



TUGAS AKHIR - SS141501

EVALUASI KETEPATAN KLASIFIKASI PENYAKIT ISPA (INFEKSI SALURAN PERNAFASAN AKUT) PADA ANAK BALITA MENGGUNAKAN METODE *CLASSIFICATION TREES* DI PUSKESMAS CUKIR KABUPATEN JOMBANG JAWA TIMUR

**ARIF BUDHIMAN
NRP 1315 105 039**

**Dosen Pembimbing
Dr. Bambang Widjanarko Otok, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

EVALUASI KETEPATAN KLASIFIKASI PENYAKIT ISPA (INFEKSI SALURAN PERNAFASAN AKUT) PADA ANAK BALITA MENGGUNAKAN METODE *CLASSIFICATION TREES* DI PUSKESMAS CUKIR KABUPATEN JOMBANG JAWA TIMUR

**ARIF BUDHIMAN
NRP 1315 105 039**

**Dosen Pembimbing
Dr. Bambang widjanarko Otok, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS141501

**EVALUATION OF CLASSIFICATION ACCURACY IN CASE
OF INFECTION ACUTE RESPIRATORY ON TODDLER
USING *CLASSIFICATION TREES* AT CUKIR PUSKESMAS
DISTRICT OF JOMBANG EAST JAVA**

**ARIF BUDHIMAN
NRP 1315 105 039**

**Supervisor
Dr. Bambang Widjanarko Otok, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

**EVALUASI KETEPATAN KLASIFIKASI ISPA
(INFEKSI SALURAN PERNAFASAN AKUT)
PADA ANAK BALITA MENGGUNAKAN
METODE *CLASSIFICATION TREES* DI
PUSKESMAS CUKIR KABUPATEN JOMBANG
JAWA TIMUR**

TUGAS AKHIR

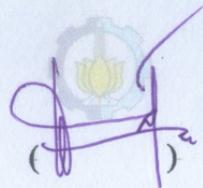
Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

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

Oleh :

Arif Budhiman
NRP. 1315 105 039

Disetujui oleh Pembimbing:
Dr. Bambang Widjanarko Otok, M.Si.
NIP. 19681124 199412 1 001



Mengetahui,
Kepala Departemen



Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JANUARI 2018

(Halaman ini sengaja dikosongkan)

Evaluasi Ketepatan Klasifikasi Penyakit ISPA (Infeksi Saluran Pernafasan Akut) Pada Anak Balita Menggunakan Metode *Classification Trees* Di Puskesmas Cukir Kabupaten Jombang

Nama Mahasiswa : Arif Budhiman
NRP : 1315 105 039
Departemen : Statistika
Dosen Pembimbing : Dr. Bambang Widjanarko Otok, M.Si

Abstrak

Infeksi Saluran Pernafasan Akut (ISPA) adalah penyakit infeksi pada saluran pernafasan terutama mengenai struktur saluran pernafasan. Penyakit ISPA merupakan masalah kesehatan tidak boleh diabaikan karena menyebabkan kematian bayi dan balita yang tinggi. Rencana Strategis (Renstra) Puskesmas di Kabupaten Jombang Jawa Timur merupakan suatu proses secara sistematis. Pada analisis Classification Trees variabel yang berpengaruh signifikan adalah variabel Suhu Tubuh. Berdasarkan hasil tersebut diketahui bahwa nilai skor variabel Suhu Tubuh memiliki nilai kontribusi terbesar yaitu 100 yang artinya variabel X_4 (Suhu Tubuh) akan menjadi pemilah awal atau sebagai simpul induk (parent nodes). Jika dilihat nilai relative cost pohon klasifikasi dengan simpul terminal sebanyak 9 simpul sudah dapat dikatakan optimal ditunjukkan dengan nilai relative cost sebesar 0.230 dan nilai kompleksitas 0.010. Hasil Ketepatan Klasifikasi Classification Trees dengan menggunakan metode 10-fold cross validation diketahui bahwa terdapat 103 anak balita yang termasuk dalam kategori ISPA (terdiagnosa penyakit ISPA) dengan 89 tepat diklasifikasikan sebagai anak balita dengan terdiagnosa penyakit ISPA dan 14 anak balita salah diklasifikasikan sebagai anak balita yang tidak terdiagnosa penyakit ISPA. Sehingga nilai sensitiviy yang diperoleh sebesar 0.864. Sementara itu jumlah anak balita yang tidak termasuk dalam kategori terdiagnosa penyakit ISPA sebanyak 53 anak balita, dimana 48 anak balita tepat diklasifikasikan Non ISPA (tidak terdiagnosa ISPA) dan 5 tepat disalahkan sebagai terdiagnosa ISPA sehingga diperoleh nilai specificity sebesar 0.906. Total akurasi yang dihasilkan untuk klasifikasi terdiagnosa pemyakit ISPA pada anak balita adalah 87.8%. Sedangkan total kesalahan klasifikasi (APER) yang dihasilkan 12.2%. Karena total akurasi cukup tinggi maka pohon klasifikasi optimal yang terbentuk sudah layak digunakan untuk pengklasifikasian.

Kata Kunci: *Classification Trees, ISPA, Relative Cost*

(Halaman ini sengaja dikosongkan)

EVALUATION OF CLASSIFICATION ACCURACY IN CASE OF INFECTION ACUTE RESPIRATORY ON TODDLER USING *CLASSIFICATION TREES* AT CUKIR PUSKESMAS DISTRICT OF JOMBANG EAST JAVA

Student Name : Arif Budhiman
Student Number : 1315 105 039
Department : Statistics
Supervisor : Dr. Bambang Widjanarko Otok, M.Si

Abstract

Acute Respiratory Infection (ARI) is an infectious respiratory tract infection, especially regarding the structure of the respiratory tract. It should be noted that ARI disease is a health problem should not be ignored because it causes high infant and toddler mortality. Strategic Plan Puskesmas in Jombang East Java is a systematic process. In the analysis of Classification Trees is known that the variables that significantly influence the variable temperature. Based on these results it is known that the value score the variable temperature rate getting result the biggest value contribution is 100 which variable X_4 (temperature) will be first sorting or as parent nodes. It can be looked value relative cost of classification trees with terminal node is 9 node have been optimum, showed with value relative cost is 0.230. In result of classification trees with using 10-fold cross validation method is known that there 103 toddler which include Acute Respiratory Infection (ARI) with 89 exactly classified as toddler which exposed Acute Respiratory Infection (ARI) disease and 14 toddler false classified as toddler which not exposed Acute Respiratory Infection (ARI) disease. So that value sensity which obtained is 0.864. While number of toddler which not exposed Acute Respiratory Infection (ARI) disease is 53 toddler, so that obtained value of specifity is 0.906. Value accuracy which resulted for classification Acute Respiratory Infection (ARI) is 87.8%. While misclassification (APER) obtained 12.2%. Because value accuracy so high enough, so classification trees optimum which formed worth it using clasifcation.

Keywords: *Acute Respiratory Infection , Classification Trees, Relative Cost*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur Alhamdulillah senantiasa penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul ***“Evaluasi Ketepatan Klasifikasi ISPA (Infeksi Saluran Pernafasan Akut) pada Anak Balita Menggunakan Metode Classification Trees di Puskesmas Cukir Kabupaten Jombang”*** dengan lancar dan tepat waktu.

Keberhasilan penyusunan Tugas Akhir ini tidak lepas dari partisipasi berbagai pihak yang telah banyak membantu. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Bambang Widjanarko Otok, M.Si selaku dosen pembimbing atas semua bimbingan, waktu, semangat dan perhatian yang telah diberikan sehingga Tugas Akhir ini dapat diselesaikan dengan baik.
2. Bapak Dr. Sutikno, M.Si dan Bapak Dr. Puhadi, M.Sc selaku tim penguji yang telah memberikan saran dan kritik yang membangun dalam kesempurnaan Tugas Akhir ini.
3. Bapak Dr. Suhartono, S.Si, M.Sc selaku Kepala Departemen Statistika ITS yang telah memberikan fasilitas dalam kelancaran Tugas Akhir ini
4. Bapak Dr. Sutikno, M.Si dan Ibu Dr. Santi Wulan Purnami, M.Si selaku Ketua Program Studi S1 dan Sekretaris Program Studi S1 yang mengawal proses berjalannya Tugas Akhir Mahasiswa S1 dengan bimbingan serta fasilitas yang diberikan.
5. Ibu Erma Oktania Permatasari, S.Si, M.Si selaku Dosen Wali penulis, seluruh dosen, dan karyawan Statistika ITS atas ilmu dan pengalaman yang telah diberikan kepada penulis.
6. Ibunda serta Ayahanda, atas semangat, kasih sayang dan doa yang tidak pernah putus kepada penulis.
7. Kakak tercinta yang tidak pernah berhenti memberi perhatian kepada penulis.
8. Keluarga besar Sunyoto yang selalu mengingatkan dan memberi dukungan kepada penulis

9. Mahasiswa Departemen Statistika Lintas Jalur Angkatan 2015, 2016 dan Reguler Angkatan 2014.

10. Semua pihak yang telah membantu penulis yang tidak dapat disebutkan satu persatu.

Semoga kebaikan dan bantuan yang telah diberikan kepada penulis dibalas dengan kebaikan yang lebih oleh Allah SWT. Amin.

Penulis menyadari bahwa Tugas Akhir ini masih terdapat kekurangan, oleh karena itu kritik dan saran yang bersifat membangun sangat diharapkan. Semoga Tugas Akhir ini dapat memberikan manfaat baik bagi penulis, pembaca, dan semua pihak.

Surabaya, Januari 2018

Penulis

Arif Budhiman

DAFTAR ISI

	Halaman
JUDUL	i
COVER PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	7
BAB II TINJAUAN PUSTAKA	
2.1 <i>Classification and Regression Trees</i>	9
2.2 Ukuran Ketepatan Klasifikasi.....	18
2.3 Tinjauan Non Statistika.....	19
2.3.1. ISPA.....	19
2.3.2. Proses Terjadinya ISPA.....	21
2.3.3. Penyebab Penyakit ISPA.....	22
2.3.4. Penularan ISPA.....	22
2.3.5. Rumah Sehat dan Faktor Lingkungan.....	22
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	27
3.2 Kerangka Konsep.....	27
3.3 Variabel Penelitian.....	28
3.4 Definisi Operasional.....	28
3.5 Langkah Analisis <i>Classification Trees</i>	30
3.6 Diagram Alir Analisis <i>Classification Trees</i>	33

BAB IV ANALISIS DATA DAN PEMBAHASAN

4.1 Karakteristik Anak Balita Terhadap Penyakit ISPA di Puskesmas Cukir Jombang	35
4.1.1 Karakteristik Penyakit ISPA Berdasarkan Riwayat Pemberian ASI	36
4.1.2 Karakteristik Penyakit ISPA Berdasarkan Berat Badan	36
4.1.3 Karakteristik Penyakit ISPA Berdasarkan Lingkar Kepala	37
4.1.4 Karakteristik Penyakit ISPA Berdasarkan Suhu Tubuh	38
4.1.5 Karakteristik Penyakit ISPA Berdasarkan Denyut Nadi	38
4.1.6 Karakteristik Penyakit ISPA Berdasarkan Respiratory Rate	39
4.1.7 Karakteristik Penyakit ISPA Berdasarkan Status Perekonomian Keluarga	40
4.2 Analisis <i>Classification Trees</i> untuk klasifikasi Diagnosa ISPA pada Anak Balita	41
4.2.1 Pembentukan Pohon Klasifikasi	42
4.2.2 Pemangkasan Pohon Klasifikasi	44
4.2.3 Penentuan Pohon Klasifikasi Optimal	45
4.3 Hasil Ketepatan Klasifikasi <i>Classification Trees</i> ...	49

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan	51
5.2 Saran	52

DAFTAR PUSTAKA	53
-----------------------------	----

LAMPIRAN	55
-----------------------	----

DAFTAR GAMBAR

	Halaman
Gambar 2.1	Ilustrasi Struktur Pohon Klasifikasi..... 11
Gambar 2.2	Ilustrasi Prosedur <i>10-fold Cross Validation</i> 18
Gambar 3.1	Kerangka Konsep 27
Gambar 3.2	Diagram Alir 33
Gambar 4.1	Deskripsi Pola Pemberian ASI Berdasarkan Status Risiko ISPA 36
Gambar 4.2	Deskripsi Tingkat Pernafasan Berdasarkan Status Penyakit ISPA..... 40
Gambar 4.3	Deskripsi Status Perekonomian Keluarga Berdasarkan Status Risiko Penyakit ISPA..... 40
Gambar 4.4	Topologi Pohon Klasifikasi Maksimal 43
Gambar 4.5	Plot <i>Relative Cost</i> dan Banyaknya Simpul Terminal..... 44
Gambar 4.6	Topologi Pohon Klasifikasi Optimal dengan Pemilahan Pemilah Indeks Gini..... 46

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 Tabulasi Silang Ketepatan Klasifikasi.....	18
Tabel 3.1 Variabel Penelitian	28
Tabel 4.1 Deskripsi Penyakit ISPA Berdasarkan Berat Badan.....	37
Tabel 4.2 Deskripsi Penyakit ISPA Berdasarkan Lingkar Kepala.....	37
Tabel 4.3 Deskripsi Penyakit ISPA Berdasarkan Suhu Tubuh.	38
Tabel 4.4 Deskripsi Penyakit ISPA Berdasarkan Denyut Nadi	39
Tabel 4.5 Skor Variabel prediktor dalam pembentukan Pohon Klasifikasi Maksimal.....	42
Tabel 4.6 Urutan Pembentukan Pohon Klasifikasi (<i>Tree Sequence</i>).....	45
Tabel 4.7 Variabel Penting Pembentukan Pohon Klasifikasi Optimal	46
Tabel 4.8 Pelabelan Kelas Simpul Terminal	47
Tabel 4.9 Karakteristik Anak Balita Berdasarkan Simpul Terminal	49
Tabel 4.10 Ketepatan Klasifikasi Risiko Penyakit ISPA pada Anak Balita.....	50

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran A. Data Kasus ISPA	55
Lampiran B. Variabel Prediktor Berskala Kontinyu	57
Lampiran C. Statistika Deskriptif antara variabel respon dan prediktor.....	58
Lampiran D Output Pohon Klasifikasi <i>10-Fold Cross Validation Estimate</i>	59
Lampiran E Hasil Ketepatan Klasifikassi.....	63
Lampiran F Informasi Simpul Terminal Pohon Klasifikasi.....	69
Lampiran G Surat Perizinan Pengambilan Data	74
Lampiran H Rekam Medik Rawat Jalan Puskesmas Cukir	75

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Infeksi Saluran Pernapasan Akut (ISPA) adalah infeksi yang mengganggu proses pernafasan yang disebabkan oleh virus atau bakteri yang menyerang hidung, pipa pernafasan, sinus, faring, laring bahkan paru-paru. Infeksi Saluran Pernapasan Akut tercatat sebagai penyakit yang paling banyak diderita oleh masyarakat khususnya para balita dan anak-anak yang biasanya mengalami gejala sakit batuk pilek setidaknya tiga hingga enam kali pertahunnya.

Penyakit ISPA tergolong penyakit yang menular melalui udara, kontak dengan orang yang terinfeksi dan dapat terkontaminasi dengan membawa virus dan bakteri. Jika penyakit ISPA tidak segera ada penanganan khusus akan menjadi penyakit yang berat karena masuk ke jaringan paru-paru dan akan menyebabkan pneumonia, sehingga berdampak kematian kepada orang yang mengalaminya terutama usia anak balita. *Period prevalence* infeksi Saluran Pernapasan Akut (ISPA) berdasarkan vonis dan diagnosis tenaga kesehatan dan keluhan penduduk sebesar 25 persen. Menurut riset kesehatan dasar, Daerah provinsi yang terjangkit penyakit ISPA tertinggi adalah di provinsi Jawa Timur dan NTT pada tahun 2007. Sedangkan untuk di daerah kabupaten Jombang menurut data dinas kesehatan kabupaten Jombang penyakit yang banyak dialami oleh warga jombang yaitu penyakit flu dan ISPA sepanjang tahun 2015 hingga pertengahan September tahun 2016. Di Puskesmas Jombang, sejumlah pasien ISPA didominasi oleh anak-anak yang menjalani rawat inap. Sementara itu pihak puskesmas mengakui ISPA menjadi tren penyakit dimusim kemarau. Penyebabnya yakni faktor cuaca yang panas, angin hingga polusi dari kendaraan bermotor.

Negara Indonesia sebagai daerah Negara tropis yang sangat berpotensi menjadi daerah endemik dari beberapa penyakit infeksi

yang setiap saat dapat menjadi ancaman bagi kesehatan masyarakat Indonesia. Pengaruh geografis dapat mendorong terjadinya peningkatan kasus maupun kematian akibat penderita ISPA, misalnya saja pencemaran lingkungan yang disebabkan oleh asap yang disebabkan oleh kebakaran hutan, gas pembuangan transportasi dan polusi udara dalam rumah karena asap dapur, asap rokok, perubahan iklim global antara lain perubahan suhu udara, kelembaban dan curah hujan merupakan ancaman kesehatan terutama pada penyakit ISPA.

Masalah utama anak balita pada masa terjangkau penyakit ISPA dapat diketahui dengan berbagai gejala yang sering dianggap biasa dari orang tua yaitu pilek, batuk yang berkelanjutan dan sesak nafas yang dapat menjadikan ancaman serius bahkan sampai kematian jika tidak ditangani dengan serius maupun intensif. Kasus ISPA yang dialami pada anak balita merupakan salah satunya penyebab kematian. Umumnya kasus Infeksi Saluran Pernafasan Akut menunjukkan gejala batuk yang biasa dan seringkali tidak diketahui sampai keadaannya sudah sangat terlambat. Oleh sebab itu, tenaga kesehatan perlu melakukan pencegahan infeksi melalui perawatan intensif terhadap anak balita yang memiliki resiko ISPA. Risiko ISPA 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 anak balita yang berisiko ISPA merupakan hal yang penting. Melalui diagnosa tersebut anak balita yang memiliki risiko ISPA dapat diberikan perawatan dan prosedur *pulse oximetry* yang bertujuan untuk memeriksa seberapa banyak oksigen yang masuk ke paru-paru dan biasanya dilakukan pada anak balita yang mengalami kesulitan bernafas.

Selain itu dokter mungkin akan menyarankan untuk melakukan pengambilan sampel dahak untuk diperiksa dilaboratorium. Pemeriksaan ini bertujuan untuk menentukan jenis virus atau bakteri penyebab ISPA. Apabila infeksi dicurigai telah masuk ke dalam paru-paru, maka pemeriksaan dengan X-Ray atau CT Scan mungkin akan direkomendasikan oleh dokter. Kedua jenis

pemeriksaan ini dilakukan untuk mengamati kondisi paru-paru pada anak balita.

Karakteristik anak balita perlu diketahui melalui pemeriksaan essensial yang dilakukan tenaga kesehatan guna menetapkan diagnosa terhadap anak balita tersebut apakah berisiko infeksi atau tidak. Hal tersebut dapat diwujudkan dengan mengklasifikasi anak balita yang berisiko ISPA atau tidak berisiko. Metode pengklasifikasian yang tepat yaitu metode *Classification Trees*. Metode *Classification Trees* merupakan bagian dari metode CART (*Classification* dan *Regression Trees*). Metode CART adalah metode nonparametrik yang dapat memilih variabel dan interaksi yang paling mempengaruhi variabel respon. Jika variabel respon yang dipelajari skala 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). Metode CART digunakan untuk menggambarkan hubungan antara variabel respon (variabel dependen atau tak bebas) dengan satu atau lebih variabel prediktor (variabel independen atau bebas). Kedua metode ini merupakan metode yang bisa diterapkan untuk data jumlah besar, variabel yang sangat banyak dengan skala variabel campuran melalui prosedur pemilahan biner. Akan tetapi CART juga memiliki kelemahan yaitu menghasilkan pohon yang kurang stabil karena CART sangat sensitif dengan data baru, bergantung dengan jumlah sampel. Jika sampel data *training* dan *testing* berubah maka pohon keputusan yang dihasilkan juga ikut berubah. (Pratiwi & Zain, 2014).

Penelitian sebelumnya mengenai risiko penyakit ISPA dilakukan oleh Bagas (2017) yang meneliti tentang faktor-faktor yang menyebabkan ISPA pada anak balita menggunakan *Propensity Score Stratification* di Puskesmas yang bertempat di Kabupaten Pasuruan, Jawa Timur. Hasil penelitian tersebut menyimpulkan bahwa faktor-faktor yang mempengaruhi risiko penyakit ISPA yaitu status perokok pasif dan status imunisasi anak.

Sedangkan penelitian penggunaan metode CART yang dilakukan oleh Aulia (2016) mengenai klasifikasi penyakit risiko infeksi pada bayi baru lahir (BBL) menghasilkan akurasi penyakit infeksi menggunakan pemilah indeks gini berdasarkan pohon klasifikasi optimal dengan variabel terpenting kondisi air ketuban sebesar 100 % dan simpul terminal yang dihasilkan sebanyak 6 simpul. Penelitian juga 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. Dan 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. Mengamati dari penelitian-penelitian sebelumnya, maka dalam penelitian ini menggunakan metode CART, khususnya *Classification Trees* untuk mendapatkan klasifikasi risiko penyakit ISPA pada anak balita.

Studi kasus yang digunakan dalam penelitian ini merupakan studi kasus dari Puskesmas Cukir yang bertempat di daerah Jombang, Jawa Timur. Karena Puskesmas Jombang merupakan tempat yang menjadi tujuan utama bagi anak-anak balita untuk melakukan pemeriksaan di kabupaten Jombang. Kasus ISPA sendiri merupakan salah satu faktor penyebab tingginya kematian anak balita di Jombang setelah demam berdarah (DBD). Oleh sebab itu, untuk menurunkan kasus ISPA di Kabupaten Jombang perlu diketahui faktor yang paling berpengaruh terhadap kasus ISPA pada anak balita tersebut. Selain itu dilakukan

pula pengklasifikasian berisiko terkena ISPA dan tidak berisiko terkena ISPA pada anak balita menggunakan metode *Classification Trees* untuk mengetahui perbedaan karakteristik diantara dua kategori pada balita tersebut berdasarkan kriteria kondisi anak balita. Tujuan dari *Classification Trees* adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian. (Lewis, 2000).

1.2 Rumusan Masalah

Menurunkan angka penyakit ISPA merupakan salah satu tujuan pembangunan kesehatan Indonesia untuk meningkatkan derajat kesehatan dan kualitas sumber daya manusia. Penyakit ISPA tidak bisa dianggap biasa, karena jika tidak ditangani serius akan mengakibatkan kematian pada seseorang yang mengalaminya. Oleh sebab itu, perlu adanya penanganan lebih lanjut untuk menurunkan kasus penyakit ISPA pada anak balita melalui pelayanan kesehatan yang serius dan tindakan medis yang tepat pada anak balita yang berisiko ISPA. Upaya yang dapat dilakukan untuk menurunkan kasus penyakit ISPA yaitu dengan mengetahui faktor paling dominan yang menyebabkan kasus infeksi penyakit ISPA pada anak balita serta melakukan pengklasifikasian anak balita berdasarkan status risiko infeksi untuk mengetahui perbedaan karakteristik anak balita dari kedua kategori status risiko penyakit ISPA tersebut. Hal tersebut dapat diketahui menggunakan metode *Classification Trees* yang menghasilkan pohon klasifikasi. Pohon klasifikasi tersebut memilah anak balita yang terdiagnosa dan tidak terdiagnosa penyakit ISPA berdasarkan faktor-faktor yang mempengaruhi.

Berdasarkan uraian tersebut, maka digunakan klasifikasi dengan pendekatan nonparametrik yaitu pohon klasifikasi dan permasalahan yang akan dibahas dalam penelitian ini adalah bagaimana karakteristik penyakit ISPA pada anak balita yang dirujuk di Puskesmas Cukir di Kabupaten Jombang Jawa Timur yang terdiagnosa dan tidak terdiagnosa penyakit ISPA.

1.3 Tujuan Penelitian

Dalam penelitian ini tujuan yang ingin dicapai dari peneliti secara umum yaitu mendapatkan klasifikasi status penyakit ISPA pada anak balita di Puskesmas Cukir Kabupaten Jombang berdasarkan faktor-faktor yang mempengaruhi dengan pendekatan *Classification Trees*. Sedangkan tujuan khusus dari penelitian ini adalah sebagai berikut.

1. Mendapatkan pohon klasifikasi maksimal dari data status terdiagnosa dan tidak terdiagnosa penyakit ISPA pada anak balita berdasarkan hasil analisis *Classification Trees*. Setelah itu, memperoleh hasil pohon klasifikasi optimal dari data status terdiagnosa penyakit ISPA dan tidak terdiagnosa pada anak balita berdasarkan hasil analisis *Classification Trees* dengan fungsi keheterogenan indeks gini.
2. Mengetahui faktor atau variabel yang paling dominan mempengaruhi penentuan hasil klasifikasi penyakit ISPA di Puskesmas Cukir Kabupaten Jombang.

1.4 Manfaat Penelitian

Adapun manfaat yang dapat dipaparkan dalam penelitian pada kasus penyakit ISPA ini adalah sebagai berikut.

1. Bagi bidang kesehatan, dapat memberikan informasi mengenai karakteristik anak balita yang terkena penyakit ISPA atau yang tidak terkena penyakit ISPA dengan menggunakan metode *Classification Trees*. Kemudian ketepatan klasifikasi yang dihasilkan dapat dijadikan informasi tambahan bagi dinas kesehatan Kabupaten Jombang dan Puskesmas Jombang dalam menentukan kebijakan pelayanan kesehatan dan tindakan medis bagi anak balita yang berisiko terkena penyakit ISPA yang berguna untuk menurunkan kasus penyakit ISPA.
2. Bagi bidang pendidikan, dengan penelitian ini diharapkan dapat berguna untuk menambah informasi dan wawasan pengetahuan pada penelitian selanjutnya khususnya mengenai klasifikasi dalam bidang kesehatan.

1.5 Batasan Masalah

Pada penelitian ini yang dijadikan batasan masalah adalah pasien yang diteliti merupakan anak balita yang terdiagnosa terkena penyakit ISPA dan tidak terdiagnosa penyakit ISPA di Puskesmas Cukir daerah Jombang.

(Halaman ini sengaja dikosongkan)

BAB II

TINJUAN PUSTAKA

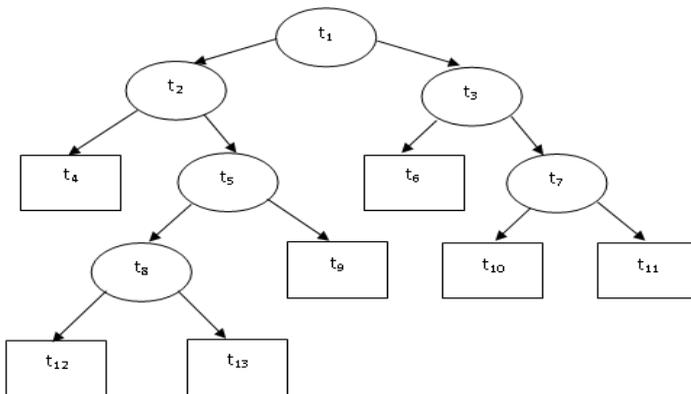
Pada Bab II diuraikan mengenai konsep landasan teori yang digunakan untuk mencapai tujuan penelitian, yaitu dengan menggunakan pendekatan metode CART pada kasus ISPA (Infeksi Saluran Pernapasan Akut). Adapun landasan teori yang digunakan sebagai berikut.

2.1 *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 yang mengandung seluruh data dengan notasi t_1 . Pada Gambar 2.1. Simpul utama (*root node*) dinotasikan sebagai t_1 , sedangkan simpul t_2 , t_3 , t_5 , t_7 dan t_8 disebut simpul dalam (*internal nodes*). Simpul akhir yang juga disebut sebagai simpul terminal (*terminal nodes*) adalah t_4 , t_6 , t_9 , t_{10} , t_{11} , t_{12} , dan t_{13} , dimana tidak terjadi lagi pemilahan. Kedalaman pohon (*depth*) dihitung dimulai dari simpul utama atau t_1 yang berada pada kedalaman 1, sedangkan t_2 dan t_3 berada pada kedalaman 2. Begitu seterusnya sampai pada simpul terminal t_{12} dan t_{13} 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*).

Langkah-langkah penerapan algoritma CART adalah sebagai berikut :

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.

a. Pemilihan Pemilah (*Classifier*)

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 homogeny dibandingkan dengan simpul induknya. Tingkat heterogenitas 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 pemilihan pemilah 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.

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

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

$p(i|t)$ adalah proporsi i pada simpul t .

$p(j|t)$ adalah proporsi 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* merupakan suatu evaluasi pemilahan oleh pemilah s pada simpul t yang didefinisikan sebagai penurunan keheterogenan dan didefinisikan 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 $\Delta_i(s, t)$ lebih tinggi merupakan pemilah yang terbaik karena hal ini memungkinkan untuk mereduksi keheterogenan 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.

b. Penentuan Simpul Terminal

Suatu simpul t akan menjadi simpul terminal atau tidak, akan dipilah kembali bila pada simpul t tidak terdapat penurunan keheterogenan secara berarti atau adanya batasan minimum n seperti hanya terdapat satu pengamatan pada tiap simpul anak. Jumlah kasus minimum dalam suatu terminal akhir umumnya adalah 5, dan apabila hal itu terpenuhi maka pengembangan pohon dihentikan.

c. Penandaan Label Kelas

Penandaan label kelas pada simpul terminal dilakukan berdasarkan aturan jumlah terbanyak. Label kelas simpul terminal t adalah j_0 yang memberi nilai dugaan kesalahan pengklasifikasian simpul t terbesar. Proses pembentukan pohon klasifikasi berhenti saat terdapat hanya satu pengamatan dalam tiap simpul anak atau adanya batasan minimum n , semua pengamatan dalam tiap simpul anak identik, dan adanya batasan jumlah level/kedalaman pohon maksimal.

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

Dengan

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

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

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

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

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

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

$R(T)$: Penduga pengganti (*Resubstitution Estimate*) ukuran kesalahan klasifikasi pada pohon T

a : Parameter *cost complexity* bagi penambahan satu simpul terminal pada pohon T

$|\tilde{T}|$: ukuran banyaknya simpul terminal 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 α . Nilai parameter kompleksitas (α) 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_1 .

Karena semakin besar pohon, semakin kecil nilai penduga penggantinya.

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

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.

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

dimana N_2 adalah jumlah pengamatan dalam data learning (L_2) dan $X(\cdot)$ bernilai 0 jika pertanyaan dalam tanda kurung salah dan bernilai 1 jika pertanyaan dalam tanda kurung benar. $R^{ts}(T_t)$ adalah total proporsi dari kesalahan *test sample estimate*. 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 $R^{ts}(T_t) = \min R^{ts}(T_t)$.

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).

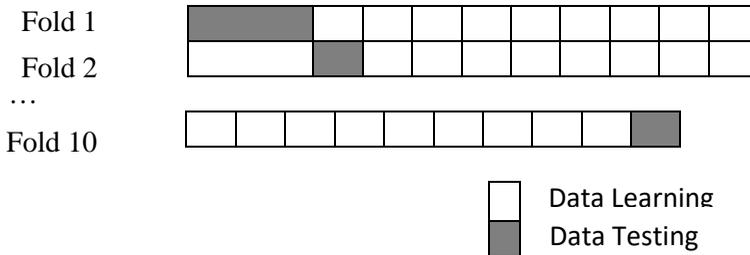
Dalam *v-fold cross validation estimate*, pohon akan terbentuk menggunakan data *learning* ($(v-1)/v$ bagian), yaitu $T_{maks}^{(v)}$ dengan $v = 1, \dots, v$ dengan kriteria pemilahan sampai simpul terminal memiliki jumlah pengamatan paling minimum. Apabila $d^{(v)}(x)$ adalah hasil pengklasifikasian dari pohon bentukan 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 dilakukan prosedur yang sama menggunakan semua pengamatan dalam L untuk membentuk deret pohon T_i . 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*.



Gambar 2.2 Ilustrasi Prosedur *10-fold Cross Validation*

2.2 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 2.1 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*.

$$APER = \frac{\text{jumlah prediksi salah}}{\text{Total Prediksi}} = \frac{n_{21} + n_{12}}{N} \quad (2.9)$$

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

2.3 Tinjauan Non Statistika

2.3.1 Infeksi Saluran Pernafasan Akut (ISPA)

Infeksi Saluran Pernafasan Akut (ISPA) dibedakan menjadi dua, ISPA atas dan bawah menurut Nelson, Infeksi saluran pernafasan atas adalah infeksi yang disebabkan oleh virus dan bakteri termasuk *common cold*, *faringitis akut*, *uvulitis akut*. Sedangkan, infeksi saluran pernafasan akut bawah merupakan infeksi yang telah didahului oleh infeksi bakteri sekunder, yang termasuk dalam penggolongan ini adalah *bronkhitis akut*, *bronkhitis kronis*, *bronkiolitis* dan *pneumonia aspirasi*.

Penyakit infeksi akut menyerang salah satu bagian atau lebih dari saluran nafas mulai hidung (saluran atas) hingga alveoli (saluran bawah) termasuk jaringan aksesoris seperti sinus, rongga telinga tengah dan pleura. Istilah ISPA meliputi tiga unsur yakni antara lain :

1. Infeksi

Infeksi merupakan masuknya kuman atau mikroorganisme ke dalam tubuh manusia dan berkembang biak sehingga menimbulkan gejala penyakit.

2. Saluran pernafasan

Saluran pernafasan merupakan organ mulai dari hidung hingga alveoli beserta organ aksesorinya seperti sinus, rongga telinga dan pleura.

3. Infeksi Akut

Infeksi yang berlangsung sampai dengan 14 hari. Batas 14 hari ditentukan untuk menunjukkan proses akut meskipun untuk beberapa penyakit yang dapat digolongkan dalam ISPA proses ini dapat berlangsung lebih dari 14 hari.

Penyakit ISPA pada balita dapat menimbulkan bermacam-macam tanda dan gejala seperti batuk, kesulitan bernafas, sakit tenggorokan, pilek, sakit telinga dan demam. Berikut gejala ISPA dibagi menjadi 3 antara lain sebagai berikut :

1. Gejala ISPA ringan

Seorang balita dinyatakan menderita ISPA ringan jika ditemukan satu atau lebih gejala-gejala sebagai berikut.

- a. batuk
- b. Serak, yaitu anak bersuara parau pada waktu mengeluarkan suara (pada waktu berbicara atau menangis)
- c. Pilek, yaitu mengeluarkan lendir atau ingus dari hidung
- d. Panas atau demam, suhu badan lebih dari 37 derajat celcius.

2. Gejala ISPA sedang

Seorang balita dinyatakan menderita ISPA sedang jika dijumpai gejala ISPA ringan disertai satu atau lebih gejala-gejala sebagai berikut.

- a. Pernafasan cepat sesuai umur yaitu untuk kelompok umur kurang dari 2 bulan frekuensi nafas 60 kali per menit atau lebih untuk umur 2 – 11 bulan dan 40 kali per menit atau lebih pada umur 12 bulan - 5 tahun.
- b. Suhu tubuh lebih dari 39 derajat celcius
- c. Tenggorokan berwarna merah

- d. Timbul bercak-bercak merah pada kulit menyerupai bercak campak
 - e. Telinga sakit atau mengeluarkan nanah dari lubang telinga
 - f. Pernafasan berbunyi seperti mengorok (mendengkur)
3. Gejala ISPA Berat

Seorang balita dinyatakan menderita ISPA berat jika dijumpai gejala-gejala ISPA ringan atau ISPA sedang disertai satu atau lebih gejala-gejala sebagai berikut.

- a. Bibir atau kulit membiru
- b. Anak tidak sadar atau kesadaran menurun
- c. Pernafasan berbunyi seperti mengorok dan anak tampak gelisah
- d. Nadi cepat lebih dari 160 kali per menit atau tidak teraba
- e. Tenggokan berwarna merah

2.3.2 Proses Terjadinya Infeksi Saluran Pernafasan

Saluran pernafasan dari hidung sampai bronkhus dilapisi oleh membran mukosa bersilia, udara yang masuk melalui rongga hidung disaring, dihangatkan dan dilembutkan. Partikel debu yang kasar dapat disaring oleh rambut yang terdapat dalam hidung, sedangkan partikel debu yang halus akan terjatuh dalam membran mukosa. Gerakan silia mendorong membran mukosa ke posterior kerongga hidung dan ke arah superior menuju faring.

Secara umum efek pencemaran udara terhadap pernafasan dapat menyebabkan pergerakan silia hidung menjadi lambat dan kaku bahkan dapat berhenti sehingga tidak dapat membersihkan saluran pernafasan akibat iritasi oleh bahan pencemar. Produksi lendir akan meningkat sehingga menyebabkan penyempitan saluran pernafasan dan makrofage di saluran pernafasan. Akibat dari dua hal tersebut akan menyebabkan kesulitan bernafas sehingga benda asing tertarik dan bakteri tidak dapat dikeluarkan dari saluran pernafasan, hal ini akan memudahkan terjadinya infeksi saluran pernafasan.

2.3.3 Penyebab Penyakit Infeksi Saluran Pernafasan Akut

ISPA dapat disebabkan oleh banyak hal, antara lain :

1. Menurut Nelson, Virus penyebab ISPA meliputi virus para influenza, adenovirus, rhinovirus, koronavirus, koksakavirus A dan B, Streptokokus dan lain-lain.
2. Perilaku individu, seperti sanitasi fisik rumah, kurangnya ketersediaan air bersih (Depkes RI, 2005:30). Untuk pencegahan ISPA dapat dilakukan dengan berbagai cara yaitu
 - a. Imunisasi
 - b. Penyehatan Lingkungan Pemukiman (PLP) polusi di dalam maupun di luar rumah
 - c. Mengatasi demam
 - d. Perbaikan makanan pendamping ASI
 - e. Penggunaan air bersih untuk kebersihan dan untuk minum

2.3.4 Penularan Infeksi Saluran Pernafasan Akut

Penyebaran melalui kontak langsung atau tidak langsung dari benda yang telah dicemari virus dan bakteri penyebab ISPA dan dapat juga ditularkan melalui udara tercemar pada penderita ISPA yang kebetulan mengandung bibit penyakit melalui sekresi berupa saliva atau sputum.

2.3.5 Rumah Sehat dan Faktor Lingkungan

a. Pengertian Rumah Sehat

Secara umum rumah dapat dikatakan sehat apabila memenuhi kriteria yaitu

1. Memenuhi kebutuhan fisiologis meliputi pencahayaan, ruang gerak yang cukup dan terhindar dari kebisingan yang mengganggu.
2. Memenuhi kebutuhan psikologis meliputi komunikasi yang harmonis antar keluarga dan penghuni rumah.
3. Memenuhi persyaratan pencegahan penularan penyakit antar penghuni rumah meliputi penyediaan air bersih, pengelolaan tinja, limbah rumah tangga, bebas vektor penyakit dan tikus, kepadatan hunian tidak berlebihan dan cukup sinar matahari pagi

Menurut Depkes RI, Puskesmas dan Rumah Sakit adalah proporsi rumah yang memenuhi kriteria sehat minimum terdapat komponen rumah dan sarana sanitasi di satu wilayah kerja pada kurun waktu tertentu. Minimum yang memenuhi kriteria sehat pada masing-masing parameter adalah sebagai berikut.

1. Minimum dari kelompok komponen rumah adalah langit-langit, dinding, lantai, jendela kamar tidur, jendela ruang keluarga, ventilasi, sarana pembuangan asap dapur dan pencahayaan.
2. Minimum kelompok fasilitas pendukung rumah sehat adalah sarana air bersih, jamban (sarana pembuangan kotoran), sarana pembuangan air limbah (SPAL) dan sarana pembuangan sampah.

Sanitasi rumah sangat erat kaitannya dengan angka kesakitan penyakit menular, terutama ISPA. Lingkungan perumahan sangat berpengaruh pada terjadinya dan tersebarnya ISPA.

Rumah yang tidak sehat dapat menjadi penyakit bagi seluruh lingkungan, jika kondisi tidak sehat bukan hanya pada satu rumah tetapi pada kumpulan rumah (lingkungan pemukiman). Timbulnya permasalahan kesehatan dilingkungan pemukiman pada dasarnya disebabkan karena tingkat kemampuan ekonomi yang rendah, karena rumah dibangun berdasarkan kemampuan penghuninya (Notoatmodjo, 2007 :168).

b. Faktor Lingkungan

Faktor lingkungan memegang peranan yang penting dalam menentukan terjadinya proses interaksi antara manusia dengan lingkungan dalam proses terjadinya penyakit. Secara garis besarnya lingkungan terdiri dari lingkungan fisik, biologis dan sosial.

Keadaan fisik sekitar manusia berpengaruh terhadap lingkungan-lingkungan biologis dan lingkungan sosial manusia. Lingkungan fisik (termasuk unsur kimia) meliputi udara, kelembaban, air dan pencemaran udara. Berkaitan dengan ISPA adalah termasuk air borne disease karena salah satu penularannya melalui udara yang tercemar dan masuk kedalam tubuh melalui saluran pernafasan, maka udara secara epidemiologi mempunyai peranan penting yang besar pada transmisi penyakit infeksi saluran pernafasan.

Kualitas udara dalam ruangan dipengaruhi oleh asap dalam ruangan yang bersumber dari perokok, penggunaan bahan bakar kayu atau arang atau asap. Disamping itu ditentukan oleh ventilasi, kepadatan penghuni, suhu ruangan, saluran pembuangan air limbah, tempat pembuangan sampah, ketersediaan air bersih, dan debu (polutan).

1. Ventilasi

Ventilasi rumah mempunyai banyak fungsi. Fungsi yang pertama adalah menjaga agar aliran udara dalam rumah tetap segar sehingga keseimbangan O₂ tetap terjaga, karena kurangnya ventilasi menyebabkan kurangnya O₂ yang berarti kadar CO₂ menjadi racun. Fungsi kedua adalah untuk membebaskan udara ruangan dari bakteribakteri, terutama bakteri patogen dan menjaga agar rumah selalu tetap dalam kelembaban yang optimum. (Notoatmodjo, 2007)

Menurut Notoatmodjo (2007: 170), ventilasi adalah proses udara segar ke dalam dan mengeluarkan udara kotor dari suatu ruangan tertutup secara alamiah maupun buatan. Berdasarkan kejadiannya ventilasi dibagi menjadidua yaitu :

a. Ventilasi Alamiah

Ventilasi alamiah berguna untuk mengalirkan udara di dalam ruangan yang terjadi secara alamiah melalui jendela, pintu dan lubang angin. Selain itu ventilasi alamiah juga menggerakkan udara sebagai hasil poros dinding ruangan, atap dan lantai.

b. Ventilasi Buatan

Ventilasi buatan dapat dilakukan dengan menggunakan alat mekanis maupun elektrik. Alat-alat tersebut di antaranya adalah kipas angin, exhauster dan AC.

2. Kepadatan Hunian

Setiap rumah harus mempunyai bagian ruangan yang sesuai fungsinya. Penentuan bentuk, ukuran dan jumlah ruangan perlu memperhatikan standar minimal jumlah ruangan. Sebuah rumah tinggal harus mempunyai ruangan yaitu kamar tidur, ruang tamu, ruang makan, dapur, kamar mandi dan kakus.

Kepadatan di dalam kamar terutama kamar balita yang tidak sesuai dengan standar akan meningkatkan suhu ruangan yang disebabkan

oleh pengeluaran panas badan yang akan meningkatkan kelembaban akibat uap air dari pernapasan tersebut. Dengan demikian, semakin banyak jumlah penghuni ruangan tidur maka semakin cepat udara ruangan mengalami pencemaran gas atau bakteri. Dengan banyaknya penghuni, maka kadar oksigen dalam ruangan menurun dan diikuti oleh peningkatan CO₂ dan dampak peningkatan CO₂ dalam ruangan adalah penurunan kualitas udara dalam ruangan.

3. Suhu Ruangan

Salah satu syarat fisiologis rumah sehat adalah memiliki suhu optimum 18-30°C. Hal ini berarti, jika suhu ruangan rumah dibawah 18°C atau di atas 30°C keadaan rumah tersebut tidak memenuhi syarat. Suhu ruangan yang tidak memenuhi syarat kesehatan menjadi faktor resiko terjadinya ISPA pada balita sebesar 4 kali. Suhu dalam ruangan berperan untuk menjaga rumah dalam kelembaban optimal untuk membebaskan bakteri dan virus (Erna, 2005: 77).

4. Saluran Pembuangan Air Limbah

Limbah merupakan pembuangan zat yang tidak terpakai yang berbentuk cair, gas, dan padat. Salah satunya adalah limbah rumah tangga yang berasal dari dapur, kamar mandi, cucian, limbah bekas industri rumah tangga dan kotoran manusia. Dalam air limbah terdapat bahan kimia yang sukar untuk dihilangkan dan berbahaya. Bahan kimia tersebut dapat memberi kehidupan bagi kuman-kuman penyebab penyakit disentri, tipus, kolera dan penyakit lainnya. Air limbah tersebut harus diolah agar tidak membahayakan kesehatan lingkungan. Air limbah harus dikelola untuk mengurangi pencemaran.

Pengelolaan air limbah yang dapat dilakukan yaitu pengelolaan limbah air bekas mandi dan cuci dialirkan ke bak kontrol dan langsung ke sumur resapan. Bak control perlu ditutup dan diberi pegangan agar memudahkan pengambilan tutup bak. Air akan tersaring pada bak resapan air yang keluar dari bak resapan sudah bebas dari pencemaran. Tempat mandi dan cuci dibuat dari batu bata, campuran semen dan pasir. Kemudian dibuat sumur resapan yang terbuat dari susunan batu bata kosong yang diberi kerikil dan lapisan ijuk. Sumur resapan diberi kerikil dan pasir. Jarak antara sumur air bersih ke sumur resapan minimum 10 m supaya tidak mencemari (Yulestra Putra, 2004)

5. Tempat Pembuangan Sampah

Pengaruh sampah terhadap kesehatan dapat disebabkan karena kontak langsung dengan sampah maupun tidak langsung akibat pembusukan, pembakaran dan pembuangan. Efek tidak langsung lainnya berupa penyakit bawaan vektor yang berkembang biak di dalam sampah.

6. Sumber Air Bersih

Air merupakan sumber kebutuhan sehari-hari bagi kehidupan manusia. Air yang bisa digunakan untuk keperluan sehari-hari harus diperhatikan kualitas dan kuantitasnya. Kualitas air yang baik jika air memenuhi syarat kesehatan seperti syarat fisik, kimia, bakteriologi dan radioaktif.

Berdasarkan Peraturan Menteri Kesehatan RI nomor 416/Menkes/Per/IX/1990, yang dimaksud air bersih adalah air yang digunakan untuk keperluan sehari-hari yang kualitasnya memenuhi syarat kesehatan dan dapat diminum apabila telah masak. Air bersih yang baik harus memenuhi syarat kualitas air bersih, yaitu :

- a. Syarat fisik yaitu tidak berwarna, tidak mempunyai rasa dan tidak berbau.
- b. Syarat kimia yaitu tidak mengandung zat kimia atau mineral yang berbahaya bagi kesehatan manusia.
- c. Syarat bakteriologis, yaitu tidak mengandung bakteri *E.coli* yang melampaui batas yang ditentukan.

BAB III METODOLOGI PENELITIAN

Pada bagian ini dijelaskan mengenai metode dan tahapan-tahapan dalam melakukan analisis untuk menyelesaikan permasalahan dalam penelitian ini yang meliputi sumber data, variabel penelitian, dan langkah penelitian. Metode analisis yang digunakan adalah klasifikasi dengan pendekatan *Classification Trees*.

3.1 Sumber Data

Data yang digunakan dalam penelitian tugas akhir ini merupakan data sekunder berupa data penyakit ISPA pada anak balita yang diperoleh dari Puskesmas Cukir yang terletak di Kabupaten Jombang tahun 2016 dan 2017. Unit penelitian dalam penelitian ini adalah anak balita yang berinfeksi penyakit ISPA dan tidak, yang jumlahnya mencapai 173 anak balita.

3.2 Kerangka Konsep

Berdasarkan teori, ISPA dan faktor-faktor yang dijelaskan pada tinjauan pustaka bahwa kejadian ISPA pada anak balita dipengaruhi oleh beberapa faktor yakni pola pemberian ASI, berat badan anak balita, lingkaran kepala, suhu tubuh, denyut nadi, respiratory rate (tingkat pernafasan), dan Status Ekonomi Keluarga. Berikut merupakan kerangka konsep yang digunakan dalam penelitian ini.



Gambar 3.1 Kerangka Konsep Penyakit ISPA

3.3 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini antara lain ditunjukkan pada Tabel 3.1

Tabel 3.1 Variabel Penelitian

Variabel	Deskripsi	Kategori dan Koding	Skala
Y	Status ISPA	0=Terdiagnosa ISPA 1=Tidak Terdiagnosa ISPA	Nominal
X ₁	Riwayat Pemberian ASI	0=NonASI 1=ASI Eksklusif (0-6 Bulan)	Nominal
X ₂	Berat Badan (kg)	-	Rasio
X ₃	Lingkar Kepala (cm)	-	Rasio
X ₄	Suhu Tubuh (°C)	-	Interval
X ₅	Denyut Nadi / permenit	-	Rasio
X ₆	Respiratory Rate (Tingkat Pernafasan)	0= tidak normal 1= normal	Nominal
X ₇	Status Ekonomi Keluarga	0= PBI 1= Non PBI	Nominal

3.4 Definisi Operasional

Berikut ini adalah definisi operasional dari masing-masing variabel yang digunakan dalam penelitian ini.

1. ISPA

Seringnya balita menderita ISPA dengan gejala klinis seperti : batuk, pilek, panas/demam dan telinga sakit atau mengeluarkan nanah dari lubang telinga dalam 6 bulan terakhir. Informasi dari gejala yang dialami anak didapatkan berdasarkan laporan dari orang tua ketika pengambilan data. Dikatakan anak mengalami ISPA, jika anak mengalami batuk, atau batuk dan pilek, atau pilek dan panas/demam

atau panas/demam dan batuk, atau telinga sakit atau mengeluarkan nanah dari lubang telinga. Dikatakan tidak mengalami ISPA, jika anak tidak mengalami gejala batuk, atau batuk dan pilek, atau pilek dan panas/demam, atau panas/demam dan batuk.

2. Pola Pemberian ASI Eksklusif

Pemberian asi berdasarkan kebiasaan ibu dalam memenuhi nutrisi bayi. Informasi didapatkan berdasarkan dari laporan ibu. Dikatakan menyusui eksklusif apabila bayi masih disusui, sejak lahir tidak pernah mendapatkan makanan dan minuman selain ASI termasuk air putih (kecuali obat-obatan dan vitamin atau mineral tetes). Dikatakan non ASI apabila bayi tidak diberikan ASI. Pemberian ASI eksklusif juga dapat mengurangi tingkat kekurangan gizi dan penyakit yang akan menimpa anak-anak (Dinkes Jatim, 2013).

3. Berat Badan Anak Balita

Berat badan ideal pada anak balita normal berusia 0-3 bulan sebesar 2.7 kg -5.7 kg, Berat badan ideal pada anak balita normal berusia 4-6 bulan adalah 5 kg-7.4 kg, berat badan anak balita normal berusia 7-9 bulan adalah 8 kg-8.9 kg, berat badan anak balita normal yang berusia 10-12 bulan adalah 9.3 kg-9.9 kg, pada anak balita normal berusia 1-2 tahun adalah 8.50 kg – 11.90 kg, sedangkan berat badan anak balita normal berusia 2-3 tahun adalah 9.50 kg – 14 kg. Pada anak balita normal yang berusia 3-4 tahun adalah 11.20-16 kg. Dan berat badan pada anak balita normal yang berusia 4-5 tahun adalah 12-18kg.

4. Lingkar Kepala

Lingkar kepala anak balita normal pada usia 1-3 bulan sebesar 33-43 cm, lingkar kepala pada anak balita normal yang berusia 4-6 bulan sebesar 38-46 cm, sedangkan lingkar kepala pada anak balita normal yang berusia 7-9 bulan adalah 40.5-48 cm, lingkar kepala pada anak balita normal yang berusia 10-12 bulan sebesar 42.5-49.5, dan lingkar kepala pada anak balita normal yang berusia 1-2 tahun adalah 44-51 cm, sedangkan lingkar kepala pada anak balita normal yang berusia 2-3 tahun sebesar 45.5-53 cm. lingkar kepala pada anak balita normal yang berusia 3-4 tahun adalah 46.5- 53.8 cm. Dan lingkar kepala pada anak balita normal yang berusia 4-5 tahun 47.5-54 cm.

5. Suhu Tubuh Bayi

Jika suhu tubuh bayi yang lebih dari 37.5°C (demam) disertai keluhan batuk, pilek, dan sesak nafas akan mengarah ke ISPA, sedangkan suhu tubuh bayi dalam keadaan normal yaitu 36 °C-37 °C. (Kementrian Kesehatan RI,2010).

6. Denyut Nadi

Denyut nadi pada anak balita normal, jika berdenyut kurang dari 100 kali / menit. Namun jika denyut nadi tidak normal yaitu lebih dari 100 kali / menit, maka kemungkinan akan mengalami penyakit infeksi.

7. Respiratory Rate

Tingkat pernafasan pada anak balita, jika dalam keadaan normal ≤ 50 kali / menit sedangkan jika tingkat pernafasan pada anak balita > 50 kali / menit maka terdeteksi anak balita mengalami resiko berinfeksi.

8. Status Ekonomi Keluarga

Kondisi ekonomi keluarga yang dihitung berdasarkan fasilitas jaminan kesehatan yang diterima. Kategori miskin yaitu yang mempunyai jaminan kesehatan masyarakat PBI (Penerima Bantuan Iuran), sedangkan kategori mampu yaitu yang tidak mempunyai jaminan kesehatan Non PBI (Mandiri) di Puskesmas Cukir.

3.5 Langkah Analisis *Classification Trees*

Langkah analisis yang dilakukan dalam penelitian tugas akhir ini mengenai klasifikasi risiko penyakit ISPA pada anak balita di Puskesmas Jombang menggunakan *Classification Trees* adalah sebagai berikut.

1. Mendeskripsikan karakteristik anak balita di Puskesmas Jombang berdasarkan dua tipe variabel respon. Apabila variabel prediktor berskala kategorik, maka dideskripsikan menggunakan *bar-chart*. Sedangkan untuk variabel prediktor yang memiliki skala kontinu dideskripsikan menggunakan rata-rata, standar deviasi, nilai maksimum dan nilai minimum.
2. Membagi data menjadi dua bagian, yaitu data learning dan data testing. Data dibagi sesuai dengan aturan *10-fold cross validation*

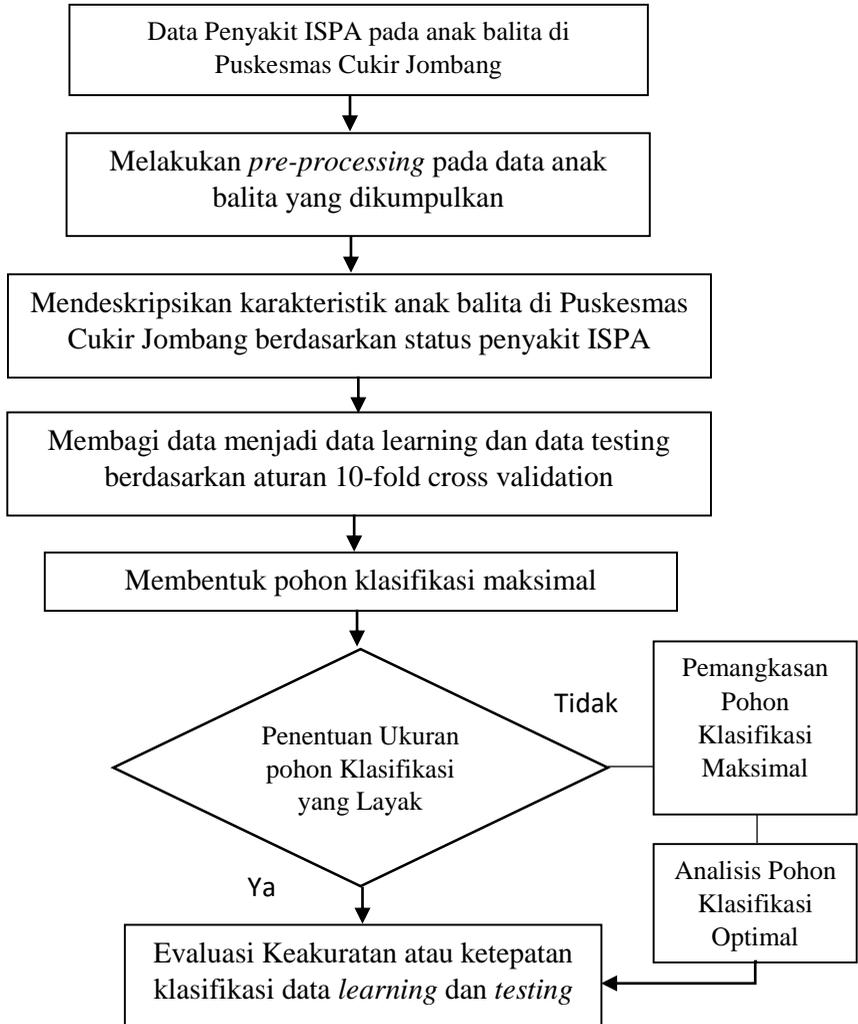
estimate, karena dalam menentukan pohon optimal menggunakan metode *10-fold cross validation estimate*.

3. Melakukan analisis klasifikasi dengan 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. Menentukan pohon klasifikasi yang layak dengan melihat besarnya nilai kompleksitas pohon klasifikasi yang terbentuk dan nilai *resubstitution relative cost*. Jika nilai kompleksitas 0.000 dan nilai *resubstitution relative cost* yang kecil (menunjukkan struktur data dari pohon klasifikasi maksimal kompleks) maka perlu dilakukan pemangkasan pohon klasifikasi maksimal (pruning) sehingga diperoleh suatu pohon klasifikasi optimal.
 - c. Menentukan pohon klasifikasi optimal dengan *10-Fold Cross Validation Estimate*.

- d. Mendapatkan karakteristik kelas simpul terminal yang dihasilkan dari penelusuran pohon klasifikasi optimal.
- e. Menghitung ketepatan klasifikasi pohon hasil bentukan dengan menggunakan data learning dan validasi dengan menggunakan data testing pada pohon klasifikasi yang terbentuk. Ukuran ketepatan klasifikasi yang digunakan adalah *total accuracy rate* (1-APER), *sensitivity* dan *specificity*.

3.6 Diagram Alir Analisis *Classification Trees*

Tahapan-tahapan dalam penelitian tugas akhir ini akan digambarkan melalui diagram alir penelitian yang disajikan pada Gambar 3.2 berikut.



Gambar 3.2 Diagram Alir

(Halaman Sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

Dalam bab ini membahas tentang analisis data faktor-faktor yang mempengaruhi anak balita terdiagnosa Infeksi Saluran Penyakit Akut (ISPA) di Puskesmas Cukir untuk menjawab permasalahan serta tujuan dari penelitian yang telah dijelaskan pada bab sebelumnya.

Pada Subbab 4.1 diberikan penjelasan hasil statistika deskriptif data penelitian yaitu tentang karakteristik status penyakit ISPA berdasarkan faktor-faktor yang mempengaruhi di Puskesmas Cukir Jombang tahun 2016 dan 2017 untuk memberikan gambaran mengenai unit yang diteliti. Lalu pada Subbab 4.2 diberikan penjelasan analisis klasifikasi status penyakit ISPA dengan pendekatan pohon klasifikasi CART.

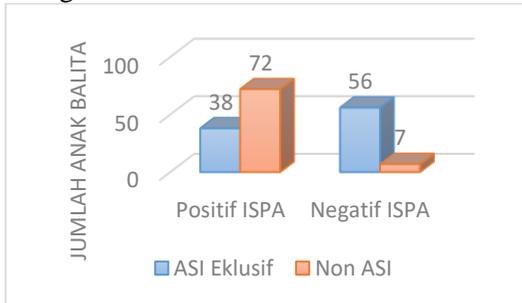
4.1 Karakteristik Status Penyakit ISPA Pada Anak Balita di Puskesmas Cukir Jombang

Anak balita yang terinfeksi penyakit ISPA di puskesmas cukir kabupaten jombang merupakan hal yang diukur dalam penelitian ini, dimana karakteristik-karakteristik dari pasien anak balita yang dikumpulkan sebagai data. Kemudian hal yang diperhatikan adalah status diagnosa pada anak balita tersebut, yakni positif atau negatif terinfeksi penyakit ISPA. Status infeksi penyakit ISPA merupakan variabel respon dalam penelitian ini. Indikator yang mempengaruhi penyakit ISPA terdapat beberapa faktor yaitu Pola pemberian ASI, berat badan, lingkar kepala, suhu tubuh, denyut nadi, respiratory rate (tingkat pernafasan), dan status ekonomi keluarga.

Jumlah sampel kasus penyakit ISPA pada anak balita di Puskesmas Cukir Jombang yang menjadi sampel pengamatan pada penelitian ini adalah 173 anak balita. Tampilan beberapa data pengamatan anak balita mengenai status penyakit ISPA disajikan dalam Lampiran A.

4.1.1 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Riwayat Pemberian ASI

Pada anak balita juga tercatat riwayat pemberian ASI pada usia 0-6 bulan, apakah ibu dapat memberikan ASI atau tidak. Untuk mengetahui perbandingan jumlah riwayat pola pemberian ASI anak balita dan status risiko ISPA di Puskesmas Cukir ditampilkan pada Gambar 4.1 sebagai berikut.



Gambar 4.1 Deskripsi Pola Pemberian ASI Berdasarkan Status Risiko ISPA

Gambar 4.1 memberikan informasi bahwa dari total 110 anak balita yang berstatus positif terkena risiko penyakit ISPA, sebesar 38 anak balita diantaranya aktif dalam pola pemberian ASI eksklusif dan 72 anak balita yang tidak diberi dalam pemberian ASI eksklusif. Sedangkan dari 63 anak balita yang berstatus negatif terkena penyakit ISPA, sebesar 56 anak balita yang aktif dalam pemberian ASI dan 7 anak balita yang tidak diberi ASI eksklusif. Sehingga dapat disimpulkan bahwa sebagian besar anak balita yang tidak berisiko penyakit ISPA diberikan ASI eksklusif.

4.1.2 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Berat Badan

Karakteristik penyakit ISPA pada anak balita di Puskesmas Cukir berdasarkan berat badan dapat dijelaskan dengan tampilan Tabel 4.1 sebagai berikut.

Tabel 4.1 Deskripsi Penyakit ISPA Berdasarkan Berat Badan

Status ISPA	Rata-rata (kg)	Standar Deviasi	Maksimum (kg)	Minimum (kg)
Terdiagnosa	9,075	1,973	14,000	3,000
Tidak terdiagnosa	8,721	2,062	12,000	2,700

Tabel 4.1 memberikan informasi mengenai penyakit ISPA berdasarkan indikator berat badan pada anak balita. Rata-rata anak balita yang terdiagnosa penyakit ISPA memiliki berat badan sebesar 9,075 kilogram. Sedangkan anak balita yang tidak terdiagnosa penyakit ISPA memiliki berat badan sebesar 8,721 kilogram. Sementara itu deviasi standar untuk masing-masing anak balita terhadap diagnosa penyakit ISPA sebesar 1,973 dan 2,062.

Anak balita yang terdiagnosa penyakit ISPA memiliki berat badan yang paling besar adalah 14 kilogram, sedangkan paling minimum memiliki berat badan sebesar 3 kilogram. Pada anak balita yang tidak terdiagnosa penyakit ISPA memiliki berat badan yang paling besar adalah 12 kilogram, sementara anak balita yang memiliki berat badan minimum sebesar 2,7 kilogram.

4.1.3 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Lingkar Kepala

Pada anak balita yang berusia 0-5 tahun juga dicatat ukuran lingkar kepalanya. Karakteristik anak balita di Puskesmas Cukir berdasarkan lingkar kepala dapat dijelaskan pada Tabel 4.2 sebagai berikut.

Tabel 4.2 Deskripsi Penyakit ISPA Berdasarkan Lingkar Kepala

Status ISPA	Rata-rata (cm)	Standar Deviasi	Maksimum (cm)	Minimum (cm)
Terdiagnosa	46,645	3,703	54,000	35,000
Tidak terdiagnosa	46,627	4,010	54,000	35,000

Tabel 4.2 menjelaskan mengenai penyakit ISPA berdasarkan indikator lingkar kepala pada anak balita. Rata-rata anak balita yang terdiagnosa penyakit ISPA memiliki lingkar kepala sebesar 46,645 cm. Sedangkan anak balita yang tidak terdiagnosa penyakit ISPA memiliki lingkar kepala sebesar 46,627 cm. Sementara itu deviasi standar untuk masing-masing anak balita terhadap diagnosa penyakit ISPA sebesar 3,703 dan 4,010.

Anak balita yang terdiagnosa penyakit ISPA memiliki lingkaran kepala yang paling besar adalah 54 cm, sedangkan paling minimum memiliki berat badan sebesar 35 cm. Pada anak balita yang tidak terdiagnosa penyakit ISPA memiliki lingkaran kepala yang paling besar adalah 54 cm, sementara anak balita yang memiliki lingkaran kepala minimum sebesar 35 cm.

4.1.4 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Suhu Tubuh

Indikator suhu tubuh anak balita berdasarkan masing-masing status diagnosa penyakit ISPA ditampilkan dalam Tabel 4.3 sebagai berikut.

Tabel 4.3 Deskripsi Penyakit ISPA Berdasarkan Suhu Tubuh

Status ISPA	Rata-rata (°C)	Standar Deviasi	Maksimum (°C)	Minimum (°C)
Terdiagnosa	38,315	1,222	40,700	36,000
Tidak terdiagnosa	36,495	0,649	39,600	36,000

Berdasarkan Tabel 4.3 dapat diketahui bahwa rata-rata suhu tubuh anak balita yang terdiagnosa penyakit ISPA adalah 38,315 °C dengan standar deviasi sebesar 1,222. Nilai standar deviasi yang cukup kecil menunjukkan bahwa suhu tubuh anak balita yang terdiagnosa penyakit ISPA tidak beragam. Sedangkan rata-rata suhu tubuh pada anak balita yang tidak terdiagnosa penyakit ISPA adalah 36,495 °C dengan standar deviasi 0,649.

Nilai maksimum dari suhu tubuh anak balita yang terdiagnosa penyakit ISPA adalah 40,700 °C, sedangkan paling minimum suhu tubuh anak balita sebesar 36,000 °C. Untuk suhu tubuh anak balita yang tidak terdiagnosa penyakit ISPA memiliki nilai suhu tubuh paling tinggi adalah 39,600 °C sedangkan paling rendah memiliki suhu tubuh sebesar 36,000 °C artinya terdapat anak balita yang mendapatkan sakit demam biasa tanpa terjangkit penyakit ISPA, karena penyakit ISPA masih banyak dipengaruhi berbagai faktor, tidak hanya satu faktor.

4.1.5 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Denyut Nadi

Anak balita harus melewati pemeriksaan denyut nadi yang merupakan salah satu faktor yang mempengaruhi dalam diagnosa penyakit ISPA. Berikut yang dapat ditampilkan dalam Tabel 4.4 mengenai karakteristik anak balita berdasarkan denyut nadi.

Tabel 4.4 Deskripsi Penyakit ISPA Berdasarkan denyut nadi

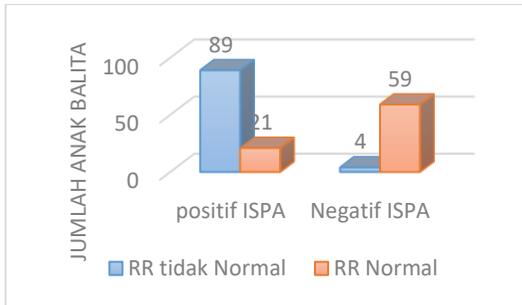
Status ISPA	Rata-rata (permenit)	Standar Deviasi	Maksimum (permenit)	Minimum (permenit)
Terdiagnosa	104,45	7,19	120,00	90,00
Tidak terdiagnosa	96,286	5,589	111,000	90,000

Berdasarkan Tabel 4.4 diperoleh informasi bahwa rata-rata denyut nadi yang terdiagnosa penyakit ISPA adalah 104,45 kali permenit dengan standar deviasi sebesar 7.19. Denyut nadi yang dimiliki cukup beragam karena nilai standar deviasi cukup besar. Sedangkan anak balita dengan tidak terdiagnosa penyakit ISPA memiliki rata-rata denyut nadi sebesar 96,286 kali per menit dengan standar deviasi sebesar 5,589.

Tabel 4.4 juga memberikan informasi mengenai denyut nadi yang paling maksimum yang dimiliki oleh anak balita yang terdiagnosa penyakit ISPA sebesar 120 kali permenit dan paling minimum sebesar 90 kali permenit. Sedangkan anak balita yang tidak terdiagnosa penyakit ISPA memiliki denyut nadi paling maksimum sebesar 111 kali permenit, untuk paling minimum sebesar 90 kali per menit artinya bayi memiliki denyut nadi normal.

4.1.6 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Respiratory Rate (RR)

Tingkat pernafasan (RR) merupakan suatu usaha pemeriksaan terhadap anak balita yang terdiagnosa penyakit. Indikator kecepatan bernafas dari anak balita dapat diketahui pula perbandingan indikator tersebut dengan status penyakit ISPA dalam Gambar 4.2 sebagai berikut.

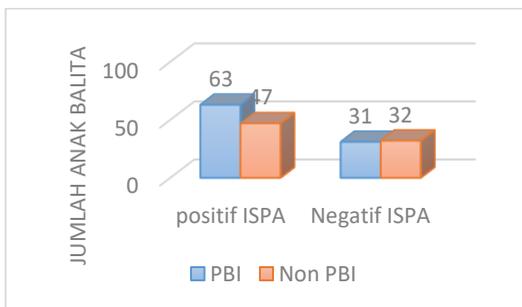


Gambar 4.2 Deskripsi Tingkat Pernafasan Berdasarkan Status Penyakit ISPA

Gambar 4.2 menunjukkan bahwa dari total 110 anak balita yang terdiagnosa penyakit ISPA berdasarkan indikator tingkat pernafasan sebesar 21 anak balita diantaranya memiliki tingkat pernafasan yang normal dan sebesar 89 anak balita memiliki tingkat pernafasan yang tidak normal. Begitu juga dengan total 63 anak balita yang terindikator negatif penyakit ISPA dengan 59 anak balita diantaranya memiliki nafas yang normal. Sehingga dapat disimpulkan bahwa sebagian besar anak balita yang tidak terdiagnosa penyakit ISPA memiliki tingkat pernafasan yang normal.

4.1.7 Karakteristik Penyakit ISPA Anak Balita Berdasarkan Status Perekonomian Keluarga

Indikator status perekonomian anak balita berdasarkan masing-masing status penyakit ISPA ditampilkan dalam Gambar 4.3 sebagai berikut.



Gambar 4.3 Deskripsi Status Perekonomian Keluarga Berdasarkan Status Risiko Penyakit ISPA

Gambar 4.3 menunjukkan bahwa dari total 110 anak balita yang positif terdiagnosa penyakit ISPA berdasarkan indikator status perekonomian keluarga terdapat 63 anak balita yang dikategorikan menerima bantuan iuran (PBI) / keluarga miskin (Gakin). Dan terdapat 47 anak balita yang dikategorikan tidak menerima bantuan iuran (PBI) / keluarga miskin (Gakin).

Sedangkan total dari 63 anak balita yang negatif terdiagnosa penyakit ISPA terdapat 31 anak balita yang dikategorikan menerima bantuan iuran (PBI) / keluarga miskin (Gakin). Dan terdapat 32 anak balita yang dikategorikan tidak menerima bantuan iuran (PBI)/ keluarga miskin (Gakin).

4.2 Analisis *Classification Trees* untuk Klasifikasi Diagnosa ISPA pada Anak Balita di Puskesmas Cukir Jombang

Berikut ini adalah hasil analisis klasifikasi resiko terdiagnosa penyakit ISPA pada anak balita di Puskesmas Cukir Jombang menggunakan metode *Classification Trees* dengan jumlah data sebanyak 173. Analisis menggunakan *Classification Trees* bertujuan untuk mengetahui ketepatan klasifikasi yang telah dihasilkan berdasarkan status resiko penyakit ISPA pada anak balita melalui pohon klasifikasi maksimal dan pohon klasifikasi optimal dengan proses pemilahan menggunakan fungsi keheterogenan indeks gini.

Berdasarkan data pengamatan yang digunakan dalam analisis tersebut terdapat 110 kasus yang dikategorikan sebagai anak balita yang terdiagnosa penyakit ISPA, sedangkan sisanya 63 kasus yang dikategorikan sebagai anak balita yang tidak terdiagnosa penyakit ISPA. 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 17 dan 18 data. Berikut ini adalah hasil dari pengolahan data Penyakit ISPA anak balita di Puskesmas Cukir Jombang.

4.2.1 Pembentukan Pohon Klasifikasi

Dalam Pembentukan pohon klasifikasi terdapat beberapa langkah yaitu pemilihan pemilah, penentuan simpul terminal, dan penandaan label kelas. 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 4.5 sebagai berikut.

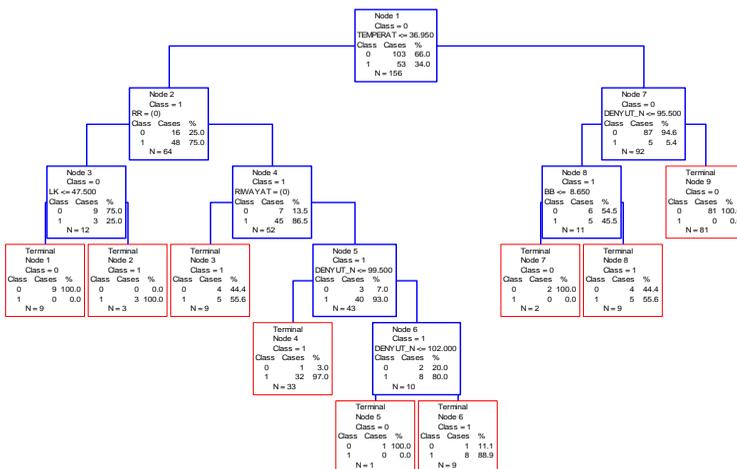
Tabel 4.5 Skor Variabel prediktor dalam pembentukan Pohon Klasifikasi Maksimal

Variabel	Nama Variabel	Score	
X4	Suhu Tubuh	100.00	
X5	Denyut Nadi	20.60	
X6	Tingkat Pernafasan	12.50	
X3	Lingkar Kepala	12.16	
X2	Berat Badan	3.01	
X1	Riwayat ASI	2.61	
X7	Status Ekonomi Keluarga	0.00	

Berdasarkan Tabel 4.5 dapat diperoleh informasi bahwa semua variabel prediktor menjadi pembangun dalam pembentukan pohon klasifikasi. Namun berdasarkan skor yang dihasilkan diketahui bahwa variabel terpenting yang memiliki kontribusi terbesar adalah variabel X_4 (Suhu Tubuh) dengan skor 100. Artinya variabel X_4 akan menjadi pemilah awal atau biasa disebut sebagai simpul induk (*parent nodes*). Suhu Tubuh mampu menurunkan tingkat heterogenan paling besar, sehingga simpul yang dihasilkan lebih homogen. Selain itu, terdapat beberapa variabel lain yang juga berpengaruh besar dalam pembentukan pohon klasifikasi yaitu variabel X_5 (Denyut Nadi)

mempunyai skor kontribusi 20.60, variabel X_6 (Tingkat Pernafasan) mempunyai skor kontribusi 12.50, variabel X_3 (lingkar kepala) memperoleh skor kontribusi sebesar 12.16, variabel X_2 (Berat Badan) memperoleh nilai skor kontribusi sebesar 3.01, variabel X_1 (Riwayat ASI) mendapatkan skor 2.61 dan variabel X_7 (ekonomi keluarga) memperoleh skor kontribusi sebesar 0.00.

Selanjutnya dilakukan proses pemilahan dengan fungsi keheterogenan indeks gini sampai 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 9 simpul dengan tingkat kedalaman 6. Topologi dari pohon klasifikasi maksimal ditampilkan dalam Gambar 4.4 sebagai berikut.

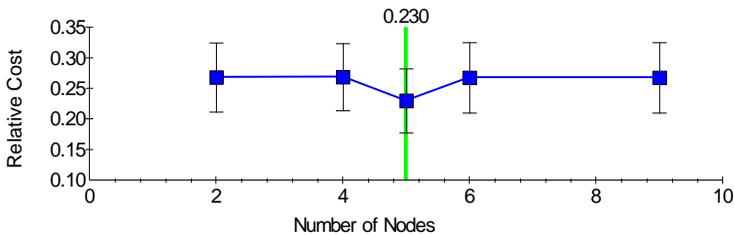


Gambar 4.4 Topologi Pohon Klasifikasi Maksimal

4.2.2 Pemangkasan Pohon Klasifikasi

Setelah memperoleh pohon klasifikasi maksimal langkah berikutnya adalah melihat apakah pohon klasifikasi maksimal 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 actual). 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 menjadi 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 *misclassification error rate*. Setiap hasil pemangkasan yang dilakukan memiliki nilai *relative cost* tertentu, kemudian dipilih hasil pemangkasan yang mempunyai nilai *relative cost* paling minimum. Dalam kasus ini, pohon klasifikasi sudah optimal dapat dilihat pada gambar Plot *relative cost* yang sudah berwarna hijau yang ditunjukkan pada Gambar 4.5.



Gambar 4.5 Plot *Relative Cost* dan Banyaknya Simpul Terminal

Jika dilihat pada plot *relative cost* pohon klasifikasi dengan simpul terminal sebanyak 9 simpul. Pohon klasifikasi optimal ditunjukkan dengan nilai *relative cost* yang dihasilkan sebesar 0.230, ditandai dengan garis berwarna hijau. Sedangkan untuk *cross-validation relative cost* menghasilkan nilai yang paling minimum yakni sebesar 0.230 ± 0.052 dengan nilai kompleksitas sebesar 0.010 sehingga tidak dilakukan pemangkasan (*pruning*). Berikut ini akan ditunjukkan urutan dalam pembentukan pohon klasifikasi sampai didapatkan pohon klasifikasi optimal yang ditampilkan pada Tabel 4.6.

Tabel 4.6 Urutan Pembentukan Pohon Klasifikasi (*Tree Sequence*)

Tree Number	Terminal Nodes	<i>Cross-Validated Relative Cost</i>	Resubstitution Relative Cost	<i>Complexity</i>
1	9	0.268+/- 0.057	0.097	0.000
2	6	0.268+/-0.057	0.107	0.002
3**	5	0.230+/-0.052	0.126	0.010
4	4	0.269+/-0.055	0.162	0.018
5	2	0.269+/-0.056	0.250	0.022
6	1	1.000+/- . 610352E-04	1.000	0.375

***pohon klasifikasi optimal*

Berdasarkan Tabel 4.6, jika dilihat pada (*cross-validation relative cost*) yang dihasilkan oleh hasil pohon klasifikasi makasimal, pohon pertama merupakan pembentukan pohon klasifikasi maksimal dengan jumlah simpul terminal sebanyak 9 simpul dan biaya kesalahan berkisar antara 0.268 dan 0.057 dengan nilai kompleksitas sebesar 0.000. Sedangkan pohon ketiga merupakan pohon klasifikasi yang menghasilkan biaya kesalahan paling minimum yakni berkisar antara 0.230 dan 0.052 dan nilai kompleksitas sebesar 0.010.

4.2.3 Penentuan Pohon Klasifikasi Optimal

Setelah memperoleh pembentukan pohon klasifikasi secara iteratif, langkah selanjutnya yang dilakukan adalah menentukan pohon klasifikasi optimal untuk mengetahui akurasi hasil klasifikasi risiko penyakit ISPA pada anak balita di Puskesmas Cukir Jombang. 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 4.7

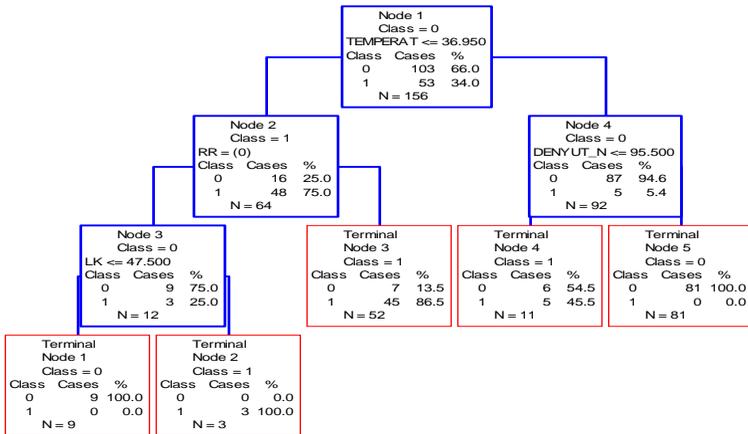
Tabel 4.7 Skor Variabel prediktor dalam pembentukan Pohon Klasifikasi Optimal

Variabel	Nama Variabel	Score	
X4	Suhu Tubuh	100.00	
X5	Denyut Nadi	17.29	

Tabel 4.7 Skor Variabel prediktor dalam pembentukan Pohon Klasifikasi Optimal (lanjutan)

Variabel	Nama Variabel	Score
X6	Tingkat Pernafasan	12.50
X3	Lingkar Kepala	12.16
X2	Berat Badan	0.00
X1	Riwayat ASI	0.00
X7	Status Ekonomi Keluarga	0.00

Berdasarkan Tabel 4.7 dapat diketahui bahwa variabel X₄ (Suhu Tubuh) merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal dengan skor kontribusi sebesar 100. Sedangkan variabel-variabel prediktor lainnya 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 5 simpul dan tingkat kedalaman sebesar 4. Berikut ini ditampilkan topologi pohon klasifikasi optimal berdasarkan faktor-faktor yang mempengaruhi Penyakit ISPA pada anak balita dalam Gambar 4.6.



Gambar 4.6 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

penyakit ISPA pada anak balita. Apabila kelas dilabelkan sebagai 0 maka anak balita terdiagnosa penyakit ISPA. Namun, jika kelas dilabelkan sebagai 1 maka anak balita tidak terdiagnosa penyakit ISPA. Berikut ini merupakan penyajian mengenai pelabelan kelas yang telah dilakukan terhadap simpul terminal.

Tabel 4.8 Pelabelan Kelas Simpul Terminal

Kelas	Simpul Terminal
	Ke-
Status terdiagnosa ISPA	1
	5
Status tidak terdiagnosa ISPA	2
	3
	4

Berdasarkan Tabel 4.8 diperoleh informasi bahwa simpul terminal yang masuk kekelas Status diagnosa ISPA adalah simpul terminal ke 1 dan simpul terminal ke 5. Karakteristik anak balita yang tidak terdiagnosa ISPA yaitu simpul terminal ke 2, 3, dan 4.

Proses pemilahan mengacu pada pohon klasifikasi optimal yang ditampilkan dalam Gambar 4.6. Pemilahan diawali dengan memilah sebanyak 156 anak balita berdasarkan suhu tubuh (X_4) dan nilai *threshold* 36,950. Kemudian data dibagi menjadi dua bagian yaitu simpul kiri (*node 2*) yang terdiri dari 64 anak balita dengan kondisi suhu tubuh kurang dari sama dengan 36,950 dan dilabeli sebagai anak balita yang tidak terdiagnosa penyakit ISPA. Sedangkan, simpul kanan (*node 4*) terdiri dari 92 anak balita dengan kondisi suhu tubuh lebih dari 36,950 dan diberi label kelas sebagai anak balita yang terdiagnosa penyakit ISPA. Selanjutnya pemilahan kembali dilakukan secara berulang karena tingkat keheterogenan masih terbilang tinggi. Pemilahan berhenti pada jumlah terminal simpul sebesar 5 simpul.

Secara menyeluruh, melalui Gambar 4.6 diperoleh informasi bahwa terdapat 2 simpul terminal diklasifikasikan sebagai anak balita yang terdiagnosa ISPA dan 3 simpul diklasifikasikan sebagai anak balita yang tidak terdiagnosa ISPA. Karakteristik simpul terminal yang mengklasifikasikan anak balita menjadi salah satu dari kedua status risiko diagnosa ISPA dijelaskan sebagai berikut.

- a. Simpul terminal satu terdiri dari 9 anak balita yang diprediksi terdiagnosa ISPA. Karakteristik anak balita pada simpul ini adalah memiliki lingkaran kepala kurang dari sama dengan 47,5 cm, tingkat pernafasan anak balita dalam keadaan tidak normal, dan suhu tubuh dalam kondisi kurang dari sama dengan 36,95°C.
- b. Simpul terminal dua terdiri dari 3 anak balita yang diprediksi tidak mengalami penyakit ISPA. Karakteristik anak balita pada simpul ini antara lain adalah memiliki lingkaran kepala lebih dari 47,5 cm, tingkat pernafasan anak balita dalam keadaan tidak normal, dan suhu tubuh dalam kondisi kurang dari sama dengan 36,95°C.
- c. Simpul terminal tiga terdiri dari 52 anak balita dengan diprediksi tidak terdiagnosa penyakit ISPA. Karakteristik anak balita yang terbentuk adalah memiliki tingkat pernafasan normal dan memiliki suhu tubuh kurang dari sama dengan 36,95°C.
- d. Simpul terminal empat terdiri dari 11 anak balita dengan diprediksi tidak mengalami penyakit ISPA yang memiliki beberapa karakteristik yang terbentuk antara lain memiliki denyut nadi kurang dari sama dengan 96 kali permenit dan memiliki keadaan suhu tubuh sebesar lebih dari 36,95°C.
- e. Simpul terminal lima terdiri dari 81 anak balita dengan diprediksi mengalami penyakit ISPA. Karakteristik anak balita yang termasuk pada simpul ini antara lain adalah memiliki denyut nadi lebih dari 96 kali permenit dan memiliki kondisi suhu tubuh sebesar lebih dari 36,95°C.

Sesuai dengan penjelasan sebelumnya mengenai karakteristik anak balita berdasarkan hasil analisis *Classification Trees*, selanjutnya ditampilkan perbedaan antara karakteristik anak balita dengan dua kategori status diagnosa penyakit ISPA, yaitu anak balita terdiagnosa penyakit ISPA dan anak balita tidak terdiagnosa penyakit ISPA. Karakteristik dari kedua kategori status diagnosa penyakit ISPA pada anak balita disajikan dalam Tabel 4.9 sebagai berikut.

Tabel 4.9 Karakteristik Anak Balita Berdasarkan Simpul Terminal

Terdiagnosa ISPA	Tidak terdiagnosa ISPA
<ul style="list-style-type: none"> • Anak balita dengan lingkaran kepala \leq 47,5 cm, tingkat pernafasan tidak normal, suhu tubuh \leq 36,95°C. • Anak balita dengan denyut nadi $>$ 96 kali permenit, kondisi suhu tubuh $>$ 36,95 °C. 	<ul style="list-style-type: none"> • Anak balita dengan lingkaran kepala $>$ 47,5 cm, tingkat pernafasan keadaan tidak normal, kondisi suhu tubuh \leq 36,95 °C. • Anak balita yang memiliki tingkat pernafasan normal, kondisi suhu tubuh \leq 36,95 °C • Anak balita dengan denyut nadi kurang \leq 96 kali permenit, kondisi suhu tubuh $>$ 36,95 °C

4.3 Hasil Ketepatan Klasifikasi *Classification Trees*

Hasil klasifikasi penyakit ISPA pada anak balita di puskesmas cukir kabupaten Jombang berdasarkan faktor-faktor yang mempengaruhi penyakit ISPA pada anak balita menggunakan metode *10-fold cross validation* diukur berdasarkan total ketepatan klasifikasi tiap kelas dan total kesalahan klasifikasi adalah sebagai berikut.

Tabel 4.10 Ketepatan Klasifikasi Risiko Penyakit ISPA pada Anak Balita

Aktual	Prediksi		Total	Total Akurasi	Sensitivity	Specificity
	ISPA	Non ISPA				
ISPA	89	14	103	0.878	0.864	0.906
Non ISPA	5	48	53			
Total	94	62	156			

Berdasarkan Tabel 4.10 dapat diketahui terdapat 103 anak balita yang termasuk dalam kategori ISPA (terdiagnosa penyakit ISPA) dengan 89 tepat diklasifikasikan sebagai anak balita dengan terdiagnosa penyakit ISPA dan 14 anak balita salah diklasifikasikan sebagai anak balita yang tidak terdiagnosa penyakit ISPA. Sehingga nilai sensitivity yang diperoleh sebesar 0.864. Sementara itu jumlah anak balita yang tidak termasuk dalam kategori terdiagnosa penyakit ISPA sebanyak 53 anak balita, dimana 48 anak balita tepat diklasifikasikan Non ISPA (Tidak Terdiagnosa ISPA) dan 5 tepat disalahkan sebagai terdiagnosa ISPA sehingga diperoleh nilai specificity sebesar 0.906.

Total akurasi (1-APER) yang dihasilkan untuk klasifikasi terdiagnosa penyakit ISPA pada anak balita di Puskesmas Cukir Jombang dengan menggunakan aturan pemilihan pemilah indeks gini sesuai dengan disajikan dalam Tabel 4.9 adalah sebesar 87.8%. Sedangkan total kesalahan klasifikasi (APER) yang dihasilkan sebesar 12,2 %. Karena total akurasi yang cukup tinggi, maka pohon klasifikasi optimal yang terbentuk sudah layak digunakan untuk pengklasifikasian risiko penyakit ISPA pada anak balita.

BAB V

KESIMPULAN DAN SARAN

Setelah analisis dan pembahasan dilakukan, didapatkan beberapa hal yang dapat disimpulkan mengenai klasifikasi Penyakit ISPA pada Anak Balita menggunakan pendekatan *Classification Trees* dan faktor-faktor yang diduga mempengaruhinya yang menjawab dari tujuan penelitian ini. Selain itu, juga terdapat beberapa hal yang dapat disarankan bagi penelitian selanjutnya.

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian adalah sebagai berikut.

1. Berdasarkan analisis *Classification Trees* diketahui bahwa variabel terpenting dan yang berpengaruh signifikan adalah variabel X_4 (Suhu Tubuh). Berdasarkan hasil tersebut diketahui bahwa variabel Suhu Tubuh memiliki kontribusi terbesar dengan nilai skor 100 yang artinya variabel Suhu Tubuh akan menjadi pemilah awal yang disebut sebagai simpul induk (*parent nodes*).
2. Berdasarkan tabel 4.10 dapat diketahui terdapat 103 anak balita yang termasuk dalam kategori ISPA (terdiagnosa penyakit ISPA) dengan 89 tepat diklasifikasikan sebagai anak balita dengan terdiagnosa penyakit ISPA dan 14 anak balita salah diklasifikasikan sebagai anak balita yang tidak terdiagnosa penyakit ISPA. Sehingga nilai sensitivy yang diperoleh sebesar 0.864. Sementara itu jumlah anak balita yang tidak termasuk dalam kategori terdiagnosa penyakit ISPA sebanyak 53 anak balita, dimana 48 anak balita tepat diklasifikasikan Non ISPA (Tidak Terdiagnosa ISPA) dan 5 tepat disalahkan sebagai terdiagnosa ISPA sehingga diperoleh nilai specificity sebesar 0.906.
3. Total akurasi (1-APER) yang dihasilkan untuk klasifikasi terdiagnosa penyakit ISPA pada anak balita di Puskesmas Cukir Jombang dengan menggunakan aturan pemilihan pemilah indeks gini sesuai dengan disajikan dalam Tabel 4.10 adalah sebesar 87.8%. Sedangkan total kesalahan klasifikasi (APER) yang dihasilkan sebesar 12,2 %. Karena total akurasi yang cukup tinggi,

maka pohon klasifikasi optimal yang terbentuk sudah layak digunakan untuk pengklasifikasian risiko penyakit ISPA pada anak balita.

5.2 Saran

Dari kesimpulan yang diperoleh, maka terdapat beberapa hal yang dapat disarankan antara lain sebagai berikut.

1. Diharapkan bagi Puskesmas Cukir agar lebih memperhatikan lagi tentang dampak terkenanya ISPA pada balita sehingga dapat mengurangi balita yang terkena ISPA.
2. Bagi peneliti selanjutnya agar menambahkan variabel independen lain yang diduga mempengaruhi ISPA pada balita di Kabupaten Jombang Jawa Timur. Selain itu, disarankan untuk melakukan kajian mengenai *Classification Trees* lebih lanjut.

DAFTAR PUSTAKA

- Agresti, A. (2007). *Categorical Data Analysis*. New York : Jhon Wiley and Sons.
- Breiman L., Friedman J.H Olshen R.A & Stone C.J. 1993. *Classification And Regression Tree*. New York, NY: Chapman And Hall.
- CART *Reference Guide*. (2000). *CART User's Guide*. San Diego: Salford System.
- Daniel, W. W. (1978). *Statistik Nonparametrik Terapan*. Jakarta : Gramedia.
- Depkes. (2008). *Laporan Nasional riset kesehatan dasar tahun 2007*. Jakarta : Depkes RI.
- Dinas Kesehatan Provinsi Jawa Timur, 2013. Profil Kesehatan Provinsi Jawa Timur. Retrieved 19 Maret, 2017, from dinkes.jatimprov.go.id: http://dinkes.jatimprov.go.id/dokumen/dokumen_publicasi.html
- Dinas Kesehatan Provinsi Jawa Timur, 2014. Profil Kesehatan Provinsi Jawa Timur. Retrieved 19 Maret, 2017, from dinkes.jatimprov.go.id: http://dinkes.jatimprov.go.id/dokumen/dokumen_publicasi.html.
- 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
- 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
- Kim H, Loh WY. *Classification Trees with unbiased multiway splits*. J Am Stat Assoc 2001, 96:589-604
- Le, C. T. (1998). *Applied Categorical Data Analysis*, USA : Jhon Wiley and Sons, Inc.

- Lewis, R. J. 2000. *An introduction to Classification and Regression Tree (CART) Analysis*. Department of Emergency Medicine, Harbor UCLA Medical Center, Torrance, California.
- Lusyanti, Merlina. 2010. *Perbandingan Metode Regresi Logistik Dengan Metode Pohon Klasifikasi Pada Data Polikotomus (Studi Kasus Pada Faktor Yang Berpengaruh Terhadap Keberhasilan Pengobatan Akupuntur Pada Penderita Obesitas Di LP3A Surabaya)*. Skripsi. Surabaya : Institut Teknologi Sepuluh Nopember (tidak dipublikasikan).
- 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.
- Notoatmodjo, S. (2010). *Metodologi Penelitian Kesehatan*. Jakarta : Rineka Cipta
- Pratiwi, F.E. dan Zain, I. 2014. *Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara*. *Sains dan Seni Pomit*, Jurusan Statistika, FMIPA, Institut Teknologi Sepuluh Nopember.
- Prawirohardjo, S. (2014). *Ilmu kebidanan (4th ed.)*. Jakarta : Yayasan Bina Pustaka Sarwono Prawirohardjo.
- Timofeev, R. 2004. *Classification and Regression Trees (CART) Theory and Application*. Berlin : Center of Applied Statistics and Economics Humboldt University.
- Yuniarto. 2009. *Klasifikasi Angkatan Kerja Provinsi Bengkulu Menggunakan Metode CART dan Regresi Logistik*. Surabaya : Institut Teknologi Sepuluh Nopember.
- Yusri. Penyebab Infeksi Saluran Pernafasan Akut. 2011 [cited 2017 28 Desember]; Available from : <http://www.kesehatan123.com/1679/penyebab-isp/>.

LAMPIRAN

Lampiran A.

Data Anak Balita Terhadap Status Penyakit ISPA yang dirujuk di Puskesmas Cukir Kabupaten Jombang Tahun 2016 dan 2017

No.	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
1	0	1	12	52.5	38	100	0	0
2	1	1	8.7	47.5	36	90	1	1
3	0	0	10	49.5	39	120	0	1
4	0	0	12.5	48.5	40	110	0	0
5	0	0	9.7	44	38.5	100	0	0
...
50	1	1	6	43	36	107	1	0
51	1	1	7.3	47	36	104	1	0
52	1	1	10	47	36.4	109	1	0
53	0	1	9.7	46	37.9	96	0	0
54	1	0	9.2	48	36.3	98	1	0
55	0	0	11.6	46	38.4	111	0	1
...
116	1	1	7.4	46	36.8	95	1	1
117	0	0	9.7	50	38.4	116	0	0
118	1	1	11.1	52	36.9	99	1	0
119	1	1	7.6	45	36	93	1	1
120	1	1	11	52.5	36.4	98	1	1
...
169	0	0	7.5	47	37	99	0	0
170	0	1	7.4	47	38	97	0	1
171	0	0	8.8	48.7	36.7	96	1	1
172	1	0	9.8	49	38	94	0	1
173	0	0	10	50	39	99	0	1

Keterangan :

Y (0) : Anak Balita Yang Terdiagnosa Penyakit ISPA

Y(1) : Anak Balita Yang Tidak Terdiagnosa Penyakit ISPA

X1 : Riwayat Pemberian ASI

X2 : Berat Badan Anak Balita (kg)

X3 : Lingkar Kepala (cm)

X4 : Suhu Tubuh (°C)

X5 : Denyut Nadi

X6 : Tingkat Pernafasan (respiratory rate)

X7 : Status Ekonomi Keluarga

Lampiran B.

Kemungkinan pemilah pada Variabel Prediktor berskala Kontinyu

Berat Badan (kilogram)	Lingkar Kepala (cm)	Suhu Tubuh	Denyut Nadi
12	52.5	38	100
8.7	47.5	36	90
10	49.5	39	120
12.5	48.5	40	110
9.7	44	38.5	100
10.7	53	38	114
10.1	54	38	100
8.5	43.5	39	112
7.8	47	36	111
7.1	44	36	108
8.7	45.5	39.7	100
12.7	43	40	110
9.5	46	39	114
10	52	36.7	109
10.1	50	36.6	107
11.5	54	36.9	105
8.8	48	36	97
9.1	48.5	36	94
9.4	47	40.2	115
9.8	47.5	40	100
...

Lampiran C. Statistika Deskriptif antara variabel respon dan prediktor

```
MTB > Describe 'BB+' - 'Nadi-';
SUBC> Mean;
SUBC> StDeviation;
SUBC> Variance;
SUBC> Minimum;
SUBC> Maximum;
SUBC> N;
SUBC> NMissing.
```

Descriptive Statistics: BB+, BB-, LK+, LK-, Temp+, Temp-, Nadi+, Nadi-

Variable	N	Mean	StDev	Variance	Minimum	Maximum
BB+	110	9.075	1.973	3.892	3.000	14.000
BB-	63	8.721	2.062	4.253	2.700	12.000
LK+	110	46.645	3.703	13.714	35.000	54.000
LK-	63	46.627	4.010	16.077	35.000	54.000
Temp+	110	38.315	1.222	1.492	36.000	40.700
Temp-	63	36.495	0.649	0.421	36.000	39.600
Nadi+	110	104.45	7.19	51.72	90.00	120.00
Nadi-	63	96.286	5.589	31.240	90.000	111.000

Lampiran D. *Output Pohon Klasifikasi 10-Fold Cross Validation Estimate*

Cross Validation 1

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

[Breiman adjusted cost, lambda = 0.060]

Node	Class	Wgt	Count	N	Prob	Cost	Parent Complexity
1	0	9.00	9	9	0.045	0.000	0.018 [0.187]
	1	0.00	0	0	1.000		
2	0	9.00	9	9	1.000		
	1	2.00	2	2	0.018	0.000	0.018 [0.252]
3	0	0.00	0	0	0.000		
	1	2.00	2	2	1.000		
4	0	52.00	52	52	0.449	0.045	0.006 [0.083]
	1	4.00	4	4	0.045		
5	0	48.00	48	48	0.955		
	1	4.00	4	4	0.024	0.373	0.006 [0.607]
6	0	3.00	3	3	0.627		
	1	1.00	1	1	0.373		
7	0	2.00	2	2	0.010	0.000	0.010 [0.280]
	1	2.00	2	2	1.000		
8	0	0.00	0	0	0.000		
	1	9.00	9	9	0.065	0.309	0.010 [0.467]
9	0	4.00	4	4	0.309		
	1	5.00	5	5	0.691		
10	0	78.00	78	78	0.390	0.000	0.015 [0.044]
	1	78.00	78	78	1.000		
11	0	0.00	0	0	0.000		
	1	0.00	0	0	0.000		

Lampiran D. (lanjutan)

Cross Validation 2

=====

TERMINAL NODE INFORMATION

=====

[Breiman adjusted cost, lambda = 0.107]

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
1	0	9.00	9	0.044	0.000	0.019 [0.326]
	1	0.00	0	0.000		
2	0	9.00	9	1.000	0.000	0.019 [0.390]
	1	2.00	2	0.019		
3	0	0.00	0	0.000	0.073	0.025 [0.158]
	1	2.00	2	1.000		
4	0	53.00	53	0.468	0.000	0.010 [0.421]
	1	7.00	7	0.073		
5	0	46.00	46	0.927	0.292	0.010 [0.574]
	1	2.00	2	0.010		
6	0	2.00	2	1.000	0.000	0.018 [0.098]
	1	0.00	0	0.000		

Lampiran D. (lanjutan)**Cross Validation 3**

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

[Breiman adjusted cost, lambda = 0.070]

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
1	0	9.00	9	0.043	0.000	0.029 [0.215]
	1	0.00	0	0.000		
2	0	9.00	9	1.000	0.000	0.029 [0.244]
	1	3.00	3	0.029		
3	0	0.00	0	0.000	0.070	0.021 [0.115]
	1	3.00	3	1.000		
4	0	7.00	7	0.070	0.393	0.010 [0.598]
	1	45.00	45	0.930		
5	0	7.00	7	0.048	0.000	0.010 [0.051]
	1	4.00	4	0.393		
5	0	3.00	3	0.607	0.000	0.010 [0.051]
	1	85.00	85	0.405		
5	0	85.00	85	1.000	0.000	0.010 [0.051]
	1	0.00	0	0.000		

Cross Validation 4

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

[Breiman adjusted cost, lambda = 0.100]

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
1	0	9.00	9	0.046	0.000	0.026 [0.324]
	1	0.00	0	0.000		
1	0	9.00	9	1.000	0.000	0.026 [0.324]
	1	0.00	0	0.000		

Lampiran D. (lanjutan)

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
2	1	3.00	3	0.026	0.000	0.026 [0.375]
	0	0.00	0	0.000		
	1	3.00	3	1.000		
3	1	57.00	57	0.467	0.077	0.023 [0.160]
	0	7.00	7	0.077		
	1	50.00	50	0.923		
4	0	2.00	2	0.010	0.000	0.010 [0.429]
	0	2.00	2	1.000		
	1	0.00	0	0.000		
5	1	9.00	9	0.064	0.321	0.010 [0.610]
	0	4.00	4	0.321		
	1	5.00	5	0.679		
6	0	76.00	76	0.388	0.000	0.012 [0.097]
	0	76.00	76	1.000		
	1	0.00	0	0.000		

Cross Validation 5

=====

TERMINAL NODE INFORMATION

=====

[Breiman adjusted cost, lambda = 0.082]

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
1	0	9.00	9	0.044	0.000	0.028 [0.300]
	0	9.00	9	1.000		
	1	0.00	0	0.000		
2	1	3.00	3	0.028	0.000	0.028 [0.342]
	0	0.00	0	0.000		

Lampiran D. (lanjutan)

Node	Class	Wgt Count	N	Prob	Cost	Parent Complexity
	1	3.00	3	1.000		
3	1	52.00	52	0.459	0.074	0.022 [0.144]
	0	7.00	7	0.074		
	1	45.00	45	0.926		
4	1	11.00	11	0.076	0.382	0.018 [0.620]
	0	6.00	6	0.382		
	1	5.00	5	0.618		
5	0	81.00	81	0.393	0.000	0.018 [0.079]
	0	81.00	81	1.000		
	1	0.00	0	0.000		

Lampiran E. Hasil Ketepatan Klasifikasi (Accuracy)Akurasi Data *Learning***Cross Validation Trees 1**=====

CROSS VALIDATION CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	0	1	
0	89.00	11.00	100.00
1	3.00	53.00	56.00
PRED. TOT.	92.00	64.00	156.00
CORRECT	0.890	0.946	
SUCCESS IND.	0.249	0.587	
TOT. CORRECT	0.910		
SENSITIVITY:	0.890	SPECIFICITY:	0.946
FALSE REFERENCE:	0.033	FALSE RESPONSE:	0.172
REFERENCE = "0", RESPONSE = "1"			

Lampiran E. (lanjutan)**Cross Validation Trees 2**

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	91.00	12.00	103.00
1	6.00	47.00	53.00
PRED. TOT.	97.00	59.00	156.00
CORRECT	0.883	0.887	
SUCCESS IND.	0.223	0.547	
TOT. CORRECT	0.885		
SENSITIVITY: 0.883		SPECIFICITY: 0.887	
FALSE REFERENCE: 0.062		FALSE RESPONSE: 0.203	
REFERENCE = "0", RESPONSE = "1"			

Cross Validation Trees 3

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	93.00	12.00	105.00
1	3.00	48.00	51.00
PRED. TOT.	96.00	60.00	156.00
CORRECT	0.886	0.941	
SUCCESS IND.	0.213	0.614	
TOT. CORRECT	0.904		
SENSITIVITY: 0.886		SPECIFICITY: 0.941	
FALSE REFERENCE: 0.031		FALSE RESPONSE: 0.200	
REFERENCE = "0", RESPONSE = "1"			

Lampiran E. (lanjutan)**Cross Validation Trees 4**

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	85.00	13.00	98.00
1	6.00	52.00	58.00
PRED. TOT.	91.00	65.00	156.00
CORRECT	0.867	0.897	
SUCCESS IND.	0.239	0.525	
TOT. CORRECT	0.878		
SENSITIVITY: 0.867		SPECIFICITY: 0.897	
FALSE REFERENCE: 0.066		FALSE RESPONSE: 0.200	
REFERENCE = "0", RESPONSE = "1"			

Cross Validation Trees 5

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	89.00	14.00	103.00
1	5.00	48.00	53.00
PRED. TOT.	94.00	62.00	156.00
CORRECT	0.864	0.906	
SUCCESS IND.	0.204	0.566	
TOT. CORRECT	0.878		
SENSITIVITY: 0.864		SPECIFICITY: 0.906	
FALSE REFERENCE: 0.053		FALSE RESPONSE: 0.226	
REFERENCE = "0", RESPONSE = "1"			

Lampiran E. (lanjutan)**Cross Validation Trees 6**

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	88.00	11.00	99.00
1	1.00	56.00	57.00
PRED. TOT.	89.00	67.00	156.00
CORRECT	0.889	0.982	
SUCCESS IND.	0.254	0.617	
TOT. CORRECT	0.923		
SENSITIVITY: 0.889		SPECIFICITY: 0.982	
FALSE REFERENCE: 0.011		FALSE RESPONSE: 0.164	
REFERENCE = "0", RESPONSE = "1"			

Cross Validation Trees 7

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	89.00	14.00	103.00
1	3.00	50.00	53.00
PRED. TOT.	92.00	64.00	156.00
CORRECT	0.864	0.943	
SUCCESS IND.	0.204	0.604	
TOT. CORRECT	0.891		
SENSITIVITY: 0.864		SPECIFICITY: 0.943	
FALSE REFERENCE: 0.033		FALSE RESPONSE: 0.219	
REFERENCE = "0", RESPONSE = "1"			

Lampiran E. (lanjutan)**Cross Validation Trees 8**

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	83.00	10.00	93.00
1	5.00	57.00	62.00

PRED. TOT.	88.00	67.00	155.00
CORRECT	0.892	0.919	
SUCCESS IND.	0.292	0.519	
TOT. CORRECT	0.903		
SENSITIVITY:	0.892	SPECIFICITY:	0.919
FALSE REFERENCE:	0.057	FALSE RESPONSE:	0.149
REFERENCE = "0", RESPONSE = "1"			

Cross Validation Trees 9

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	83.00	9.00	92.00
1	5.00	58.00	63.00

PRED. TOT.	88.00	67.00	155.00
CORRECT	0.902	0.921	
SUCCESS IND.	0.309	0.514	
TOT. CORRECT	0.910		
SENSITIVITY:	0.902	SPECIFICITY:	0.921
FALSE REFERENCE:	0.057	FALSE RESPONSE:	0.134
REFERENCE = "0", RESPONSE = "1"			

Lampiran E. (lanjutan)**Cross Validation Trees 10**

```

=====
CROSS VALIDATION CLASSIFICATION TABLE
=====

```

Actual Class	Predicted Class		Actual Total
	0	1	
0	87.00	7.00	94.00
1	5.00	56.00	61.00
PRED. TOT.	92.00	63.00	155.00
CORRECT	0.926	0.918	
SUCCESS IND.	0.319	0.524	
TOT. CORRECT	0.923		
SENSITIVITY: 0.926		SPECIFICITY: 0.918	
FALSE REFERENCE: 0.054		FALSE RESPONSE: 0.111	
REFERENCE = "0", RESPONSE = "1"			

Lampiran F.

Informasi Simpul Terminal Pohon Klasifikasi Optimal

=====

NODE INFORMATION

=====

```
*****
*   Node 1: TEMPERAT   *
*         N: 156       *
*****
```

```
*****
*   Node 2             *
*         N: 64        *
*****
```

```
*****
*   Node 4             *
*         N: 92        *
*****
```

Node 1 was split on TEMPERAT

A case goes left if TEMPERAT \leq 36.950

Improvement = 0.283 Complexity Threshold = 0.375

Node	Cases	Wgt Counts	Cost Class
1	156	156.00	0.500 0
2	64	64.00	0.146 1
4	92	92.00	0.100 0

Class	Weighted Counts		
	Top	Left	Right
0	103.00	16.00	87.00
1	53.00	48.00	5.00

Class	Within Node Probabilities		
	Top	Left	Right
0	0.500	0.146	0.900
1	0.500	0.854	0.100

Surrogate	Split	Assoc.	Improve.
1 RR	s 1	0.632	0.271
2 DENYUT_N	s 98.500	0.471	0.157
3 RIWAYAT	s 1	0.437	0.136
4 BB	s 9.350	0.203	0.005
5 LK	s 45.250	0.068	0.003
Competitor	Split	Improve.	
1 RR	0	0.271	
2 DENYUT_N	98.500	0.157	

Lampiran F. (lanjutan)

3 RIWAYAT	0		0.136
4 LK		51.750	0.018
5 STATUS_E	0		0.011

```
*****
*           Node 2: RR           *
*           N: 64                 *
*****
```

```
*****
*           Node 3           *
*           N: 12           *
*****
=====
= Terminal Node 3 =
= N: 52           =
=====
```

Node 2 was split on RR

A case goes left if RR = (0)

Improvement = 0.035 Complexity Threshold = 0.022

Node	Cases	Wgt Counts	Cost Class
2	64	64.00	0.146 1
3	12	12.00	0.393 0
-3	52	52.00	0.074 1

Class	Weighted Counts		
	Top	Left	Right
0	16.00	9.00	7.00
1	48.00	3.00	45.00

Class	Within Node Probabilities		
	Top	Left	Right
0	0.146	0.607	0.074
1	0.854	0.393	0.926

Surrogate	Split	Assoc.	Improve.
1 LK	s 36.500	0.067	0.007

Competitor	Split	Improve.
1 RIWAYAT	0	0.011
2 DENYUT_N	96.500	0.008
3 LK	36.500	0.007
4 BB	8.850	0.005
5 STATUS_E	0	0.002

Lampiran F. (lanjutan)

```
*****
*   Node 3: LK   *
*   N: 12       *
*****
```

```
=====
= Terminal Node 1 =
=   N: 9         =
=====
```

```
=====
= Terminal Node 2 =
=   N: 3         =
=====
```

Node 3 was split on LK

A case goes left if LK <= 47.500

Improvement = 0.034 Complexity Threshold = 0.028

Node	Cases	Wgt	Counts	Cost	Class
3	12		12.00	0.393	0
-1	9		9.00	0.000	0
-2	3		3.00	0.000	1

Class	Weighted Counts		
	Top	Left	Right
0	9.00	9.00	0.00
1	3.00	0.00	3.00

Class	Within Node Probabilities		
	Top	Left	Right
0	0.607	1.000	0.000
1	0.393	0.000	1.000

Surrogate		Split	Assoc.	Improve.
1 BB	s	8.350	0.828	0.026
2 DENYUT_N	r	94.500	0.333	0.008
3 STATUS_E	s 1		0.142	0.008

Competitor		Split	Improve.
1 BB		8.350	0.026
2 STATUS_E	1		0.008
3 DENYUT_N		94.500	0.008
4 TEMPERAT		36.850	0.003

Lampiran F. (lanjutan)

```
*****
*   Node 4: DENYUT_N *
*           N: 92     *
*****
```

```
=====
=   Terminal Node 4   =           =   Terminal Node 5   =
=     N: 11           =           =     N: 81           =
=====
```

Node 4 was split on DENYUT_N

A case goes left if DENYUT_N <= 95.500

Improvement = 0.049 Complexity Threshold = 0.018

Node	Cases	Wgt Counts	Cost Class
4	92	92.00	0.100 0
-4	11	11.00	0.382 1
-5	81	81.00	0.000 0

Class	Weighted Counts		
	Top	Left	Right
0	87.00	6.00	81.00
1	5.00	5.00	0.00

Class	Within Node Probabilities		
	Top	Left	Right
0	0.900	0.382	1.000
1	0.100	0.618	0.000

Competitor	Split	Improve.
1 RR	0	0.019
2 RIWAYAT	0	0.008
3 TEMPERAT	37.200	0.006
4 BB	8.750	0.006
5 LK	48.850	0.005

```
=====
```

Lampiran F. (lanjutan)

TERMINAL NODE INFORMATION

=====

[Breiman adjusted cost, lambda = 0.082]

Node	Class	Wgt	Count	N	Parent	
					Prob	Cost Complexity
1	0	9.00	9	0.044	0.000	0.028
					[0.300]	
0		9.00	9	1.000		
1		0.00	0	0.000		
2	1	3.00	3	0.028	0.000	0.028
					[0.342]	
0		0.00	0	0.000		
1		3.00	3	1.000		
3	1	52.00	52	0.459	0.074	0.022
					[0.144]	
0		7.00	7	0.074		
1		45.00	45	0.926		
4	1	11.00	11	0.076	0.382	0.018
					[0.620]	
0		6.00	6	0.382		
1		5.00	5	0.618		
5	0	81.00	81	0.393	0.000	0.018
					[0.079]	
0		81.00	81	1.000		
1		0.00	0	0.000		

Node	Learn		
1	9.000	9.000	0.000
2	3.000	0.000	3.000
3	52.000	7.000	45.000
4	11.000	6.000	5.000
5	81.000	81.000	0.000

Lampiran G. Surat Perizinan Pengambilan Data



PEMERINTAH KABUPATEN JOMBANG
DINAS KESEHATAN

JL. KH. Wahid Hasyim No. 131 Jombang. Kode Pos : 61411
Telp/Fax. (0321) 866197 Email : dinkesjombang@yahoo.com
Website : www.jombangkab.go.id

Jombang, 4 Desember 2017

Nomor : 070/11220/415.17/2017
Sifat : Biasa
Lampiran : -
Perihal : Izin Penelitian

K e p a d a
Yth. 1. Kepala Puskesmas **Mojoagung**
2. Kepala Puskesmas **Cukir**
3. Kepala Puskesmas **Ploso**
di -

JOMBANG

Menindaklanjuti Surat dari Badan Pelayanan Perizinan Kabupaten Jombang Nomor : 072/245/415.35/2017 tanggal 04 Desember 2017 tentang Izin Penelitian. Maka mohon berkenan Puskesmas Saudara sebagai tempat penelitian mahasiswa Fakultas Matematika, Komputasi dan Sains Data, Institut Teknologi Sepuluh Nopember Surabaya . Adapun nama mahasiswanya adalah :

Nama : ARIF BUDHIMAN
NIM : 0621154500039
Judul Penelitian : Evaluasi Ketepatan Klasifikasi ISPA pada Anak Balita menggunakan metode Classification Trees di Puskesmas Kabupaten Jombang.
Catatan : - Tidak mengganggu kegiatan pelayanan
- Segala sesuatu yang terkait dengan kegiatan / pembimbingan di lapangan agar dimusyawarahkan bersama dengan yang bersangkutan.

Demikian atas perhatian dan kerjasama yang baik diucapkan terima kasih.

Plt. KEPALA DINAS KESEHATAN
KABUPATEN JOMBANG



Drg. INNA SILESTYOWATI, M.Kes.
NIP. 196906232002122001

Lampiran H. Rekam Medik Rawat Jalan Puskesmas Cukir


PUSKESMAS CUKIR
NO. RM 0 1 7 2 0 5

NO	TGL / JAM	SUBJEKTIF (ANAMNESIS)	OBJEKTIF (PEMERIKSAAN)	ASSESSMENT (DIAGNOSA)	PLANNING TERAPI & TINDAKAN	TTD PEMERIKSA
		REKAM MEDIK RAWAT JALAN Nama : _____ Tgl lahir : _____ Alamat : _____ Riwayat Penyakit Dahulu : _____ Riwayat Alergi : _____				
	12 JAN 2018 27, 18	KIA Demam (2)	O: BB: 1/2 TB: TD: N: RR: 7, 36.0 TFU: letak: Dji: Odema: O: BB: 6.6 T: 37.4 N: 100 RR: 32 O: BB: 6.6 T: 37.8 N: 100 RR: 30 HA : ausc : Tcr : Jc : Rnc :	200.1 106.	BEG. & polio 1. Revisi Parasetamol 3 x 1/2 - III BC 1 x 1/2 - V dr. R. MAULIDINA Nip: 19831201001 dr. R. MAULIDINA Nip: 19831201001	

RM 25a

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Tak Kenal Maka Taa'ruf. Penulis yang bernama lengkap Arif Budhiman, yang biasa disapa Arif atau Tejjos merupakan anak ketiga dari pasangan Bpk. Anang Sumarhyanto dan ibu Harini dari tiga bersaudara. Penulis dilahirkan di Kota Santri (Jombang) pada tanggal 6 Juli 1992 dan bertempat tinggal di Jalan Pattimura Gg V Blok D/56 Kelurahan sengon RT 28 RW 6 Kecamatan Jombang Kabupaten Jombang, Jawa Timur. Sebelum memasuki dunia perkuliahan, penulis juga telah menempuh pendidikan formal, yang pernah ditempuh penulis adalah SDN Jombatan III Jombang (Betari), SMPN 2 Jombang, SMA PGRI 1 Jombang, Diploma tiga Statistika FMIPA ITS dan pada tahun 2015 diterima menjadi mahasiswa jurusan lintas jalur Statistika FMKSD ITS dengan NRP 1315 105 039. Mahasiswa ini juga pernah aktif dalam beberapa organisasi mahasiswa di ITS yaitu pada periode tahun 2012/2013 sebagai Staff Tabligh Syiar FORSIS, Staff BPM (Badan Pelaksana Mentoring) JMII ITS, Cinta Rebana ITS (CR), IBC Bulu tangkis ITS. Selain itu, penulis juga aktif dalam berbagai bidang kegiatan kemahasiswaan lainnya yang tidak bisa di sebutkan satu persatu. Apabila pembaca ingin berdiskusi mengenai tugas akhir ini atau matari yang berhubungan dengan ini dan segala kritik serta saran dari pembaca untuk penulis bisa disampaikan melalui email ariftejjos@gmail.com.