disimpulkan bahwa sebagian besar bayi baru lahir yang tidak berisiko infeksi memiliki sisa air ketuban dalam kondisi jernih, sedangkan bayi baru lahir yang berisiko infeksi didominasi oleh bayi yang memiliki sisa air ketuban keruh.

#### 4.1.4 Karakteristik Bayi Baru Lahir Berdasarkan Status Ketuban Pecah Dini

Ketuban pecah dini (KPD) merupakan peristiwa dimana ketuban pecah lebih dahulu sebelum proses persalinan. Ketuban pecah dini merupakan salah satu faktor yang mempengaruhi dalam diagnosa risiko infeksi pada bayi baru lahir. Untuk mengetahui perbandingan antara status ketuban pecah dini dan status risiko infeksi diperlukan tabel kontingensi seperti yang disajikan pada Gambar 4.6 sebagai berikut.



**Gambar 4.6** Dekripsi Status KPD Berdasarkan Status Risiko Infeksi

Berdasarkan Gambar 4.6 dapat diperoleh informasi bahwa dari total 125 bayi baru lahir di RSUD Sidoarjo yang tidak memiliki risiko infeksi, 50% bayi diantaranya tidak mengalami ketuban pecah dini, sementara 2% bayi dilahirkan dengan kondisi ketuban pecah dini. Sedangkan dari total 115 bayi yang berisiko infeksi, 31% bayi diantaranya tidak mengalami ketuban pecah dini dan 17% bayi dilahirkan dengan kondisi ketuban pecah dini. Hal tersebut menunjukkan bahwa sebagian besar bayi yang lahir di RSUD Sidoarjo tidak mengalami ketuban pecah dini, baik bayi yang tidak memiliki risiko infeksi maupun yang memiliki risiko infeksi.

#### 4.1.5 Karakteristik Bayi Baru Lahir Berdasarkan Suhu Tubuh Bayi

Indikator suhu tubuh bayi sewaktu lahir berdasarkan masing-masing status risiko infeksi ditampilkan dalam Tabel 4.3 sebagai berikut.

**Tabel 4.3** Deskripsi Bayi Baru Lahir Berdasarkan Suhu Tubuh

<b>Status Risiko Infeksi</b>	<b>Rata-Rata (°C)</b>	<b>Deviasi Standar</b>	<b>Maksimum (°C)</b>	<b>Minimum (°C)</b>
Negatif	36,68	0,385	38,9	35,5
Positif	36,74	0,571	39,5	35

Berdasarkan Tabel 4.3 dapat diketahui bahwa rata-rata suhu tubuh saat lahir untuk bayi yang tidak memiliki risiko infeksi adalah 36,68. Nilai deviasi standar yang cukup kecil menunjukkan bahwa suhu tubuh bayi yang tidak memiliki risiko infeksi kurang beragam. Sedangkan rata-rata suhu tubuh sewaktu lahir untuk bayi yang memiliki risiko infeksi adalah 36,74. Nilai deviasi standar sebesar 0,571. Nilai deviasi standar dari suhu tubuh bayi yang memiliki risiko infeksi juga cukup kecil yang menunjukkan bahwa suhu tubuh yang dimiliki bayi dengan risiko infeksi kurang beragam.

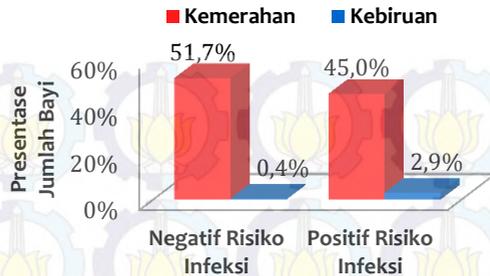
Nilai maksimum dari suhu

Untuk bayi yang memiliki risiko

bahwa ada bayi baru lahir dengan risiko infeksi menderita demam dan juga ada yang menderita hipotermia ring

#### **4.1.6 Karakteristik Bayi Baru Lahir Berdasarkan Warna Kulit**

Pemeriksaan pada bayi baru lahir selanjutnya adalah pemeriksaan warna kulit. Bayi dikatakan normal, jika warna kulitnya adalah kemerahan atau merah muda. Sedangkan apabila kulit bayi kebiruan, perlu adanya pemeriksaan lebih lanjut terhadap bayi tersebut. Kemudian akan dilakukan perbandingan antara indikator warna kulit bayi dengan status risiko infeksi pada Gambar 4.7 sebagai berikut.



**Gambar 4.7** Dekripsi Warna Kulit Berdasarkan Status Risiko Infeksi

Gambar 4.7 memberikan informasi bahwa dari total keseluruhan 125 bayi yang lahir di RSUD Sidoarjo dengan status risiko infeksi negatif, sebesar 51,7% bayi diantaranya memiliki warna kulit merah muda atau kemerahan dan hanya 0,4% bayi yang memiliki warna kulit kebiruan saat lahir. Jumlah bayi yang berisiko infeksi sebanyak 115 bayi dengan 45% bayi diantaranya lahir dengan warna kulit kemerahan dan 2,9% bayi memiliki warna kulit kebiruan. Sehingga dapat disimpulkan bahwa sebagian besar bayi yang lahir memiliki warna kulit kemerahan, baik bayi dengan status risiko negatif maupun positif.

#### 4.1.7 Karakteristik Bayi Baru Lahir Berdasarkan Denyut Jantung

Bayi baru lahir melewati pemeriksaan denyut jantung yang merupakan salah satu faktor yang mempengaruhi dalam diagnosa risiko infeksi. Berikut ditampilkan dalam Tabel 4.4 mengenai karakteristik bayi baru lahir berdasarkan indikator denyut jantung.

**Tabel 4.4** Deskripsi Bayi Baru Lahir Berdasarkan Denyut Jantung

Status Risiko Infeksi	Rata-Rata (kali/menit)	Deviasi Standar	Maksimum (kali/menit)	Minimum (kali/menit)
Negatif	128,48	9,0959	146	110
Positif	128,64	10,149	160	120

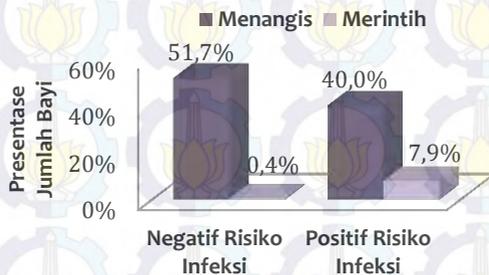
Berdasarkan Tabel 4.4 diperoleh informasi bahwa rata-rata denyut jantung bayi yang tidak memiliki risiko infeksi adalah 128,48 kali per menit dengan deviasi standar sebesar 9,0959.

Denyut jantung yang dimiliki cukup beragam oleh bayi yang berstatus negatif risiko infeksi, karena nilai deviasi standar cukup besar. Sedangkan, bayi baru lahir dengan risiko infeksi memiliki rata-rata denyut jantung sebesar 128,64 kali per menit dengan deviasi standar sebesar 10,149. Nilai deviasi standar yang dimiliki oleh bayi yang positif risiko infeksi cukup besar, sehingga denyut jantung bayi-bayi tersebut cukup beragam.

Tabel 4.4 juga memberikan informasi mengenai denyut jantung yang paling maksimum yang dimiliki oleh bayi berstatus negatif terhadap risiko infeksi sebesar 146 kali per menit dan paling minimum sebesar 110 kali per menit. Sedangkan bayi yang memiliki risiko infeksi menunjukkan denyut jantung maksimumnya sebesar 160 kali per menit dan paling minimum sebesar 120 kali per menit. Artinya, seluruh bayi memiliki denyut jantung normal.

#### 4.1.8 Karakteristik Bayi Baru Lahir Berdasarkan Refleksi Rangsangan

Indikator refleksi rangsangan merupakan salah satu faktor yang mempengaruhi klasifikasi risiko infeksi pada bayi baru lahir. Bayi baru lahir diberi stimulus dan dilihat reaksi secara refleksi dari bayi. Selanjutnya dilakukan perbandingan antara refleksi rangsangan bayi dan status risiko infeksi yang disajikan pada Gambar 4.8 sebagai berikut.



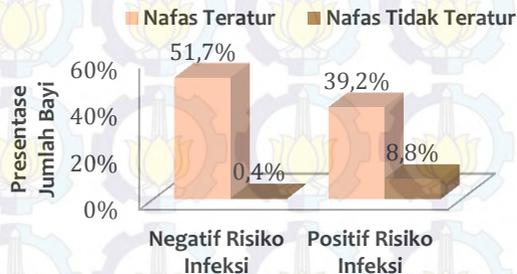
Gambar 4.8 Dekripsi Refleksi Rangsangan Berdasarkan Status Risiko Infeksi

Gambar 4.8 memberikan informasi bahwa dari total 125 bayi baru lahir yang tidak memiliki risiko infeksi, sebesar 51,7%

bayi diantaranya memberikan refleks menangis setelah diberi stimula dan 0,4% bayi memberikan refleks merintih. Sedangkan dari 115 bayi yang memiliki risiko infeksi, 40% bayi diantaranya memberikan refleks menangis saat diberi stimulan dan 7,9% bayi memberikan refleks merintih. Sehingga dapat disimpulkan bahwa sebagian besar bayi memberikan respon menangis keras, baik bayi yang tidak memiliki risiko infeksi maupun yang memiliki risiko infeksi.

#### 4.1.9 Karakteristik Bayi Baru Lahir Berdasarkan Kecepatan Bernafas

Kecepatan bernafas merupakan pemeriksaan terhadap usaha bernafas dari bayi baru lahir, apakah pernafasan bayi teratur atau tidak. Indikator kecepatan bernafas dari bayi baru lahir dapat diketahui pula perbandingan indikator tersebut dengan status risiko infeksi dalam Gambar 4.9 sebagai berikut.

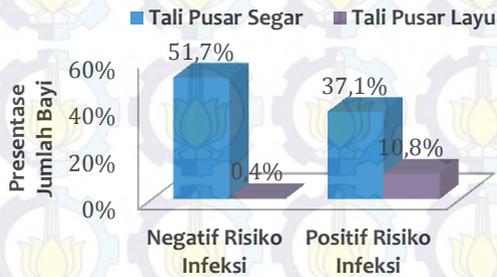


**Gambar 4.9** Dekripsi Kecepatan Bernafas Berdasarkan Status Risiko Infeksi

Gambar 4.9 menunjukkan bahwa dari total 125 bayi yang tidak berisiko infeksi, sebanyak 51,7% bayi diantaranya memiliki kecepatan bernafas teratur dan 0,4% bayi memiliki kecepatan bernafas tidak teratur. Begitu juga dengan total 115 bayi baru lahir yang berisiko infeksi dengan 39,2% bayi diantaranya memiliki kecepatan bernafas teratur dan 8,8% bayi memiliki kecepatan bernafas tidak teratur.

#### 4.1.10 Karakteristik Bayi Baru Lahir Berdasarkan Kondisi Tali Pesar

Pemeriksaan kondisi tali pesar pada bayi baru lahir termasuk dalam asuhan keperawatan. Tali pesar bayi baru lahir diperiksa apakah dalam kondisi segar atau layu. Kemudian untuk mengetahui hubungan antara kondisi tali pesar dan status risiko infeksi pada bayi baru lahir disajikan tabel kontingensi antara kedua variabel tersebut dalam Gambar 4.10 sebagai berikut.



**Gambar 4.10** Dekripsi Kondisi Tali Pesar Berdasarkan Status Risiko Infeksi

Gambar 4.10 menyajikan informasi bahwa sebanyak 125 bayi baru tidak berisiko infeksi, dimana 51,7% bayi diantaranya memiliki tali pesar dalam kondisi segar dan 0,4% bayi memiliki tali pesar dalam kondisi layu. Sedangkan jumlah bayi baru lahir yang berisiko infeksi sebanyak 115 bayi, dengan 37,1% bayi diantaranya memiliki tali pesar dalam kondisi segar dan 10,8% bayi memiliki tali pesar dalam kondisi layu. Hal tersebut menunjukkan bahwa sebagian besar bayi baru lahir di RSUD Sidoarjo memiliki tali pesar dalam kondisi segar, baik bayi yang didiagnosa tidak berisiko infeksi maupun yang berisiko infeksi.

#### 4.2 Analisis *Classification Trees* untuk Klasifikasi Risiko Infeksi pada Bayi Baru Lahir di RSUD Sidoarjo

Berikut ini adalah hasil analisis klasifikasi risiko infeksi pada bayi baru lahir di Rumah Sakit Umum Daerah (RSUD) Sidoarjo menggunakan metode *Classification Trees* dengan jumlah data sebanyak 240. Data bayi baru lahir di RSUD Sidoarjo

yang digunakan dalam analisis merupakan data hasil *pre-processing* dengan menghilangkan kasus *missing value*. Analisis menggunakan *Classification Trees* bertujuan untuk mengetahui ketepatan klasifikasi yang telah dihasilkan berdasarkan status risiko infeksi pada bayi baru lahir melalui pohon klasifikasi maksimal dan pohon klasifikasi optimal dengan proses pemilahan menggunakan fungsi keheterogenan indeks gini. Jumlah data yang digunakan dalam analisis *Classification Trees* sejumlah 200 data, dimana 40 data sisanya akan digunakan sebagai data bayi baru untuk diklasifikasikan menggunakan pohon klasifikasi yang telah dihasilkan. Berdasarkan data pengamatan yang digunakan dalam analisis tersebut terdapat 101 kasus yang dikategorikan sebagai bayi negatif risiko infeksi, sedangkan sisanya 99 kasus dikategorikan sebagai bayi positif risiko infeksi. Karena jumlah data sampel pengamatan berukuran kecil (kurang dari 3000), maka penentuan pohon klasifikasi optimal menggunakan metode *v-fold cross validation estimate*. Jumlah *fold* yang digunakan sebanyak 10 *fold* ( $v=10$ ), sehingga data akan dibagi menjadi 10 bagian, dimana masing-masing bagian berjumlah 20 data. Namun, sebelum pembentukan pohon klasifikasi akan ditentukan terlebih dahulu kemungkinan pemilah untuk masing-masing variabel prediktor.

#### **4.2.1 Kemungkinan Pemilah untuk Setiap Variabel Prediktor**

Dalam tahap pembentukan pohon klasifikasi diperlukan variabel terpenting yang dijadikan sebagai pemilah terbaik. Setiap variabel memiliki *threshold* untuk menentukan *goodness of split*. Variabel yang berperan sebagai pemilah dipilih melalui beberapa kemungkinan pemilah dari setiap variabel prediktor berdasarkan skala data pada variabel prediktor tersebut. Apabila skala pengukuran pada variabel prediktor merupakan skala kontinyu, maka kemungkinan pemilah pada variabel prediktor tersebut diperoleh dari median yang dihitung pada tiap dua nilai amatan sampel yang berbeda secara berurutan. Kemungkinan pemilah pada variabel prediktor yang memiliki skala kontinyu ditampilkan

dalam Lampiran B. Perhitungan beberapa kemungkinan pemilah dari setiap variabel prediktor ditunjukkan dalam Tabel 4.5 sebagai berikut.

**Tabel 4.5** Kemungkinan Pemilah Setiap Variabel Prediktor

Variabel	Nama Variabel	Skala Data	Jumlah Kategori/ Nilai Amatan Sampel	Kemungkinan Pemilah
$X_1$	Berat Badan Lahir	Rasio	53	$53-1 = 52$ pemilah
$X_2$	Lahir Prematur	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_3$	Kondisi Air Ketuban	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_4$	Ketuban Pecah Dini	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_5$	Suhu Tubuh	Interval	21	$21-1 = 20$ pemilah
$X_6$	Warna Kulit	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_7$	Denyut Jantung	Rasio	19	$19-1 = 18$ pemilah
$X_8$	Refleks Rangsangan	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_9$	Kecepatan Bernafas	Nominal	2	$2^{2-1}-1 = 1$ pemilah
$X_{10}$	Kondisi Tali Pusing	Nominal	2	$2^{2-1}-1 = 1$ pemilah

Tabel 4.5 menunjukkan banyaknya kemungkinan pemilah pada masing-masing variabel prediktor. Kemungkinan pemilah yang dihasilkan untuk variabel  $X_1$  (berat badan lahir) sebanyak 52 pemilah. Untuk variabel  $X_2$  (lahir prematur), variabel  $X_3$  (kondisi air ketuban), variabel  $X_4$  (ketuban pecah dini), variabel  $X_6$  (warna kulit), variabel  $X_8$  (refleks rangsangan), variabel  $X_9$  (kecepatan bernafas) dan variabel  $X_{10}$  (kondisi tali pusing) memiliki jumlah kemungkinan pemilah yang sama yaitu sebanyak 1 pemilah. Jika dilihat variabel  $X_5$  (suhu tubuh) menghasilkan kemungkinan

pemilah sebanyak 20 pemilah. Sedangkan, untuk variabel  $X_7$  (denyut jantung) kemungkinan pemilah yang dihasilkan sebanyak 18 pemilah.

#### 4.2.2 Pembentukan Pohon Klasifikasi Maksimal

Pembentukan pohon klasifikasi maksimal merupakan tahap pertama dari metode *Classification Trees*. Dalam tahap ini diawali dengan pemilihan *classifier* atau pemilah yakni dengan memilih variabel prediktor terpenting yang akan digunakan sebagai pemilah awal. Berdasarkan variabel prediktor yang digunakan dalam penelitian ini, proses pemilihan pemilah dilakukan dengan mengacu pada skor kontribusi terbesar yang dihasilkan dari tiap variabel prediktor. Besar kontribusi yang dihasilkan oleh masing-masing variabel prediktor dapat dilihat dalam Tabel 4.6 sebagai berikut.

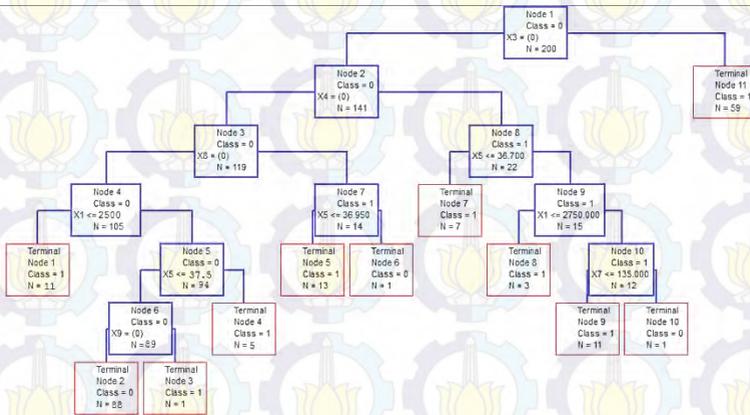
**Tabel 4.6** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal

Variabel	Nama Variabel	Skor
$X_3$	Kondisi Air Ketuban	100,000
$X_9$	Kecepatan Bernafas	46,110
$X_8$	Refleks Rangsangan	41,599
$X_4$	Ketuban Pecah Dini	35,237
$X_{10}$	Kondisi Tali Puser	29,357
$X_1$	Berat Badan Lahir	25,813
$X_5$	Suhu Tubuh	23,686
$X_7$	Denyut Jantung	16,557
$X_2$	Lahir Prematur	11,336
$X_6$	Warna Kulit	6,917

Berdasarkan Tabel 4.6 dapat diketahui bahwa variabel terpenting yang memiliki kontribusi terbesar adalah variabel  $X_3$  (kondisi air ketuban) dengan skor 100. Artinya, variabel  $X_3$  akan menjadi pemilah awal atau biasa disebut sebagai simpul induk (*parent nodes*). Faktor kondisi air ketuban mampu menurunkan tingkat keheterogenan paling besar, sehingga simpul yang dihasilkan akan lebih homogen. Variabel  $X_9$  (kecepatan bernafas) mempunyai skor kontribusi sebesar 41,599 dan untuk skor

kontribusi yang diberikan oleh variabel yang lain selengkapnya dapat dilihat dalam Tabel 4.6.

Selanjutnya dilakukan proses pemilahan dengan fungsi keheterogenan indeks gini hingga terbentuk pohon klasifikasi maksimal. Pohon klasifikasi maksimal menghasilkan simpul terminal dengan jumlah paling maksimal atau terbanyak dan tingkat kedalaman tertinggi. Jumlah simpul terminal yang dihasilkan oleh pohon klasifikasi maksimal adalah sebanyak 11 simpul dengan tingkat kedalaman sebesar 7. Topologi dari pohon klasifikasi maksimal ditampilkan dalam Gambar 4.11 sebagai berikut. Sedangkan pohon klasifikasi maksimal secara lengkap ditampilkan dalam Lampiran H.

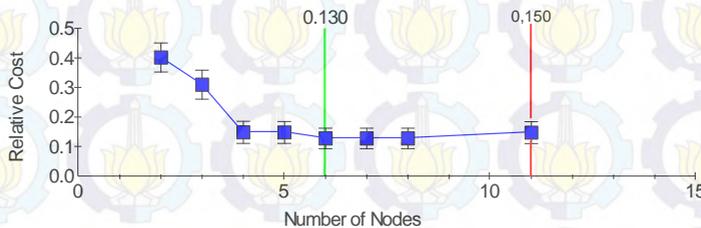


Gambar 4.11 Topologi Pohon Klasifikasi Maksimal

### 4.2.3 Pemangkasan Pohon Klasifikasi Maksimal

Setelah memperoleh pohon klasifikasi maksimal, langkah berikutnya adalah melihat apakah pohon klasifikasi maksimal tersebut perlu dilakukan proses pemangkasan (*pruning*) atau tidak. Tahap pemangkasan bertujuan untuk menghindari adanya kasus *overfitting* (akurasi yang dihasilkan melebihi kenyataan) atau *underfitting* (nilai akurasi yang dihasilkan kurang jika dibandingkan dengan nilai aktual). Proses pemangkasan pohon dilakukan berdasarkan pendekatan *cost complexity minimum* dan

dalam menentukan pohon klasifikasi optimal menggunakan metode *10-fold cross validation estimate*. Dalam hal ini data yang digunakan dibagi menjadi 10 bagian, dimana 9 bagian sebagai data *learning* dan 1 bagian sebagai data *testing*. Metode *cross validation* ini memberikan kesempatan bagi seluruh data untuk menjadi *learning* dan *testing*. Data *learning* digunakan untuk membangun pohon klasifikasi, sedangkan data *testing* digunakan untuk mengetahui pendugaan kesalahan klasifikasi atau mengestimasi *missclassification error rate*. Setiap hasil pemangkasan yang dilakukan memiliki nilai *relative cost* tertentu, kemudian dipilih hasil pemangkasan yang mempunyai nilai *relative cost* paling minimum. Plot *relative cost* dari setiap pemangkasan yang dilakukan pada pohon klasifikasi risiko infeksi pada bayi baru lahir di RSUD Sidoarjo yang ditampilkan dalam Gambar 4.12 sebagai berikut.



**Gambar 4.12** Nilai *Relative Cost* Hasil Pemangkasan Pohon Klasifikasi Maksimal

Gambar 4.12 menjelaskan bahwa pohon klasifikasi maksimal yang terbentuk dengan jumlah simpul terminal sebanyak 11 simpul, perlu dilakukan pemangkasan seperti yang ditunjukkan oleh garis berwarna merah dengan nilai *relative cost* sebesar 0,150. Sedangkan garis berwarna hijau dalam Gambar 4.10 menunjukkan bahwa pohon klasifikasi tersebut telah optimal dengan nilai *relative cost* paling minimum yaitu sebesar 0,130. Hasil pemangkasan yang keempat tersebut menghasilkan pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 6 simpul. Urutan mulai dari pohon klasifikasi maksimal yang

dibentuk hingga pemangkasan yang dilakukan terhadap pohon tersebut dapat dilihat dalam Tabel 4.7 berikut.

Tabel 4.7 Urutan Pohon Klasifikasi

<i>Tree Number</i>	<i>Terminal Nodes</i>	<i>Cross-Validated Relative Cost</i>	<i>Resubstitution Relative Cost</i>	<i>Complexity Parameter</i>
1	11	0,150 ± 0,037	0,070	0,000
2	8	0,130 ± 0,035	0,080	0,002
3	7	0,130 ± 0,035	0,090	0,005
4**	6	0,130 ± 0,035	0,100	0,005
5	5	0,150 ± 0,037	0,110	0,005
6	4	0,151 ± 0,037	0,140	0,015
7	3	0,312 ± 0,049	0,262	0,061
8	2	0,404 ± 0,049	0,404	0,071
9	1	1,000 ± 0,000	1,000	0,298

\*\*pohon klasifikasi optimal

Tabel 4.7 menunjukkan biaya kesalahan (*cross-validated relative cost*) yang dihasilkan oleh setiap hasil pemangkasan pohon klasifikasi maksimal. Pohon pertama merupakan pembentukan pohon klasifikasi maksimal dengan jumlah simpul terminal sebanyak 11 simpul dan biaya kesalahan berkisar antara 0,150 dan 0,037 dengan nilai kompleksitas sebesar 0,000. Sedangkan pohon keempat merupakan pohon klasifikasi yang menghasilkan biaya kesalahan paling minimum yakni berkisar antara 0,130 dan 0,035 dan nilai kompleksitas sebesar 0,005.

#### 4.2.4 Penentuan Pohon Klasifikasi Optimal

Setelah memperoleh hasil pemangkasan pohon klasifikasi secara iteratif, langkah berikutnya adalah menentukan pohon klasifikasi optimal untuk mengetahui akurasi hasil klasifikasi risiko infeksi pada bayi baru lahir di RSUD Sidoarjo. Sebelum itu perlu dilakukan penentuan terhadap variabel terpenting yang memberikan kontribusi terbesar dalam pembentukan pohon klasifikasi. Berikut ini akan ditunjukkan variabel terpenting dalam pembentukan pohon klasifikasi optimal dalam Tabel 4.8.

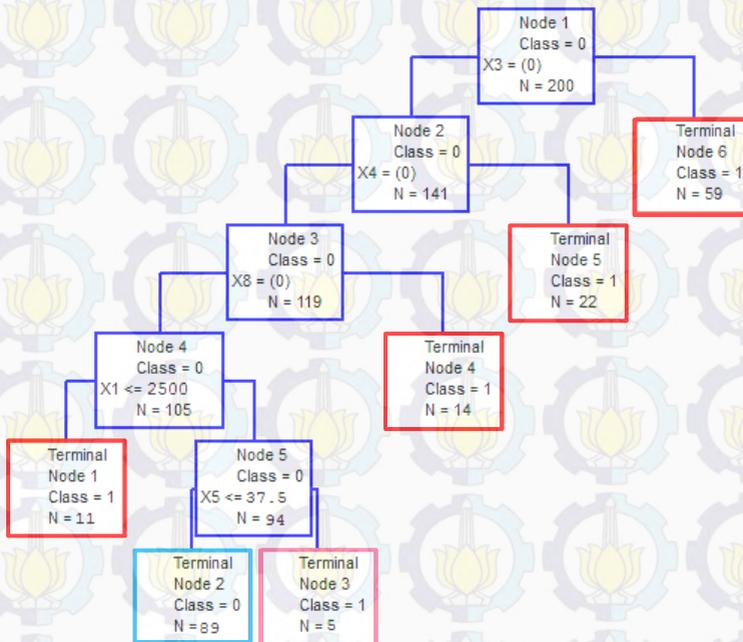
**Tabel 4.8** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal

Variabel	Nama Variabel	Skor
X <sub>3</sub>	Kondisi Air Ketuban	100,000
X <sub>9</sub>	Kecepatan Bernafas	41,599
X <sub>8</sub>	Refleks Rangsangan	41,599
X <sub>4</sub>	Ketuban Pecah Dini	35,237
X <sub>10</sub>	Kondisi Tali Puser	29,357
X <sub>1</sub>	Berat Badan Lahir	24,050
X <sub>5</sub>	Suhu Tubuh	17,775
X <sub>2</sub>	Lahir Prematur	10,102
X <sub>7</sub>	Denyut Jantung	9,664
X <sub>6</sub>	Warna Kulit	6,582

Berdasarkan Tabel 4.8 dapat diketahui bahwa variabel X<sub>3</sub> (kondisi air ketuban) merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal dengan skor kontribusi sebesar 100. Sedangkan variabel-variabel prediktor yang lain mempunyai skor kontribusi kurang dari 50. Proses pemilahan untuk membentuk pohon klasifikasi optimal dengan fungsi keheterogenan indeks gini menghasilkan pohon klasifikasi dengan simpul terminal sebanyak 6 simpul dan tingkat kedalaman sebesar 6. Berikut ini ditampilkan topologi pohon klasifikasi optimal berdasarkan faktor-faktor yang mempengaruhi risiko infeksi pada bayi baru lahir dalam Gambar 4.12. Selain itu, pohon klasifikasi optimal secara lengkap ditampilkan dalam Lampiran I.

Topologi pohon klasifikasi optimal yang ditunjukkan dalam Gambar 4.13 dapat menjelaskan bahwa variabel yang menentukan klasifikasi risiko infeksi pada bayi baru lahir adalah variabel X<sub>3</sub> (kondisi air ketuban) seperti yang ditunjukkan dalam Tabel 4.8. Kemudian simpul induk memilah menjadi dua simpul anak (*child nodes*), yaitu simpul kiri dan simpul kanan berdasarkan kondisi air ketuban (X<sub>3</sub>) saat bayi tersebut dilahirkan. Simpul kiri dilabelkan sebagai kelas 0 (negatif risiko infeksi) dan simpul kanan dilabelkan sebagai kelas 1 (positif risiko infeksi). Bayi baru lahir akan diklasifikasikan masuk ke dalam simpul kiri jika sisa air ketuban yang dimiliki dalam kondisi jernih.

Sebaliknya, jika sisa air ketuban yang dimiliki saat bayi tersebut lahir dalam kondisi keruh atau berwarna hijau, maka bayi tersebut akan diklasifikasikan ke dalam simpul kanan atau dapat dikatakan bahwa bayi tersebut positif berisiko infeksi. Pemilahan tersebut terjadi secara berulang sampai tidak dapat dilakukan pemilahan pada simpul tersebut dan menghasilkan simpul terminal. Pemilahan pertama dilakukan berdasarkan variabel kondisi air ketuban ( $X_3$ ) yang digunakan untuk memilah simpul induk, sedangkan untuk pemilahan selanjutnya dilakukan berdasarkan variabel ketuban pecah dini ( $X_4$ ), refleks rangsangan ( $X_8$ ), berat badan bayi saat lahir ( $X_1$ ) dan suhu tubuh bayi ( $X_5$ ).



Gambar 4.13 Topologi Pohon Klasifikasi Optimal

Simpul terminal merupakan simpul akhir dalam hasil pemilahan yang dibedakan menjadi dua kategori dalam pelabelan kelas. Pelabelan kelas dilakukan untuk mengetahui karakteristik

dari status risiko infeksi pada bayi baru lahir. Apabila kelas dilabelkan sebagai kelas 0 atau bayi baru lahir negatif berisiko infeksi, maka simpul tersebut akan berwarna biru. Namun, jika kelas dilabelkan sebagai kelas 1 atau bayi baru lahir memiliki risiko infeksi, maka simpul tersebut akan berwarna merah seperti yang ditunjukkan dalam Gambar 4.13. Semakin kuat warna yang dihasilkan oleh masing-masing simpul terminal, menandakan semakin besar perbedaan proporsi antara dua status risiko infeksi tersebut. Informasi tentang simpul terminal disajikan secara lengkap dalam Lampiran G. Berikut ini merupakan penyajian mengenai pelabelan kelas yang telah dilakukan terhadap simpul terminal.

**Tabel 4.9** Pelabelan Kelas Simpul Terminal

Kelas	Simpul Terminal ke-	Persentase (%)
Status Risiko Infeksi Negatif	2	96,6
	1	63,6
	3	60,0
Status Risiko Infeksi Positif	4	92,9
	5	81,8
	6	100

Berdasarkan Tabel 4.9 diperoleh informasi bahwa simpul terminal yang memiliki persentase kelas sebesar 100% dalam penandaan label kelas adalah simpul terminal 6 yang dilabeli sebagai positif risiko infeksi (kategori 1). Karakteristik untuk bayi baru lahir yang tidak memiliki risiko infeksi ditunjukkan oleh simpul terminal 2. Sedangkan, simpul terminal yang menunjukkan karakteristik untuk bayi baru lahir dengan risiko infeksi yaitu simpul terminal 1, 3, 4, 5 dan 6. Simpul terminal 3 diberi label kelas 1 (positif risiko infeksi) dengan perbedaan proporsi antara dua kelas kurang signifikan, sehingga warna yang dihasilkan pada simpul terminal 3 cenderung berwarna merah muda.

Proses pemilahan mengacu pada pohon klasifikasi optimal yang ditampilkan dalam Gambar 4.13. Pemilahan diawali dengan memilah data sebanyak 200 bayi berdasarkan kondisi air ketuban ( $X_3$ ) dan nilai *threshold* 0 (kondisi air ketuban jernih). Kemudian data dibagi menjadi dua bagian yaitu simpul kiri (*node 2*) yang terdiri dari 141 bayi dengan kondisi air ketuban yang jernih, dan dilabeli sebagai bayi yang tidak memiliki risiko infeksi karena proporsi terbesar yang terletak pada simpul tersebut adalah bayi yang status risiko infeksi negatif. Sedangkan, simpul kanan (*terminal node 6*) terdiri dari 59 bayi dengan kondisi air ketuban yang keruh atau berwarna hijau, setelah itu diberi label kelas sebagai bayi baru lahir yang memiliki risiko infeksi. Selanjutnya, pemilahan kembali dilakukan secara berulang karena tingkat heterogenan masih tinggi. Pemilahan berhenti dengan jumlah terminal simpul sebanyak 6 simpul.

Secara keseluruhan, melalui Gambar 4.13 diperoleh informasi bahwa hanya terdapat 1 simpul terminal yang diklasifikasikan sebagai bayi baru lahir yang tidak berisiko infeksi dan 5 simpul diklasifikasikan sebagai bayi baru lahir dengan risiko infeksi. Karakteristik simpul terminal yang mengklasifikasikan bayi baru lahir menjadi salah satu dari kedua status risiko infeksi dijelaskan sebagai berikut.

- a. Simpul terminal 1 terdiri dari 11 bayi yang diprediksi berisiko infeksi. Karakteristik bayi pada simpul ini adalah memiliki berat badan lahir kurang dari sama dengan 2500 gram, refleks bayi setelah diberi rangsangan menangis keras atau terdapat gerakan melawan, tidak mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih.
- b. Simpul terminal 2 terdiri dari 89 bayi yang diprediksi tidak memiliki risiko infeksi. Karakteristik bayi pada simpul ini antara lain adalah bayi memiliki suhu tubuh kurang dari sama dengan 37, berat badan lahir bayi lebih dari 2500 gram, refleks bayi setelah diberi rangsangan menangis keras atau terdapat gerakan melawan, tidak mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih.

- c. Simpul terminal 3 terdiri dari 5 bayi yang diprediksi memiliki risiko infeksi. Karakteristik bayi yang terbentuk adalah suhu tubuh bayi lebih dari 37, berat badan lahir lebih dari 2500 gram, refleks bayi setelah diberi rangsangan menangis keras atau terdapat gerakan melawan, tidak mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih.
- d. Simpul terminal 4 terdiri dari 14 bayi yang diprediksi memiliki risiko infeksi. Bayi yang termasuk dalam simpul ini memiliki beberapa karakteristik yaitu ketika bayi diberi rangsangan, bayi menunjukkan refleks merintih atau menangis lemah, tidak mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih.
- e. Simpul terminal 5 terdiri dari 22 bayi yang diprediksi berisiko infeksi. Bayi yang termasuk dalam simpul ini memiliki beberapa karakteristik antara lain adalah ibu bayi mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih.
- f. Simpul terminal 6 terdiri dari 59 bayi yang diprediksi memiliki risiko infeksi. karakteristik bayi yang termasuk simpul ini antara lain adalah sisa air ketuban dalam kondisi keruh atau berwarna hijau.

Sesuai dengan pemaparan sebelumnya mengenai karakteristik bayi baru lahir berdasarkan hasil analisis *Classification Trees*, selanjutnya ditampilkan perbedaan antara karakteristik bayi baru lahir dengan dua kategori status risiko infeksi, yakni bayi negatif risiko infeksi dan bayi positif risiko infeksi. Perbedaan karakteristik tersebut dapat dilihat melalui kriteria kondisi bayi sesaat setelah lahir. Karakteristik dari kedua kategori status risiko infeksi pada bayi baru lahir disajikan pada Tabel 4.10 sebagai berikut.

**Tabel 4.10** Karakteristik Bayi Baru Lahir Berdasarkan Simpul Terminal

<b>Negatif Risiko Infeksi</b>	<b>Positif Risiko Infeksi</b>
✓ Bayi dengan suhu tubuh $\leq 37,5$ , berat badan lahir $> 2500$ gram yang menangis keras saat diberi rangsangan, tidak mengalami ketuban pecah dini dan sisa air ketuban dalam kondisi jernih	<ul style="list-style-type: none"> <li>✓ Bayi dengan sisa air ketuban dalam kondisi keruh</li> <li>✓ Bayi dengan sisa air ketuban jernih dan mengalami ketuban pecah dini</li> <li>✓ Bayi yang merintih ketika diberi rangsangan, tidak mengalami ketuban pecah dini dan kondisi sisa air ketuban jernih</li> <li>✓ Bayi dengan berat badan <math>\leq 2500</math> gr yang menangis ketika diberi rangsangan, tidak mengalami ketuban pecah dini dan kondisi sisa air ketuban jernih</li> <li>✓ Bayi dengan suhu tubuh <math>&gt; 37,5</math> yang memiliki berat badan lahir <math>&gt; 2500</math> gr, menangis ketika diberi rangsangan, tidak mengalami ketuban pecah dini dan kondisi sisa air ketuban jernih</li> </ul>

#### 4.2.5 Hasil Ketepatan Klasifikasi *Classification Trees*

Hasil klasifikasi bayi baru lahir di RSUD Sidoarjo berdasarkan faktor-faktor yang mempengaruhi bayi berisiko infeksi menggunakan metode *10-fold cross validation* diukur berdasarkan total ketepatan klasifikasi tiap kelas dan total kesalahan klasifikasi untuk data *learning* dan data *testing*. Perhitungan ketepatan klasifikasi pada tiap *fold* ditampilkan dalam Lampiran F. Berikut ini merupakan hasil ketepatan klasifikasi risiko infeksi pada bayi baru lahir dari data *learning* untuk setiap *fold* yang ditampilkan dalam Tabel 4.11.

**Tabel 4.11** Evaluasi Performansi Hasil Klasifikasi pada Data *Learning*

<i>Fold</i>	<i>Sensitivity</i>	<i>Specificity</i>	Total Akurasi
1	0,977	0,935	0,956
2	0,989	0,957	0,972
3	0,977	0,935	0,956
4	0,956	0,966	0,961
5	0,977	0,946	0,961
6	0,978	0,967	0,972
7	0,989	0,946	0,967
8	0,978	0,956	0,967
9	0,989	0,946	0,967
10	0,977	0,350	0,956
Rata-rata	0,979	0,890	0,964
Deviasi Standar	0,010	0,190	0,006

Sesuai dengan ketepatan klasifikasi yang disajikan dalam Tabel 4.11 dapat diketahui bahwa total akurasi, *sensitivity*, *specificity* dari tiap *fold* sudah cukup tinggi. Kemudian diperoleh rata-rata ketepatan klasifikasi data *learning* sebesar 96,4% dengan rata-rata nilai *sensitivity* dan *specificity* masing-masing sebesar 0,978 dan 0,89. Nilai deviasi standar yang dimiliki untuk hasil ketepatan klasifikasi cukup kecil yang menunjukkan bahwa total akurasi dari tiap *fold* sudah stabil. Dalam perhitungan total akurasi untuk masing-masing *fold* dapat dilihat melalui persentase simpul terminal yang tepat diklasifikasikan atau yang mengalami kesalahan klasifikasi.

Selanjutnya, untuk mengetahui layak atau tidaknya pohon klasifikasi optimal yang terbentuk untuk bayi baru lahir di RSUD Sidoarjo berdasarkan status risiko infeksi perlu dilakukan validasi menggunakan data *testing*. Hasil ketepatan klasifikasi risiko infeksi pada bayi baru lahir menggunakan data *testing* disajikan dalam Tabel 4.12 sebagai berikut.

**Tabel 4.12** Ketepatan Klasifikasi Risiko Infeksi pada Bayi Baru Lahir untuk Data *Testing*

Aktual	Prediksi		Total	Total Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
	Negatif	Positif				
Negatif	94	7	101	0,935	0,931	0,939
Positif	6	93	99			
<b>Total</b>	100	100	200			

Berdasarkan Tabel 4.12 dapat diketahui terdapat 101 bayi baru lahir yang termasuk dalam kategori negatif (tidak memiliki risiko infeksi) dengan 94 bayi tepat diklasifikasikan sebagai bayi baru lahir dengan status risiko infeksi negatif dan 7 bayi salah diklasifikasikan sebagai bayi baru lahir yang memiliki risiko infeksi. Sehingga nilai *sensitivity* yang diperoleh sebesar 0,931. Sementara itu jumlah bayi baru lahir yang termasuk dalam kategori positif (memiliki risiko infeksi) sebanyak 99 bayi, dimana 93 bayi tepat diklasifikasikan sebagai bayi baru lahir berisiko infeksi, sisanya terdapat 6 bayi yang salah diklasifikasikan sebagai bayi baru lahir yang tidak memiliki risiko infeksi. Sehingga diperoleh nilai *specificity* sebesar 0,939.

Total akurasi (1-APER) yang dihasilkan untuk klasifikasi risiko infeksi pada bayi baru lahir di RSUD Sidoarjo dengan menggunakan aturan pemilihan pemilah indeks gini sesuai dengan yang disajikan dalam Tabel 4.12 adalah sebesar 93,5%, sedangkan total kesalahan klasifikasi (APER) yang dihasilkan sebesar 6,5%. Karena total akurasi yang cukup tinggi, maka pohon klasifikasi optimal yang terbentuk sudah layak digunakan untuk pengklasifikasian risiko infeksi pada bayi yang baru lahir berikutnya.

#### 4.2.6 *Odds Ratio* Bayi Baru Lahir Positif Risiko Infeksi

*Odds Ratio* dapat menunjukkan kecenderungan atau risiko bayi baru lahir positif terinfeksi dengan kondisi tertentu dibandingkan dengan kondisi lain berdasarkan kriteria yang sama. Berikut ini merupakan perhitungan *odds ratio* pada bayi baru lahir yang memiliki risiko infeksi.

a. Kondisi sisa air ketuban keruh

$$\text{Odd Ratio (1)} = \frac{101 \times 59}{1 \times 40} = 148,975$$

Risiko bayi positif infeksi dengan kondisi sisa air ketuban keruh 148,975 kali lebih besar daripada bayi dengan kondisi sisa air ketuban jernih.

b. Status tidak mengalami ketuban pecah dini

$$\text{Odd Ratio (2)} = \frac{97 \times 18}{4 \times 22} = 19,84$$

Risiko bayi positif infeksi yang mengalami ketuban pecah dini 19,84 kali lebih besar daripada bayi yang tidak mengalami ketuban pecah dini.

c. Refleks merintih dari bayi terhadap stimulan

$$\text{Odd Ratio (3)} = \frac{96 \times 13}{1 \times 9} = 138,67$$

Risiko bayi positif infeksi yang menerima stimulan dengan respon merintih 138,67 kali lebih besar daripada bayi yang menerima stimulan dengan respon menangis.

d. Berat badan lahir bayi rendah

$$\text{Odd Ratio (4)} = \frac{4 \times 95}{1 \times 5} = 76,92$$

Risiko bayi positif infeksi yang memiliki berat  $\leq$  2500 gram adalah 76,92 kali lebih besar daripada bayi yang mempunyai berat badan lahir  $>$  2500 gram.

e. Suhu tubuh bayi kurang dari  $37,5^{\circ}\text{C}$

$$\text{Odd Ratio (5)} = \frac{3 \times 1}{94 \times 2} = 0,016$$

Risiko bayi positif infeksi adalah 0,016 kali lebih kecil daripada bayi yang memiliki suhu tubuh  $>$   $37,5^{\circ}\text{C}$   $\leq$   $^{\circ}\text{C}$

f. Suhu tubuh bayi lebih dari  $37,5^{\circ}\text{C}$

$$\text{Odd Ratio (6)} = \frac{1}{0,016} = 62,5$$

Risiko bayi positif infeksi yang memiliki suhu tubuh  $> 37,5^{\circ}\text{C}$  adalah 62,5 kali lebih besar daripada bayi yang memiliki suhu  $\leq$   $^{\circ}\text{C}$ .

#### 4.2.7 Ilustrasi Klasifikasi Risiko Infeksi

Pohon klasifikasi optimal yang terbentuk dapat digunakan untuk mengklasifikasi bayi baru lahir terhadap status risiko infeksi berdasarkan faktor-faktor yang mempengaruhi. Berikut ini merupakan ilustrasi dalam mengklasifikasikan bayi baru lahir untuk mengetahui status risiko infeksi berdasarkan pohon klasifikasi optimal seperti yang ditampilkan dalam Gambar 4.12. Klasifikasi dilakukan dengan cara menelusuri pohon klasifikasi tersebut dari atas ke bawah. Dalam ilustrasi diberikan 40 data bayi baru lahir dengan indikator-indikator yang dimiliki bayi berdasarkan data rekam medis di RSUD Sidoarjo pada akhir bulan Oktober 2015 seperti yang ditunjukkan dalam Tabel 4.13 sebagai berikut.

**Tabel 4.13** Data Rekam Medis Bayi Baru Lahir

Bayi	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>	X <sub>10</sub>
1	2570	0	1	0	36,8	0	120	0	0	0
2	3500	0	1	0	38,0	0	148	0	0	0
3	3100	0	0	0	36,9	0	124	0	0	0
4	3300	0	0	0	36,8	0	124	0	0	0
5	1600	1	0	0	36,5	0	136	1	1	0
...	...	...	...	...	...	...	...	...	...	...
36	2100	1	1	1	36,2	0	142	0	0	0
37	3100	0	0	1	36,0	0	140	0	0	0
38	3200	0	0	0	36,6	0	140	0	0	0
39	3200	0	0	0	36,8	0	124	0	0	0
40	2600	0	1	0	36,9	0	130	0	0	0

Berdasarkan data rekam medis bayi baru lahir yang ditampilkan dalam Tabel 4.13 dilakukan pengklasifikasian

dengan cara menelusuri pohon klasifikasi optimal. Berikut ini dijelaskan penelusuran pohon klasifikasi optimal yang dilakukan untuk menentukan status risiko infeksi pada bayi baru lahir berdasarkan ilustrasi data rekam medis.

Untuk bayi yang pertama dilihat kondisi sisa air ketuban ( $X_3$ ). Apabila kondisi sisa air ketuban keruh (kategori 1), maka bayi pertama diklasifikasikan sebagai bayi positif risiko infeksi. Bayi yang kedua juga memiliki kondisi sisa air ketuban keruh, sehingga bayi kedua diklasifikasikan sebagai bayi positif risiko infeksi.

Selanjutnya, bayi ketiga dilihat kembali kondisi sisa air ketuban. Apabila kondisi sisa air ketuban jernih (kategori 0), maka diperiksa status ketuban pecah dini ( $X_4$ ). Apabila ibu bayi melahirkan dengan kondisi tidak mengalami ketuban pecah dini (kategori 0), maka diperiksa apakah bayi menerima rangsangan ( $X_8$ ) dengan respon menangis (kategori 0) atau merintih (kategori 1). Jika respon bayi menangis, maka dilihat berat badan lahir bayi tersebut. Apabila bayi tersebut memiliki berat badan lahir ( $X_1$ ) kurang dari sama dengan 2500 gram, maka bayi tersebut berstatus positif risiko infeksi. Bayi ketiga memiliki berat badan lahir lebih dari 2500 gram, oleh sebab itu dilanjutkan untuk melihat suhu tubuh bayi ( $X_5$ ). Suhu tubuh bayi ketiga sebesar  $36,9^{\circ}\text{C}$  (kurang dari  $37$ ), ketiga diklasifikasikan sebagai bayi baru lahir dengan status negatif risiko infeksi.

Kemudian bayi yang keempat dilihat kembali kondisi sisa air ketuban ( $X_3$ ). Apabila kondisi sisa air ketuban jernih (kategori 0), maka diperiksa status ketuban pecah dini ( $X_4$ ). Apabila ibu bayi melahirkan dengan kondisi tidak mengalami ketuban pecah dini (kategori 0), maka diperiksa apakah bayi menerima rangsangan ( $X_8$ ) dengan respon menangis (kategori 0) atau merintih (kategori 1). Bayi keempat memberikan respon menangis, sehingga dilihat berat badan lahir bayi tersebut. Bayi keempat memiliki berat badan lahir 3300 gram ( $> 2500$  gram), oleh sebab itu dilanjutkan untuk melihat suhu tubuh bayi ( $X_5$ ). Suhu tubuh bayi keempat sebesar  $36,8^{\circ}$

sehingga bayi ketiga diklasifikasikan sebagai bayi baru lahir dengan status negatif risiko infeksi.

Selanjutnya bayi kelima dilihat pula kondisi sisa air ketuban ( $X_3$ ). Karena bayi kelima memiliki sisa air ketuban jernih, maka dilihat status ketuban pecah dini ( $X_4$ ). Bayi kelima tidak mengalami ketuban pecah dini, sehingga perlu dilihat kembali refleks yang diberikan oleh bayi setelah diberi stimulan ( $X_8$ ). Bayi kelima memberikan respon merintih (kategori 1), sehingga bayi kelima diklasifikasikan sebagai bayi positif risiko infeksi. Begitu seterusnya dilakukan pada bayi keenam hingga bayi terakhir untuk mengetahui status risiko infeksi pada bayi tersebut. Hasil klasifikasi risiko infeksi pada bayi baru lahir tersebut akan menjadi variabel respon ( $Y$ ) sebagai kelas seperti yang ditunjukkan dalam Tabel 4.14 sebagai berikut. Untuk hasil klasifikasi secara lengkap ditampilkan dalam lampiran L.

**Tabel 4.14** Hasil Klasifikasi Bayi Baru Lahir

Bayi	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$	$X_{10}$	Status Risiko Infeksi
1	2570	0	1	0	36,8	0	120	0	0	0	Positif
2	3500	0	1	0	38,0	0	148	0	0	0	Positif
3	3100	0	0	0	36,9	0	124	0	0	0	Negatif
4	3300	0	0	0	36,8	0	124	0	0	0	Negatif
5	1600	1	0	0	36,5	0	136	1	1	0	Positif
...	...	...	...	...	...	...	...	...	...	...	....
36	2100	1	1	1	36,2	0	142	0	0	0	Positif
37	3100	0	0	1	36,0	0	140	0	0	0	Positif
38	3200	0	0	0	36,6	0	140	0	0	0	Negatif
39	3200	0	0	0	36,8	0	124	0	0	0	Negatif
40	2600	0	1	0	36,9	0	130	0	0	0	Positif

Setelah diperoleh hasil klasifikasi dari data rekam medis bayi baru lahir berdasarkan pohon klasifikasi optimal hasil analisis *Classification Trees*, langkah berikutnya adalah membandingkan dengan status risiko infeksi dari diagnosa dokter terhadap bayi baru lahir tersebut. Berikut disajikan tabulasi silang beserta evaluasi ketepatan klasifikasi pada Tabel 4.15.

**Tabel 4.15** Tabulasi Silang Hasil Ketepatan Klasifikasi Bayi Baru Lahir

Aktual	Prediksi		Total	Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
	0	1				
0	21	3	24	0,9000	0,8750	0,9375
1	1	15	16			
Total	22	18	40			

Tabel 4.15 memberikan informasi bahwa perbandingan antara hasil klasifikasi bayi baru lahir berdasarkan diagnosa dokter dengan pohon klasifikasi optimal diperoleh akurasi sebesar 90% dengan nilai *sensitivity* dan *specificity* masing-masing sebesar 87,5% dan 93,75%. Sebanyak 21 dari 24 bayi tepat diklasifikasi sebagai bayi negatif risiko infeksi dan 3 bayi diklasifikasikan sebagai bayi positif risiko infeksi. Sedangkan sebanyak 15 dari 16 bayi tepat diklasifikasikan sebagai bayi positif risiko infeksi dan 1 bayi salah diklasifikasikan sebagai bayi negatif risiko infeksi.



*“Halaman ini sengaja dikosongkan”*

## BAB V KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan untuk pengklasifikasian risiko infeksi pada bayi baru lahir di RSUD Sidoarjo pada bulan September-Oktober 2015 diperoleh kesimpulan sebagai berikut.

1. Jumlah bayi baru lahir sebanyak 247 jiwa dengan jumlah kasus *missing value* sebanyak 7 kasus pada 2 variabel, yakni variabel suhu tubuh bayi ( $X_5$ ) dan denyut jantung ( $X_7$ ) dengan jumlah data yang mengandung *missing* masing-masing sebanyak 1 dan 6 kasus. Karena kasus *missing value* kurang dari 30%, maka data yang mengandung *missing* dihilangkan dalam analisis. Sehingga jumlah data menjadi 240 bayi dengan persentase status risiko infeksi negatif dan positif masing-masing sebesar 48% dan 52%.
2. Dalam analisis *Classification Trees* digunakan 200 data untuk membentuk pohon klasifikasi dan 40 data sisanya digunakan sebagai data rekam medis bayi yang baru lahir untuk memperoleh klasifikasi status risiko infeksi berdasarkan hasil pohon klasifikasi. Berdasarkan analisis *Classification Trees* diperoleh kondisi sisa air ketuban sebagai variabel yang paling berpengaruh untuk mengklasifikasi risiko infeksi pada bayi baru lahir di RSUD Sidoarjo. Selain itu, diperoleh juga hasil akurasi dengan menggunakan aturan pemilihan pemilah indeks gini berdasarkan data *testing* sebesar 93,5% dengan persentase *sensitivity* dan *specificity* masing-masing sebesar 93,1% dan 93,9%. Jumlah simpul terminal yang dihasilkan pada pohon klasifikasi optimal sebanyak 6 simpul dengan tingkat kedalaman sebesar 6.
3. Karakteristik yang dihasilkan oleh bayi baru lahir dengan status negatif risiko infeksi adalah bayi dengan suhu tubuh  $\leq$   $t_{b \quad h}$  0 gram yang menangis keras

saat diberi stimulus, tidak mengalami ketuban pecah dini dan dan sisa air ketuban dalam kondisi jernih.

4. Karakteristik bayi baru lahir dengan status positif risiko infeksi adalah bayi dengan sisa air ketuban dalam kondisi keruh; bayi dengan kondisi sisa air ketuban jernih dan mengalami KPD; bayi yang merintih ketika diberi stimulan, tidak mengalami KPD dan kondisi sisa air ketuban jernih; bayi dengan  $\text{b t b h} \leq 0$  gr yang menangis ketika diberi stimulan, tidak mengalami KPD dan kondisi sisa air ketuban jernih; serta bayi dengan suhu tubuh  $> 37$ , yang memiliki berat badan lahir  $> 2500$  gr, menangis ketika diberi stimulan, tidak mengalami KPD dan kondisi sisa air ketuban jernih.

## 5.2 Saran

Rekomendasi atau saran yang dapat diberikan untuk penelitian selanjutnya yaitu sebaiknya data yang digunakan dalam analisis memiliki ukuran yang lebih besar (*big data*). Hal tersebut bertujuan untuk meningkatkan akurasi pengklasifikasian. Bagi pihak Rumah Sakit Umum Daerah (RSUD) Sidoarjo diperlukan beberapa variabel lainnya untuk memudahkan proses klasifikasi status risiko infeksi pada bayi baru lahir seperti faktor eksternal, yakni faktor dari kondisi ibu sebelum dan setelah melahirkan serta lingkungan sekitar rumah yang tidak didapatkan pada penelitian saat ini. Sehingga hasil klasifikasi bayi baru lahir terhadap status risiko infeksi akan lebih akurat.

## DAFTAR PUSTAKA

- Agresti, A. (2007). *Categorical data analysis*. New York: John Wiley dan Sons..
- BPS. (2014). *Laporan eksekutif kesehatan provinsi Jawa Timur tahun 2014*. Surabaya: BPS Jatim.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and regression trees*. New York: Chapman Hall.
- Carolus, W., Rompis, J., & Wilar, R. (2013). *Hubungan APGAR skor dan berat badan lahir dengan sepsis neonatorum*. Bagian Ilmu Kesehatan Anak Fakultas Kedokteran Universitas Sam Ratulangi, Manado.
- CART Reference Guide. (2000). *CART User's Guide*. San Diego: Salford System.
- Depkes. (2008). *Laporan Nasional riset kesehatan dasar tahun 2007*. Jakarta: Depkes RI.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis (7th ed.)*. New Jersey: Prentice Hall.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis (6th ed.)*. New Jersey: Prentice Hall.
- Kementrian Kesehatan RI. (2010). *Buku saku pelayanan kesehatan neonatal esensial*. Jakarta: Kemenkes.
- Khasanah, P. M. (2015). *Klasifikasi penderita retardasi mental di Rumah Sakit Jiwa Menur provinsi Jawa Timur menggunakan synthetic minority oversampling technique (SMOTE) – classification and regression trees (CART)*. Tugas program sarjana tidak dipublikasikan, Institut Teknologi Sepuluh Nopember, Surabaya.
- Lewis, R. J. (2000). *An introduction to classification and regression trees (CART) analysis*. Paper presented at the annual meeting of the Society for Academic Emergency Medicine, California.

- Marmi & Rahardjo, K. (2012). *Asuhan neonatus, bayi, balita, dan anak prasekolah*. Yogyakarta: Pustaka Pelajar.
- Margasari, A. (2014). *Penerapan metode CART (classification and regression trees) dan analisis regresi logistik biner pada klasifikasi profil mahasiswa FMIPA Universitas Brawijaya*. Skripsi program sarjana tidak dipublikasikan, Universitas Brawijaya, Malang.
- Prawirohardjo, S. (2014). *Ilmu kebidanan (4th ed.)*. Jakarta: Yayasan Bina Pustaka Sarwono Prawirohardjo.
- Rucy, Y. N. (2014). *Klasifikasi risiko infeksi perinatal bayi di RSUD Sidoarjo dengan regresi logistik biner*. Tugas akhir program sarjana tidak dipublikasikan, Institut Teknologi Sepuluh Nopember, Surabaya.
- Sinseng, H. (2008). *Perbedaan nilai APGAR antara persalinan normal dengan persalinan riwayat ketuban pecah dini di RSUD Dr. Moewardi Surakarta*. Skripsi program sarjana tidak dipublikasikan, Universitas Sebelas Maret, Surakarta.
- Utomo, MT. (2010). Risk factors of neonatal sepsis : a preliminary study in Dr. Soetomo Hospital. *Indonesian Journal of Tropical and Infections Disease*.
- Walpole, R. E. (1995). *Pengantar statistika*. Jakarta: Gramedia Pustaka Utama.
- WHO. (2002). *National Neonatal-Perinatal Database Report*. (2010). *International Classification of Disease 2010*.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: practical machine learning tools and techniques (3rd ed)*. United States: Morgan Kaufmann.
- Yefri, R., Mayetti, & Machmud, R. (2010). Kolonisasi kuman dan kejadian omfalitis pada tiga regimen perawatan tali pusar pada bayi baru lahir. *Jurnal Sari Pediatri*, 8, 341-347.

## LAMPIRAN

### Lampiran A.

Data Pengamatan Bayi yang Lahir dan yang dirujuk di RSUD Sidoarjo Bulan September-Oktober 2015

No.	Y	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>	X <sub>10</sub>
1	0	2900	0	0	1	36,8	0	140	0	0	0
2	0	4100	0	0	0	38,9	0	120	0	0	0
3	1	2360	0	0	0	36,8	0	124	0	1	0
4	1	2610	0	1	1	37	0	120	0	0	0
5	0	3200	0	0	0	37	0	140	0	0	0
6	1	3200	0	0	0	36,8	0	120	0	1	0
7	1	3000	0	1	0	36,8	0	120	0	0	0
8	1	3000	0	1	1	36,5	1	120	0	0	1
9	1	3710	0	1	1	36,8	0	140	0	0	1
10	1	2900	0	1	0	36,8	0	120	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
121	1	2900	1	0	0	35,5	0	120	1	1	1
122	0	3400	0	0	0	36,8	0	120	0	0	0
123	0	4000	0	0	0	36,6	0	125	0	0	0
124	0	3200	0	0	0	36,8	0	140	0	0	0
125	1	2000	0	0	0	38,3	0	120	0	0	1
126	1	1900	1	0	1	35,5	1	130	0	0	1
127	0	2900	0	0	0	37	0	136	0	0	0
128	0	2800	0	0	0	36,8	0	128	0	0	0
129	1	2800	0	0	0	36,7	0	140	1	1	0
130	0	3800	0	0	0	35,6	0	130	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
241	1	2300	1	1	1	36,4	0	142	0	0	0
242	1	2500	1	1	1	36,7	0	120	0	0	1
243	1	2100	1	1	1	36,2	0	142	0	0	0
244	0	3100	0	0	1	36	0	140	0	0	0
245	0	3200	0	0	0	36,6	0	140	0	0	0
246	0	3200	0	0	0	36,8	0	124	0	0	0
247	1	2600	0	1	0	36,9	0	130	0	0	0

**Keterangan :**

$Y(0)$  : Bayi baru lahir berstatus negatif risiko infeksi

$Y(1)$  : Bayi baru lahir berstatus positif risiko infeksi

$X_1$  : Berat badan lahir (gram)

$X_2$  : Status lahir prematur

$X_3$  : Kondisi air ketuban

$X_4$  : Status ketuban pecah dini

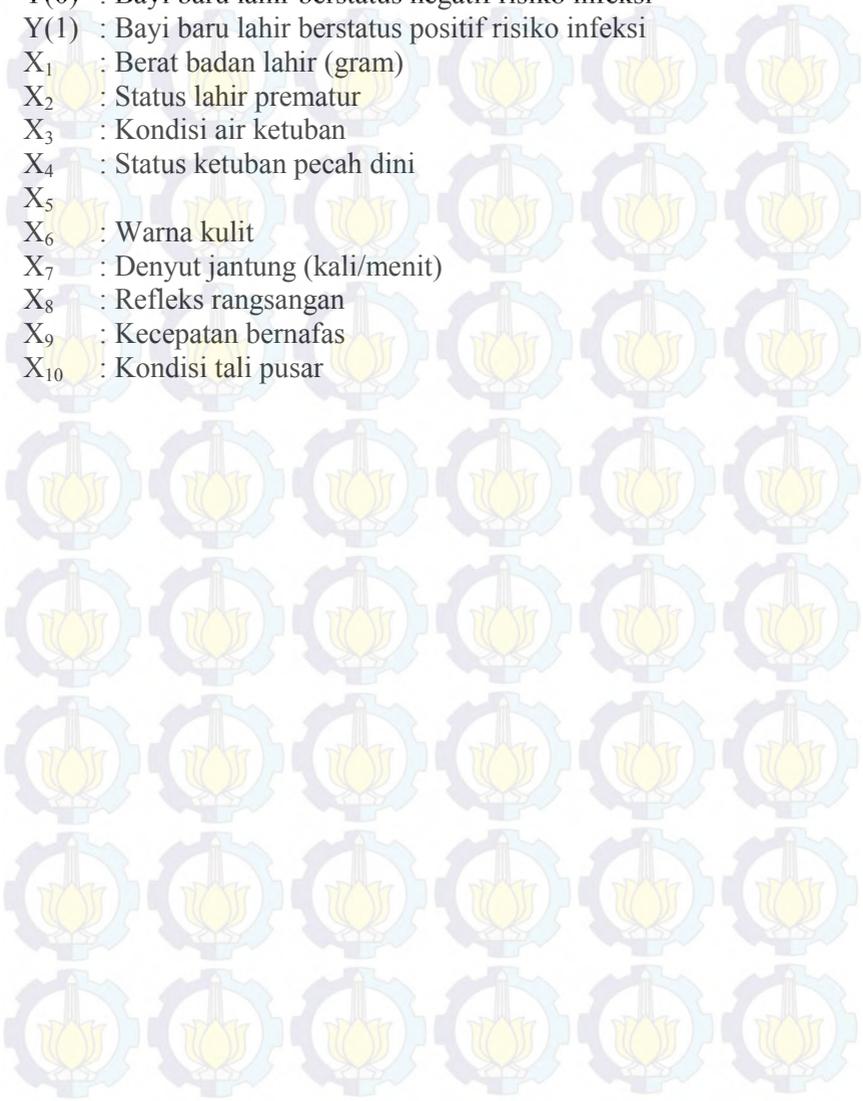
$X_5$  : Warna kulit

$X_7$  : Denyut jantung (kali/menit)

$X_8$  : Refleks rangsangan

$X_9$  : Kecepatan bernafas

$X_{10}$  : Kondisi tali pusar



**Lampiran B.**

Kemungkinan pemilah pada Variabel Prediktor berskala kontinyu

<b>Berat Badan Lahir (gram)</b>	<b>Denyut Jantung (kali/menit)</b>	<b>Suhu Tubuh</b>
1500,0	115,0	35,25
1700,0	121,0	35,55
1850,0	123,0	35,80
1950,0	124,5	36,10
2050,0	125,5	36,25
2150,0	127,0	36,35
2250,0	129,0	36,45
2327,5	131,0	36,55
2357,5	133,5	36,65
2380,0	135,5	36,75
2425,0	137,0	36,85
2460,0	139,0	36,95
2485,0	141,0	37,20
2525,0	143,0	37,50
2575,0	145,0	37,65
2605,0	147,0	37,85
2635,0	149,0	38,15
...	...	...

**Lampiran C.****Statistik Deskriptif dan Tabulasi Silang antara Variabel Respon dan Prediktor**

## LEARNING SAMPLE VARIABLE STATISTICS

VARIABLE		CLASS		OVERALL
		0	1	
Y	MEAN	0.000	1.000	0.495
	SD	0.000	0.000	0.501
	N	101.000	99.000	200.000
	SUM	0.000	99.000	99.000
X1	MEAN	3182.277	3031.414	3107.600
	SD	459.267	619.547	548.395
	N	101.000	99.000	200.000
	SUM	321410.000	300110.000	621520.000
X2	MEAN	0.050	0.131	0.090
	SD	0.218	0.339	0.287
	N	101.000	99.000	200.000
	SUM	5.000	13.000	18.000
X3	MEAN	0.000	0.596	0.295
	SD	0.000	0.493	0.457
	N	101.000	99.000	200.000
	SUM	0.000	59.000	59.000
X4	MEAN	0.040	0.333	0.185
	SD	0.196	0.474	0.389
	N	101.000	99.000	200.000
	SUM	4.000	33.000	37.000
X5	MEAN	36.712	36.721	36.717
	SD	0.393	0.526	0.462
	N	101.000	99.000	200.000
	SUM	3707.900	3635.400	7343.300
X6	MEAN	0.010	0.061	0.035
	SD	0.100	0.240	0.184
	N	101.000	99.000	200.000
	SUM	1.000	6.000	7.000
X7	MEAN	128.822	127.394	128.115
	SD	9.110	9.401	9.260
	N	101.000	99.000	200.000
	SUM	13011.000	12612.000	25623.000
X8	MEAN	0.010	0.182	0.095
	SD	0.100	0.388	0.294
	N	101.000	99.000	200.000
	SUM	1.000	18.000	19.000
X9	MEAN	0.010	0.182	0.095
	SD	0.100	0.388	0.294
	N	101.000	99.000	200.000
	SUM	1.000	18.000	19.000

### Lampiran C. (lanjutan)

X10	MEAN	0.010	0.232	0.120
	SD	0.100	0.424	0.326
	N	101.000	99.000	200.000
	SUM	1.000	23.000	24.000

---

AUTOMATIC LEVEL SETTINGS		
NAME	LEVELS	MINIMUM
X2	2	0
X3	2	0
X4	2	0
X6	2	0
X8	2	0
X9	2	0
X10	2	0

### Crosstabs

#### Lahir Prematur \* Status Risiko Infeksi Crosstabulation

Count

		Status Risiko Infeksi		Total
		Negatif	Positif	
Lahir Prematur	Lahir cukup bulan atau lebih	96	86	182
	Lahir Prematur	5	13	18
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	,143	,001
	Cramer's V	,143	,001
	Contingency Coefficient	,141	,001
N of Valid Cases		200	

### Lampiran C. (lanjutan)

#### Kondisi Air Ketuban \* Status Risiko Infeksi Crosstabulation

Count

		Status Risiko Infeksi		Total
		Negatif	Positif	
Kondisi Air	Jernih	101	40	141
Ketuban	Keruh atau hijau	0	59	59
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,653	,000
Nominal	Cramer's V	,653	,000
	Contingency Coefficient	,547	,000
N of Valid Cases		200	

#### Ketuban Pecah Dini \* Status Risiko Infeksi Crosstabulation

Count

		Status Risiko Infeksi		Total
		Negatif	Positif	
Ketuban	Tidak	97	66	163
Pecah Dini	KPD	4	33	37
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,378	,000
Nominal	Cramer's V	,378	,000
	Contingency Coefficient	,354	,000
N of Valid Cases		200	

### Lampiran C. (lanjutan)

#### Warna Kulit \* Status Risiko Infeksi Crosstabulation

Count

		Status Risiko Infeksi		Total
		Negatif	Positif	
Warna Kulit	Kemerahan	100	93	193
	Kebiruan	1	6	7
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,138	,023
Nominal	Cramer's V	,138	,023
	Contingency Coefficient	,137	,023
N of Valid Cases		200	

#### Refleks Rangsangan \* Status Risiko Infeksi Crosstabulation

Count

		Status Risiko Infeksi		Total
		Negatif	Positif	
Refleks	Menangis	100	81	181
Rangsangan	Merintih	1	18	19
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,293	,000
Nominal	Cramer's V	,293	,000
	Contingency Coefficient	,281	,000
N of Valid Cases		200	

### Lampiran C. (lanjutan)

#### Kecepatan Bernafas \* Status Risiko Infeksi Crosstabulation

Count		Status Risiko Infeksi		Total
		Negatif	Positif	
Kecepatan Bernafas	Teratur	100	81	181
	Tidak Teratur	1	18	19
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,293	,000
Nominal	Cramer's V	,293	,000
	Contingency Coefficient	,281	,000
N of Valid Cases		200	

#### Kondisi Tali Pusar \* Status Risiko Infeksi Crosstabulation

Count		Status Risiko Infeksi		Total
		Negatif	Positif	
Kondisi Tali Pusar	Segar	100	76	176
	Layu	1	23	24
Total		101	99	200

#### Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,342	,000
Nominal	Cramer's V	,342	,000
	Contingency Coefficient	,324	,000
N of Valid Cases		200	

## Lampiran D.

### Output Pohon Klasifikasi 10-Fold Cross Validation Estimate

\*\*\*\*\* CROSS VALIDATION TREE 1 \*\*\*\*\*

=====

TERMINAL NODE INFORMATION

=====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	14	0.079	1	0.070	0	1	0.070	0.068
					1	13	0.930	
2	86	0.473	0	0.024	0	84	0.976	.123472E-03
					1	2	0.024	
3	2	0.011	1	0.494	0	1	0.494	.123472E-03
					1	1	0.506	
4	1	0.006	1	0.000	0	0	0.000	0.006
					1	1	1.000	
5	15	0.084	1	0.262	0	4	0.262	0.040
					1	11	0.738	
6	9	0.051	1	0.000	0	0	0.000	0.005
					1	9	1.000	
7	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	
8	52	0.292	1	0.000	0	0	0.000	0.292
					1	52	1.000	

\*\*\*\*\* CROSS VALIDATION TREE 2 \*\*\*\*\*

=====

TERMINAL NODE INFORMATION

=====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	4	0.022	1	0.246	0	1	0.246	0.011
					1	3	0.754	
2	86	0.473	0	0.012	0	85	0.988	0.006
					1	1	0.012	
3	1	0.006	1	0.000	0	0	0.000	0.006
					1	1	1.000	
4	3	0.017	1	0.328	0	1	0.328	0.006
					1	2	0.672	
5	10	0.056	1	0.000	0	0	0.000	0.005
					1	10	1.000	
6	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	
7	9	0.050	1	0.218	0	2	0.218	0.005
					1	7	0.782	

**Lampiran D. (lanjutan)**

8	1	0.005	0	0.000	0	1	1.000	0.005
9	10	0.056	1	0.000	0	0	0.000	0.003
10	55	0.309	1	0.000	0	0	0.000	0.309

\*\*\*\*\* CROSS VALIDATION TREE 3 \*\*\*\*\*

===== TERMINAL NODE INFORMATION =====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	5	0.028	1	0.196	0	1	0.196	0.017
2	86	0.473	0	0.024	0	84	0.976	0.006
3	1	0.006	1	0.000	0	0	0.000	0.006
4	3	0.017	1	0.328	0	1	0.328	0.006
5	11	0.062	1	0.000	0	0	0.000	0.005
6	1	0.005	0	0.000	0	1	1.000	0.005
7	19	0.106	1	0.207	0	4	0.207	0.062
8	54	0.303	1	0.000	0	0	0.000	0.303

\*\*\*\*\* CROSS VALIDATION TREE 4 \*\*\*\*\*

===== TERMINAL NODE INFORMATION =====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	87	0.483	0	0.034	0	84	0.966	0.006
2	1	0.006	1	0.000	0	0	0.000	0.006
3	5	0.028	1	0.200	0	1	0.200	0.017
4	6	0.033	1	0.000	0	0	0.000	0.003
5	3	0.017	0	0.333	0	2	0.667	0.006

### Lampiran D. (lanjutan)

6	11	0.061	1	0.182	0	2	0.182	0.006
					1	9	0.818	
7	13	0.072	1	0.000	0	0	0.000	0.006
					1	13	1.000	
8	1	0.006	0	0.000	0	1	1.000	0.006
					1	0	0.000	
9	53	0.294	1	0.000	0	0	0.000	0.294
					1	53	1.000	
***** CROSS VALIDATION TREE 5 *****								
=====								
TERMINAL NODE INFORMATION								
=====								
Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	5	0.028	1	0.196	0	1	0.196	0.017
					1	4	0.804	
2	87	0.478	0	0.023	0	85	0.977	0.006
					1	2	0.023	
3	1	0.006	1	0.000	0	0	0.000	0.006
					1	1	1.000	
4	3	0.017	1	0.328	0	1	0.328	0.006
					1	2	0.672	
5	12	0.067	1	0.000	0	0	0.000	0.067
					1	12	1.000	
6	6	0.034	1	0.000	0	0	0.000	0.002
					1	6	1.000	
7	11	0.061	1	0.268	0	3	0.268	0.005
					1	8	0.732	
8	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	
9	3	0.017	1	0.000	0	0	0.000	0.003
					1	3	1.000	
10	51	0.287	1	0.000	0	0	0.000	0.287
					1	51	1.000	
***** CROSS VALIDATION TREE 6 *****								
=====								
TERMINAL NODE INFORMATION								
=====								
Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	4	0.022	1	0.000	0	0	0.000	0.022
					1	4	1.000	
2	89	0.489	0	0.023	0	87	0.977	0.006
					1	2	0.023	

**Lampiran D. (lanjutan)**

3	1	0.006	1	0.000	0	0	0.000	0.006
4	3	0.017	1	0.328	0	1	0.328	0.006
5	11	0.062	1	0.000	0	0	0.000	0.005
6	1	0.005	0	0.000	0	1	1.000	0.005
7	17	0.095	1	0.115	0	2	0.115	0.073
8	54	0.303	1	0.000	0	0	0.000	0.303
***** CROSS VALIDATION TREE 7 *****								
=====								
TERMINAL NODE INFORMATION								
=====								
Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	5	0.028	1	0.196	0	1	0.196	0.017
2	85	0.467	0	0.012	0	84	0.988	0.006
3	1	0.006	1	0.000	0	0	0.000	0.006
4	3	0.017	1	0.328	0	1	0.328	0.006
5	13	0.073	1	0.000	0	0	0.000	0.005
6	1	0.005	0	0.000	0	1	1.000	0.005
7	6	0.034	1	0.000	0	0	0.000	0.002
8	11	0.061	1	0.268	0	3	0.268	0.005
9	1	0.005	0	0.000	0	1	1.000	0.005
10	3	0.017	1	0.000	0	0	0.000	0.003
11	51	0.287	1	0.000	0	0	0.000	0.287

## Lampiran D. (lanjutan)

\*\*\*\*\* CROSS VALIDATION TREE 8 \*\*\*\*\*

=====

TERMINAL NODE INFORMATION

=====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	88	0.484	0	0.023	0	86	0.977	0.011

					1	2	0.023	
--	--	--	--	--	---	---	-------	--

2	2	0.011	1	0.000	0	0	0.000	0.011
					1	2	1.000	

3	5	0.028	1	0.196	0	1	0.196	0.017
					1	4	0.804	

4	12	0.067	1	0.000	0	0	0.000	0.005
					1	12	1.000	

5	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	

6	18	0.101	1	0.164	0	3	0.164	0.068
					1	15	0.836	

7	54	0.303	1	0.000	0	0	0.000	0.303
					1	54	1.000	

\*\*\*\*\* CROSS VALIDATION TREE 9 \*\*\*\*\*

=====

TERMINAL NODE INFORMATION

=====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
1	2	0.011	1	0.494	0	1	0.494	0.002

					1	1	0.506	
--	--	--	--	--	---	---	-------	--

2	46	0.253	0	0.022	0	45	0.978	0.006
					1	1	0.022	

3	1	0.006	1	0.000	0	0	0.000	0.006
					1	1	1.000	

4	40	0.220	0	0.000	0	40	1.000	0.003
					1	0	0.000	

5	1	0.006	1	0.000	0	0	0.000	0.006
					1	1	1.000	

6	4	0.022	1	0.000	0	0	0.000	0.022
					1	4	1.000	

7	11	0.062	1	0.000	0	0	0.000	0.005
					1	11	1.000	

8	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	

9	21	0.117	1	0.187	0	4	0.187	0.074
					1	17	0.813	

10	53	0.298	1	0.000	0	0	0.000	0.298
					1	53	1.000	

**Lampiran D. (lanjutan)**

\*\*\*\*\* CROSS VALIDATION TREE 10 \*\*\*\*\*

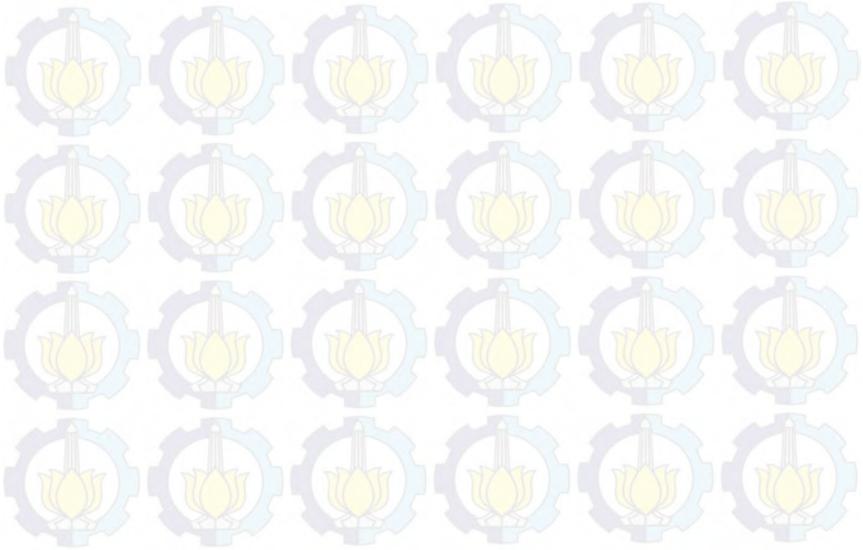
=====

TERMINAL NODE INFORMATION

=====

Node	N	Prob	Class	Cost	Class	N	Prob	Complexity Threshold
------	---	------	-------	------	-------	---	------	----------------------

1	4	0.022	1	0.246	0	1	0.246	0.011
					1	3	0.754	
2	86	0.473	0	0.024	0	84	0.976	.123472E-03
					1	2	0.024	
3	2	0.011	1	0.494	0	1	0.494	.123472E-03
					1	1	0.506	
4	12	0.067	1	0.000	0	0	0.000	0.005
					1	12	1.000	
5	1	0.005	0	0.000	0	1	1.000	0.005
					1	0	0.000	
6	21	0.117	1	0.187	0	4	0.187	0.074
					1	17	0.813	
7	54	0.303	1	0.000	0	0	0.000	0.303
					1	54	1.000	



## Lampiran E.

### Output Kesalahan Klasifikasi (*Missclassification*)

=====							
MISCLASSIFICATION BY CLASS							
=====							
(Cross Validation)							
Class	Prior Prob	Wgt	Count	Count	Wgt	Misclass	Cost
0	0.500	101.00	101	101	7.00	7	0.069
		(101.00)	101	101	7.00	7	(0.069)
1	0.500	99.00	99	99	3.00	3	0.030
		(99.00)	99	99	6.00	6	(0.061)
Total	1.000	200.00	200	200	10.00	10	
		(200.00)	200	200	13.00	13	

**Lampiran F.**Hasil Ketepatan Klasifikasi (*Accuracy*)Akurasi Data *Learning**Cross Validation Tree 1*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	85	2	87			
1	6	87	93	0,956	0,977	0,935
<b>Total</b>	91	89	180			

*Cross Validation Tree 2*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	87	1	88			
1	4	88	92	0,972	0,989	0,957
<b>Total</b>	91	89	180			

*Cross Validation Tree 3*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	85	2	87			
1	6	87	93	0,956	0,977	0,935
<b>Total</b>	91	89	180			

*Cross Validation Tree 4*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	87	4	91			
1	3	86	89	0,961	0,956	0,966
<b>Total</b>	90	90	180			

**Lampiran F. (lanjutan)**

*Cross Validation Tree 5*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	86	2	88			
1	5	87	92	0,961	0,977	0,946
<b>Total</b>	91	89	180			

*Cross Validation Tree 6*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	88	2	90			
1	3	87	90	0,972	0,978	0,967
<b>Total</b>	91	89	180			

*Cross Validation Tree 7*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	86	1	87			
1	5	88	93	0,967	0,989	0,946
<b>Total</b>	91	89	180			

*Cross Validation Tree 8*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	87	2	89			
1	4	87	91	0,967	0,978	0,956
<b>Total</b>	91	89	180			

**Lampiran F. (lanjutan)***Cross Validation Tree 9*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	86	1	87	0,967	0,989	0,946
1	5	88	93			
<b>Total</b>	91	89	180			

*Cross Validation Tree 10*

Aktual	Prediksi		Total	Akurasi	Sensitivity	Specificity
	0	1				
0	85	2	87	0,956	0,977	0,935
1	6	87	93			
<b>Total</b>	91	89	180			

**Akurasi Data Testing**

=====

CROSS VALIDATION CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	0	1	
0	94.00	7.00	101.00
1	6.00	93.00	99.00
PRED. TOT.	100.00	100.00	200.00
CORRECT	0.931	0.939	
SUCCESS IND.	0.426	0.444	
TOT. CORRECT	0.935		

SENSITIVITY: 0.931 SPECIFICITY: 0.939

FALSE REFERENCE: 0.060 FALSE RESPONSE: 0.070

REFERENCE = "0", RESPONSE = "1"

-----

## Lampiran G. Informasi Simpul Terminal Pohon Klasifikasi Optimal

```

=====
NODE INFORMATION
=====
*****
*                               Node 1: X3                               *
*                               N: 200                               *
*****

*****
*                               Node 2                               *
*                               N: 141                               *
*****

Node 1 was split on X3
A case goes left if X3 = (0)
Improvement = 0.212                               Complexity Threshold = 0.298

Node   Cases  Wgt  Counts      Cost  Class
-----
  1       200    200.00   200.00   0.500  0
  2       141    141.00   141.00   0.288  0
 -6         59     59.00    59.00   0.000  1

Class      Top      Left      Right
-----
  0      101.00   101.00    0.00
  1       99.00    40.00   59.00

Class      Top      Left      Right
-----
  0       0.500   0.712   0.000
  1       0.500   0.288   1.000

Surrogate      Split      Assoc.      Improve.
-----
  1 X10         s 0         0.102       0.058
  2 X6          s 0         0.017       0.009

Competitor      Split      Improve.
-----
  1 X4           0         0.071
  2 X10          0         0.058
  3 X8           0         0.043
  4 X9           0         0.043
  5 X1          2380.000     0.040

*****
*                               Node 2: X4                               *
*                               N: 141                               *
*****

*****
*                               Node 3                               *
*                               N: 119                               *
*****

Node 2 was split on X4
A case goes left if X4 = (0)
Improvement = 0.075                               Complexity Threshold = 0.071

Node   Cases  Wgt  Counts      Cost  Class
-----
  2       141    141.00   141.00   0.288  0
  3       119    119.00   119.00   0.188  0
 -5         22     22.00    22.00   0.179  1

```

**Lampiran G. (lanjutan)**

```

Weighted Counts
Class      Top      Left      Right
0          101.00   97.00     4.00
1          40.00    22.00     18.00

Within Node Probabilities
Class      Top      Left      Right
0          0.712   0.812     0.179
1          0.288   0.188     0.821

Surrogate      Split      Assoc.      Improve.
1 X1           s 0      4250.000    0.091      0.010
2 X6           s 0              0.047      0.005

Competitor      Split      Improve.
1 X8            0              0.065
2 X9            0              0.065
3 X1            0      2380.000    0.065
4 X2            0              0.040
5 X10           0              0.035

*****
*                Node 3: X8                *
*                N: 119                    *
*****

*****
*                Node 4                *
*                N: 105                *
*****

=====
*                Terminal Node 4        *
*                N: 14                  *
=====

Node 3 was split on X8
A case goes left if X8 = (0)
Improvement = 0.088      Complexity Threshold = 0.061

Node      Cases  Wgt  Counts      Cost  Class
3         119   119.00  0.188 0
4         105   105.00  0.087 0
-4        14    14.00   0.070 1

Weighted Counts
Class      Top      Left      Right
0          97.00   96.00     1.00
1          22.00    9.00     13.00

Within Node Probabilities
Class      Top      Left      Right
0          0.812   0.913     0.070
1          0.188   0.087     0.930

Surrogate      Split      Assoc.      Improve.
1 X9           s 0              0.714      0.088
2 X7           s 0      148.000    0.215      0.021
3 X2           s 0              0.147      0.021
4 X1           r      1700.000  0.143      0.014
5 X5           r      35.550    0.073      0.007

Competitor      Split      Improve.
1 X9            0              0.088
2 X1            0      2380.000    0.049
3 X2            0              0.021
4 X7            0      148.000    0.021
5 X10           0              0.020
    
```

**Lampiran G. (lanjutan)**

```

*****
*                               *
*           Node 4: X1           *
*           N: 105              *
*                               *
*****
=====
= Terminal Node 1              = *           Node 5           *
=                               = *           N: 100           *
=====
*****

Node 4 was split on X1
A case goes left if X1 <= 2500.000
Improvement = 0.027              Complexity Threshold = 0.015

Node      Cases  Wgt  Counts      Cost  Class
4         105    5    105.00     0.087 0
-1        5     5     5.00      0.197 1
5         100    5    100.00     0.051 0

Weighted Counts
Class     Top      Left      Right
0         96.00    7.00     89.00
1          9.00    4.00     5.00

Within Node Probabilities
Class     Top      Left      Right
0         0.913   0.197    0.949
1         0.087   0.803    0.051

Competitor      Split      Improve.
1 X5            37.500     0.027
2 X9            0         0.017
3 X10           0         0.011
4 X7           131.000   0.004
5 X2            0         .701440E-03

*****
*                               *
*           Node 5: X5           *
*           N: 100              *
*                               *
*****
=====
= Terminal Node 2              = = Terminal Node 3           =
=                               = =           N: 3           =
=====
*****

Node 5 was split on X5
A case goes left if X5 <= 37.500
Improvement = 0.012              Complexity Threshold = 0.005

Node      Cases  Wgt  Counts      Cost  Class
5         100    5    100.00     0.051 0
-2        97     5    97.00      0.032 0
-3         3     3     3.00      0.329 1

Weighted Counts
Class     Top      Left      Right
0         89.00    86.00    3.00
1          5.00    3.00     2.00

Within Node Probabilities
Class     Top      Left      Right
0         0.949   0.968    0.329
1         0.051   0.032    0.671

```

**Lampiran G. (lanjutan)**

Competitor		Split		Improve.	
1 X9	0			0.009	
2 X10	0			0.004	
3 X7		129.000		0.002	
4 X1		3550.000		.724964E-03	
5 X2	1			.795152E-04	

=====

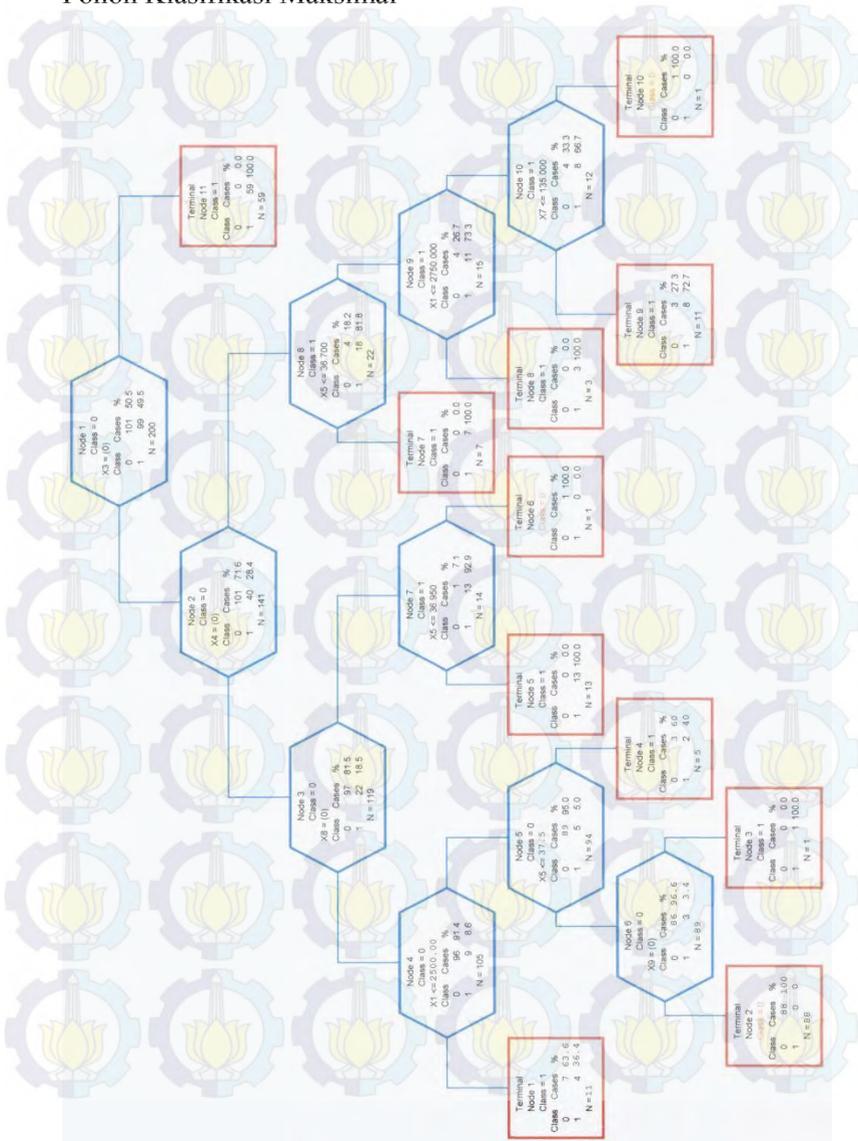
TERMINAL NODE INFORMATION

=====

[Breiman adjusted cost, lambda = 0.028]

Node	Class	Wgt	Count	N	Prob	Cost	Parent Complexity
1	1	5.00	5	5	0.025	0.197	0.015 [0.333]
0		1.00	1	1	0.197		
1		4.00	4	4	0.803		
2	0	97.00	97	97	0.480	0.032	0.005 [0.046]
0		94.00	94	94	0.968		
1		3.00	3	3	0.032		
3	1	3.00	3	3	0.015	0.329	0.005 [0.497]
0		1.00	1	1	0.329		
1		2.00	2	2	0.671		
4	1	14.00	14	14	0.071	0.070	0.061 [0.143]
0		1.00	1	1	0.070		
1		13.00	13	13	0.930		
5	1	22.00	22	22	0.111	0.179	0.071 [0.231]
0		4.00	4	4	0.179		
1		18.00	18	18	0.821		
6	1	59.00	59	59	0.298	0.000	0.298 [0.022]
0		0.00	0	0	0.000		
1		59.00	59	59	1.000		

## Lampiran H. Pohon Klasifikasi Maksimal



## Lampiran I. Pohon Klasifikasi Optimal



### Lampiran J.

#### Ilustrasi Perhitungan Skor Variabel Terpenting

Berikut ini merupakan ilustrasi dari nilai *goodness of split* (*improvement*) dan *surrogates* untuk setiap simpul yang memilah simpul anak pada pohon klasifikasi.

Simpul	Improvement	Surrogates
1	$X_3$ 0,2122	$X_{10}$ 0,0581
		$X_6$ 0,0094
2	$X_4$ 0,0748	$X_1$ 0,0104
		$X_6$ 0,0045
		$X_9$ 0,0883
3	$X_8$ 0,0883	$X_7$ 0,0205
		$X_2$ 0,0214
		$X_1$ 0,0136
		$X_5$ 0,0072
4	$X_1$ 0,0271	-
5	$X_5$ 0,0119	-

Sehingga skor yang dimiliki oleh masing-masing variabel yang berasal dari penjumlahan *improvement* dan *surrogates* pada setiap simpul.

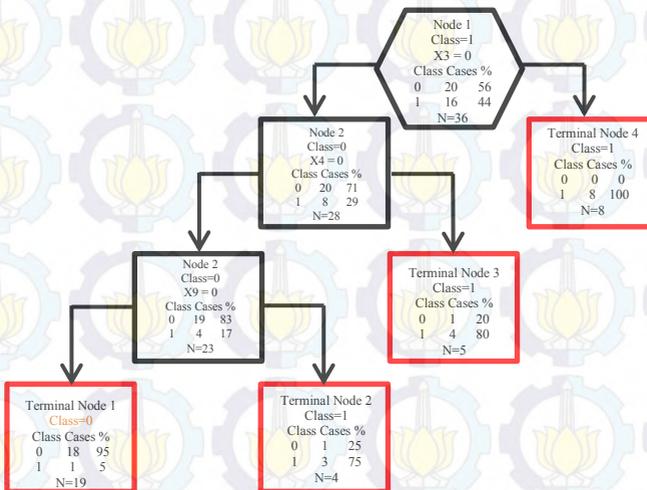
Variabel	Total <i>Improvement</i>	Skor
$X_1$	0,0511	24,081
$X_2$	0,0214	10,085
$X_3$	0,2122	100,000
$X_4$	0,0748	35,250
$X_5$	0,0191	9,001
$X_6$	0,0139	6,550
$X_7$	0,0205	9,661
$X_8$	0,0883	41,612
$X_9$	0,0883	41,612
$X_{10}$	0,0581	27,380

## Lampiran K.

### Ilustrasi 10-fold Cross Validation Estimate

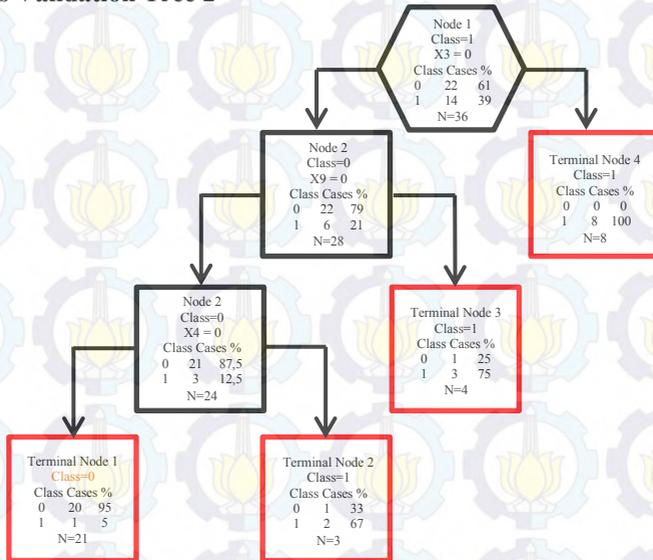
Misal diambil 50 data dari data pengamatan pada penelitian ini untuk diilustrasikan menggunakan metode *10-fold Cross Validation Estimate*. Kemudian 40 data diambil untuk membentuk pohon klasifikasi dan 10 data untuk memvalidasi bahwa pohon yang dihasilkan layak atau baik. Berikut ini ditunjukkan hasil pohon klasifikasi bentukan dari masing-masing *fold*.

#### Cross Validation Tree 1

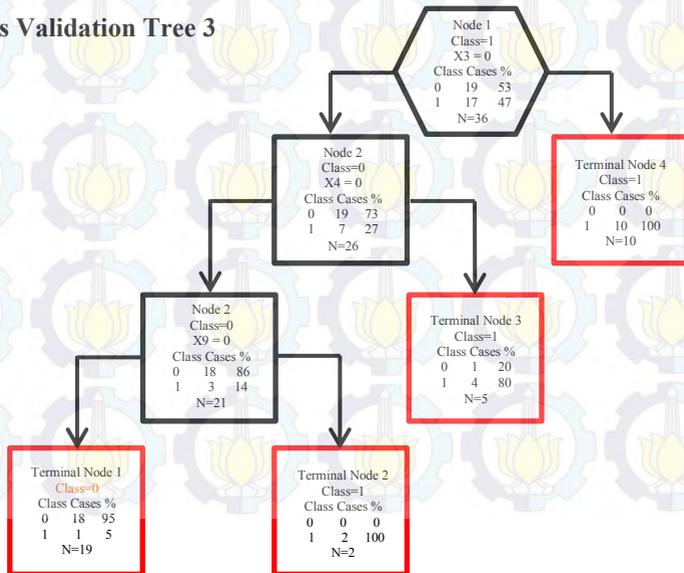


## Lampiran K. (lanjutan)

### Cross Validation Tree 2

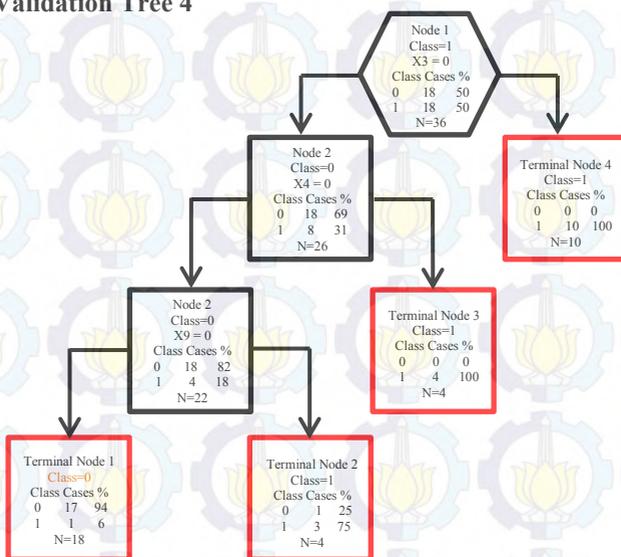


### Cross Validation Tree 3

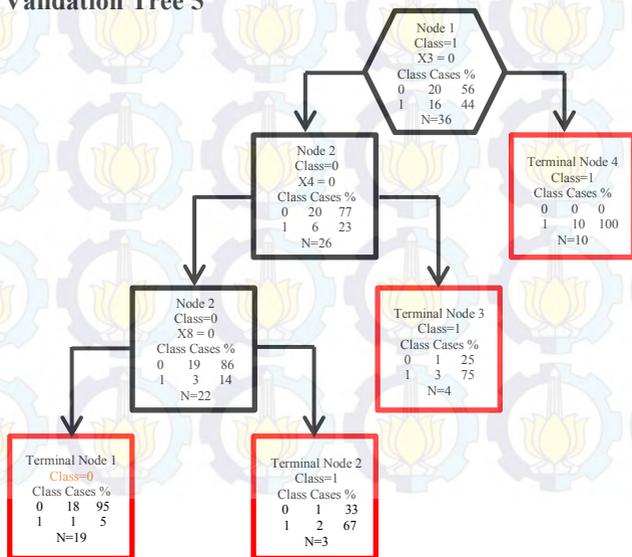


**Lampiran K. (lanjutan)**

**Cross Validation Tree 4**

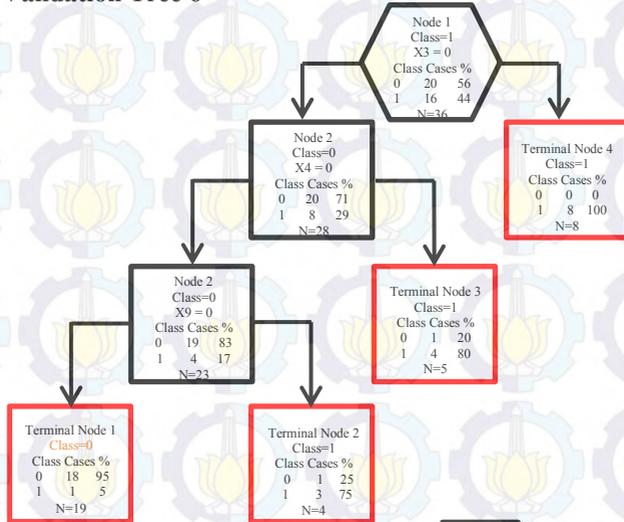


**Cross Validation Tree 5**

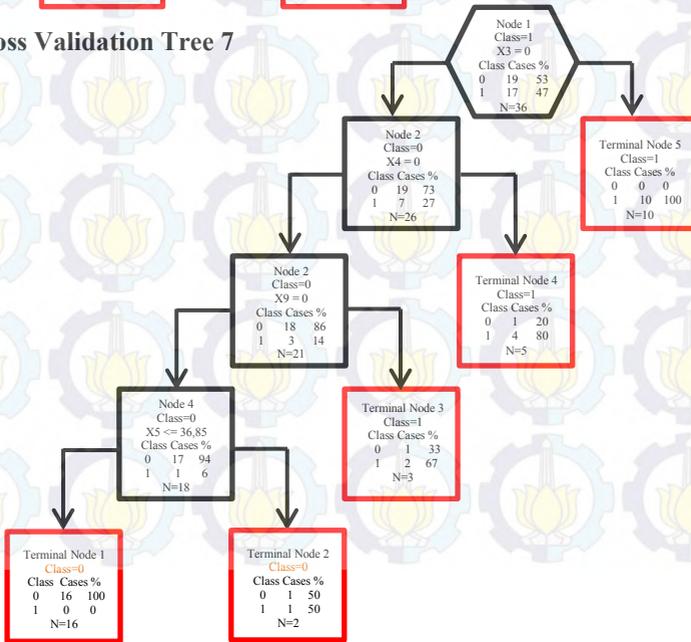


**Lampiran K. (lanjutan)**

**Cross Validation Tree 6**

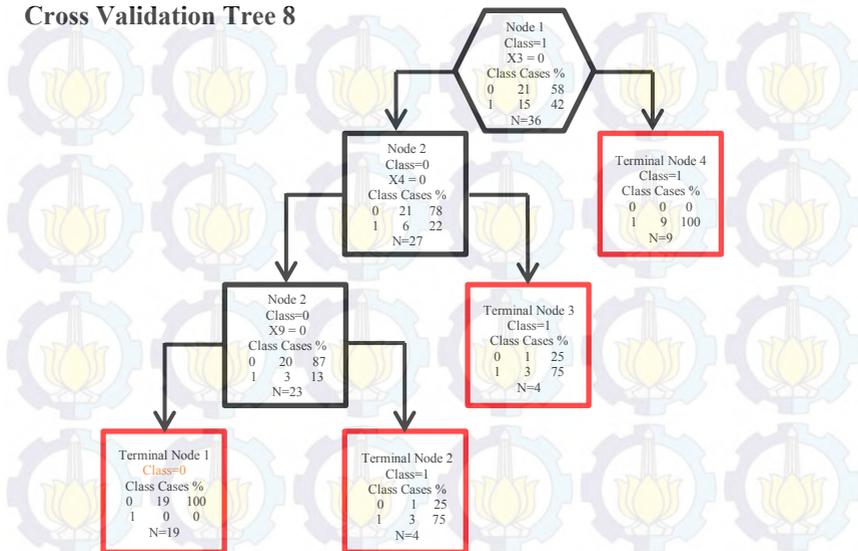


**Cross Validation Tree 7**

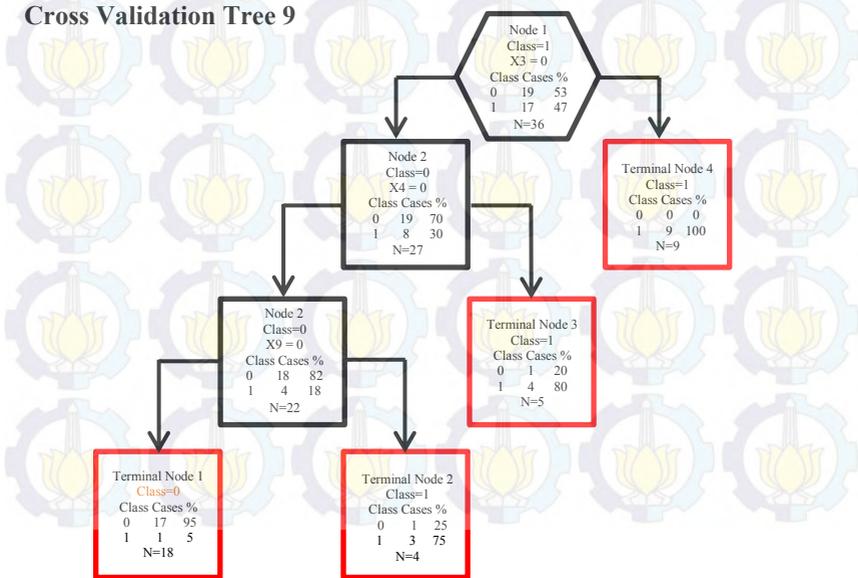


**Lampiran K. (lanjutan)**

**Cross Validation Tree 8**

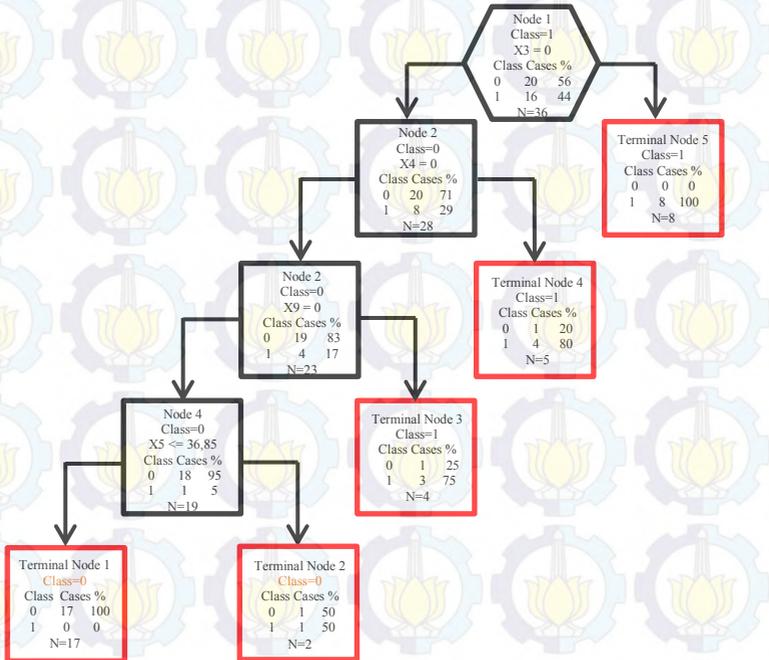


**Cross Validation Tree 9**



## Lampiran K. (lanjutan)

### Cross Validation Tree 10



Kemudian menjumlahkan hasil pengklasifikasian pada masing-masing *fold* dengan bentuk pohon klasifikasi yang berbeda-beda.

<i>Fold</i>	Simpul Terminal	$X(d^{(v)}(x))$	$R^{cv}(T_t^v)$
1	1	1	0,083333
	2	1	
	3	1	
	4	0	
2	1	1	0,083333
	2	1	
	3	1	
	4	0	

**Lampiran K. (lanjutan)**

<i>Fold</i>	Simpul Terminal	$X(d^{(v)}(x))$	$R^{cv}(T_r^v)$
3	1	1	0,055556
	2	0	
	3	1	
	4	0	
4	1	1	0,055556
	2	1	
	3	0	
	4	0	
5	1	1	0,083333
	2	1	
	3	1	
	4	0	
6	1	1	0,083333
	2	1	
	3	1	
	4	0	
7	1	0	0,083333
	2	1	
	3	1	
	4	1	
	5	0	
8	1	0	0,055556
	2	1	
	3	1	
	4	0	
9	1	1	0,083333
	2	1	
	3	1	
	4	0	
10	1	0	0,083333
	2	1	
	3	1	
	4	1	
	5	0	
<b><math>R^{cv}(T_r)</math></b>			<b>0,075</b>

**Lampiran L.****Hasil Klasifikasi Data Rekam Medis Bayi Baru Lahir**

Pasien	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y	Status Risiko Infeksi
1	2570	0	1	0	36,8	0	120	0	0	0	1	Positif
2	3500	0	1	0	38	0	148	0	0	0	1	Positif
3	3100	0	0	0	36,9	0	124	0	0	0	0	Negatif
4	3300	0	0	0	36,8	0	124	0	0	0	0	Negatif
5	1600	1	0	0	36,5	0	136	1	1	0	1	Positif
6	3100	0	0	0	36,5	0	120	0	0	0	0	Negatif
7	3200	0	0	0	36,7	0	142	0	0	0	0	Negatif
8	3000	0	0	0	36,8	0	120	0	0	0	0	Negatif
9	2900	0	1	1	36,5	1	136	0	0	1	1	Positif
10	3500	0	0	0	36,8	0	140	0	0	0	0	Negatif
11	3400	0	0	0	36	0	130	0	0	0	0	Negatif
12	3200	0	1	0	36,8	0	140	0	0	0	1	Positif
13	4900	0	1	0	36,8	0	120	0	0	0	1	Positif
14	2830	0	0	0	36	0	120	0	0	0	0	Negatif
15	2600	0	0	1	36,3	0	136	0	0	0	1	Positif
16	3180	0	0	0	36,9	0	120	0	0	0	0	Negatif
17	2600	1	1	0	36,6	0	140	0	1	0	1	Positif
18	3800	0	1	0	39,5	0	160	0	1	0	1	Positif
19	3500	0	1	0	36,2	0	140	0	0	0	0	Positif
20	3000	0	0	0	36,7	0	120	0	0	0	0	Negatif
21	2900	0	0	0	36	0	144	0	0	0	0	Negatif
22	3700	0	0	0	36,5	0	136	0	0	0	0	Negatif
23	3000	0	0	0	36,8	0	120	0	0	0	0	Negatif
24	2400	0	0	0	36,5	0	120	0	0	0	0	Positif
25	3700	0	0	0	36,8	0	120	0	0	0	0	Negatif
26	3000	1	1	0	36,9	0	124	0	0	1	1	Positif
27	3000	0	0	0	36,8	0	120	0	0	0	0	Negatif
28	3100	0	0	0	36,7	0	125	0	0	0	0	Negatif
29	3500	0	1	1	36,5	0	150	0	0	0	1	Positif
30	2900	0	0	0	36	0	120	0	0	0	0	Negatif
31	3600	0	0	0	36,8	0	120	0	0	0	0	Negatif
32	3500	0	0	1	36,7	0	138	0	0	0	1	Positif

**Lampiran L. (lanjutan)**

Pasien	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y	Status Risiko Infeksi
33	3500	0	0	0	36,8	0	120	0	0	0	0	Negatif
34	2300	1	1	1	36,4	0	142	0	0	0	1	Positif
35	2500	1	1	1	36,7	0	120	0	0	1	1	Positif
36	2100	1	1	1	36,2	0	142	0	0	0	1	Positif
37	3100	0	0	1	36	0	140	0	0	0	0	Positif
38	3200	0	0	0	36,6	0	140	0	0	0	0	Negatif
39	3200	0	0	0	36,8	0	124	0	0	0	0	Negatif
40	2600	0	1	0	36,9	0	130	0	0	0	1	Negatif

**Lampiran M.****Surat Keterangan Penelitian di RSUD Sidoarjo**

 <p><b>PEMERINTAH KABUPATEN SIDOARJO</b>  <b>RUMAH SAKIT UMUM DAERAH</b>          Jl. Mojopahit No. 667 Telepon (031) 8961649. Fax. 8943237  <b>SIDOARJO - Kode Pos 61215</b></p>		
Sidoarjo, 23 Nopember 2015		
Kepada		
Nomor	: 070/ 3341 /404.6.8/2015	Yth. Ketua Jurusan Statistika ITS Surabaya
Sifat	: Penting	Kampus ITS Sukolilo Surabaya
Lampiran	: -	Di
Perihal	: Jawaban Permohonan Ijin Penelitian	<b>SURABAYA</b>
<p>Menindak lanjuti surat Saudara tanggal 3 Nopember 2015 nomor : 2147/IT2.1.L3/KM.05.00/2015 perihal : Permohonan Ijin Penelitian, bersama ini disampaikan bahwa pada prinsipnya kami tidak keberatan dan dapat menyetujui permohonan ijin Saudara.</p> <p>Sebagai tindak lanjut tentang tarif pelayanan kesehatan Rumah Sakit Umum Daerah Kabupaten Sidoarjo yang menerapkan pola pengelolaan keuangan badan layanan umum daerah, untuk biaya pemanfaatan Rumah Sakit Umum Daerah Kabupaten Sidoarjo sebagai tempat pelatihan, PKL, Penelitian dan lain-lain, maka setiap mahasiswa non kesehatan yang melaksanakan Penelitian dikenakan biaya sebesar Rp. 150.000.- (<i>Seratus Lima Puluh Ribu Rupiah</i>) masing - masing per bulan atas nama:</p> <p>1. AULIA RAHMA SAFITRI <span style="float: right;">NIM : 1312 100 066</span></p> <p>Demikian atas perhatiannya kami ucapkan terima kasih.</p> <p>an. DIREKTUR RUMAH SAKIT UMUM DAERAH          KABUPATEN SIDOARJO          Wadiv Perencanaan dan Pendidikan</p> <p style="text-align: center;">   <b>drg. SYAFI Satriawarman, Sp. Pros</b>          Pembina FK I          NIP. 19630718 199103 1 004</p>		



*“Halaman ini sengaja dikosongkan”*

## BIODATA PENULIS



Penulis bernama Aulia Rahma Safitri yang akrab disapa dengan Aulia, merupakan anak dari pasangan Sjaiful Muttaqien dan Nur Fitri. Penulis merupakan anak kedua dari tiga bersaudara yang dilahirkan di kota Surabaya pada tanggal 10 Desember 1993. Penulis memulai jenjang pendidikan yang pertama di Taman kanak-kanak Aisyiyah Gadung (1999-2000), kemudian melanjutkan di SD Muhammadiyah 6 Gadung, Surabaya (2000-2003) dan SD Muhammadiyah 1 Sidoarjo (2003-2006). Setelah menamatkan pendidikan sekolah dasar penulis melanjutkan pendidikan di SMPN 3 Sidoarjo (2006-2009) dan di SMAN 1 Sidoarjo (2009-2012). Pada tahun 2012, penulis melanjutkan pendidikan di Perguruan Tinggi Negeri Institut Teknologi Sepuluh Nopember dan diterima sebagai mahasiswi Jurusan Statistika dengan NRP 1312100066. Selama kuliah, penulis juga aktif dalam kegiatan organisasi dan pernah menjadi staff Sekretaris HIMASTA-ITS pada tahun 2013-2014, kemudian pada tahun 2014-2015 penulis menjabat sebagai Sekretaris Umum HIMASTA-ITS. Selain itu penulis juga aktif dalam kepanitiaan seperti panitia PRS (Pekan Raya Statistika) dan panitia YELP (*Youth Enviromental Leader Program*). Pengalaman magang, penulis memiliki pengalaman magang di PT. Gajah Tunggal Tbk., pabrik ban ternama yang berlokasi di Tangerang.

Penulis menerima segala kritikan, masukan, dan saran yang bersifat membangun demi meningkatkan manfaat Tugas Akhir ini. Oleh karena itu, penulis siap membantu terkait dengan metode dan ilmu statistik, jika dibutuhkan pembaca dapat menghubungi penulis melalui email : [arahma10@gmail.com](mailto:arahma10@gmail.com).



*“Halaman ini sengaja dikosongkan”*