



TUGAS AKHIR - SS 091324

KLASIFIKASI KEGIATAN PARTISIPASI EKONOMI PEREMPUAN
DI JAWA TIMUR DENGAN PENDEKATAN CART
(*CLASSIFICATION AND REGRESSION TREES*)

Sharfina Widyandini
NRP 1311 106 002

Dosen Pembimbing
Dr. Vita Ratnasari, S.Si, M.Si

JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2014



TUGAS AKHIR - SS 091324

**KLASIFIKASI KEGIATAN PARTISIPASI EKONOMI PEREMPUAN
DI JAWA TIMUR DENGAN PENDEKATAN CART
(*CLASSIFICATION AND REGRESSION TREES*)**

Sharfina Widyandini
NRP 1311 106 002

Dosen Pembimbing
Dr. Vita Ratnasari, S.Si, M.Si

JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2014



FINAL PROJECT - SS 091324

**CLASSIFICATION OF WOMEN ECONOMIC PARTICIPATION
ACTIVITY IN EAST JAVA BY USING CART (CLASSIFICATION
AND REGRESSION TREES)**

Sharfina Widyandini
NRP 1311 106 002

Supervisor
Dr. Vita Ratnasari, S.Si, M.Si

DEPARTMENT OF STATISTICS
Faculty of Mathematics and Natural Sciences
Institut Teknologi Sepuluh Nopember
Surabaya 2014

KLASIFIKASI KEGIATAN PARTISIPASI EKONOMI PEREMPUAN DI JAWA TIMUR DENGAN PENDEKATAN CART (*CLASSIFICATION AND REGRESSION TREES*)

Nama Mahasiswa : Sharfina Widyandini
NRP : 1311 106 002
Jurusan : Statistika FMIPA-ITS
Dosen Pembimbing : Dr. Vita Ratnasari, S.Si, M.Si.

Abstrak

Seiring dengan bertambah pesatnya pertumbuhan penduduk di Indonesia, peran dan partisipasi seluruh masyarakat Indonesia, termasuk perempuan sangat dibutuhkan demi terwujudnya suatu pembangunan yang berkesinambungan. Bentuk kegiatan partisipasi ekonomi perempuan dibedakan menjadi dua kategori yaitu bekerja dan tidak bekerja. Sumber data yang digunakan ialah data sekunder dari SUSENAS tahun 2011, dengan unit penelitian individu, perempuan usia kerja khususnya umur 15-64 tahun di Propinsi Jawa Timur. Kombinasi data learning dan testing yang digunakan pada data kabupaten adalah 90%;10% dengan ketepatan klasifikasi data testing sebesar 64,6%, sementara kotamadya dengan kombinasi 85%;15% menghasilkan ketepatan klasifikasi data testing sebesar 60,1%. Pohon optimal kabupaten menghasilkan 18 simpul terminal dan 8 kedalaman. Pohon optimal di kotamadya menghasilkan 6 simpul terminal dan 5 kedalaman. Hasil karakteristik dan klasifikasi menunjukkan bahwa terdapat perbedaan kegiatan partisipasi dengan faktor yang mempengaruhinya, antara perempuan di kabupaten dan di kotamadya Propinsi Jawa Timur. Variabel umur menjadi variabel dengan skor tertinggi atau dengan kata lain memiliki kontribusi terbesar dalam pemilahan pada pohon klasifikasi optimal.

Kata Kunci: Partisipasi Perempuan, Data Learning-Testing, Klasifikasi

“Halaman Ini Sengaja Dikosongkan”

CLASSIFICATION OF WOMEN ECONOMIC PARTICIPATION ACTIVITY IN EAST JAVA BY USING CART (CLASSIFICATION AND REGRESSION TREES)

Name of student : Sharfina Widyandini
ID : 1311 106 002
Department : Statistika FMIPA-ITS
Advisor : Dr. Vita Ratnasari, S.Si, M.Si.

Abstract

Along with the rapid growth of population in Indonesia, the role and participation of all Indonesian people, including women, are needed in order to realize a sustainable development. The form of women's economic participation activities can be divided into two categories they are working and not working. Source of data used are secondary data from SUSENAS in 2011, by using individual research, the carrier women especially by the age of 15-64 years in East Java province. The combination of learning and testing data used in district area is 90 %, 10 % to the accuracy of testing data classification of 64.6 % while municipalities with a combination of 85 %, 15 % resulted in a classification testing data accuracy of 60.1 %. Optimal tree produces 18 terminal and 8 nodes. The optimal tree in municipality produces 6 terminals and 5 nodes. The results of characteristics and classification indicate that there are differences in the participation activities with the factors that influence it, among women in districts and municipalities in East Java province. Age variable becomes variable with the highest score or in other words has the largest contribution in sorting the optimal classification tree.

Keywords : *Women's participation, learning and testing data, classification*

LEMBAR PENGESAHAN

KLASIFIKASI KEGIATAN PARTISIPASI EKONOMI PEREMPUAN DI JAWA TIMUR DENGAN PENDEKATAN CART (*CLASSIFICATION AND REGRESSION TRESS*)

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana

Pada

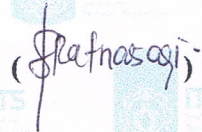
Program Studi S-1 Jurusan Statistika
Fakultas Matematika Dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh:

SHARFINA WIDYANDINI
NRP. 1311 106 002

Disetujui oleh Pembimbing Tugas Akhir :

Dr. Vita Ratnasari, S.Si, M.Si.
NIP. 19700910 199702 2 001



Mengetahui,
Ketua Jurusan Statistika FMIPA - ITS



Dr. Muhammad Mashuri, MT.
NIP. 19620408 198701 1 001

SURABAYA, JANUARI 2014

KATA PENGANTAR

Alhamdulillah atas segala rahmat dan hidayah yang telah dilimpahkan Allah SWT, serta bimbingan dan petunjuk kepada penulis sehingga Tugas Akhir dengan judul “KLASIFIKASI KEGIATAN PARTISIPASI EKONOMI PEREMPUAN DENGAN PENDEKATAN CART (CLASSIFICATION AND REGRESSION TREES)” sebagai salah satu syarat untuk memperoleh gelar sarjana dalam ilmu statistik pada Fakultas Matematika dan Ilmu Pengetahuan Alam ini dapat terselesaikan dengan baik, tepat pada waktunya. Dengan segala kerendahan hati penulis mengucapkan terima kasih kepada.

1. Mama, papa, kedua adikku dan mas yang tiada henti memberikan dukungan moriil maupun materiil serta doa kepada penulis. Terima kasih yang tak terhingga atas waktu, tenaga, pikiran, nasehat, perhatian dan arahnya.
2. Bapak Dr. Muhammad Mashuri. MT selaku Ketua Jurusan Statistika FMIPA ITS sekaligus dosen wali.
3. Ibu Dr. Vita Ratnasari. M.Si. selaku dosen pembimbing yang banyak memberikan arahan, bimbingan juga motivasi dalam penyelesaian Tugas Akhir ini.
4. Bapak Prof. Dr. I Nyoman Budiantara. M.Si selaku ketua laboratorium sosial pemerintahan sekaligus dosen penguji.
5. Ibu Dr. Ismaini Zain. M.Si. selaku dosen penguji yang banyak memberikan masukan demi kesempurnaan Tugas Akhir ini.
7. Bapak/Ibu dosen dan karyawan jurusan statistika ITS yang telah banyak membantu penulis selama masa perkuliahan. Semoga ilmu yang telah diberi akan selalu bermanfaat.
8. Teman-teman seperjuangan angkatan 2008 yang masih setia mendampingi dan berbagi, juga yang telah banyak membantu penulis. Henny, Listy, Lunny, Nindy, Wida, Ditago, Febty, Pucin, Karin, Rindang, dkk.
9. Teman-teman lintas jalur angkatan 2011 dan 2012 serta semua pihak yang tidak dapat disebut satu per satu. Terima kasih atas doa, semangat, motivasi, kekompakan yang luar biasa.

Semoga segala amal dan kebaikan yang telah mereka berikan kepada penulis mendapatkan balasan yang setimpal dari Allah SWT.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna, oleh sebab itu segala bentuk saran dan kritik yang dapat membantu untuk memperbaiki Tugas Akhir sangat diperlukan. Akhirnya penulis berharap semoga laporan ini dapat memberi manfaat bagi semua pihak yang membutuhkan.

Surabaya, Januari 2014

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	
LEMBAR PENGESAHAN	
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xviii
BAB I. PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat.....	3
1.5 Batasan Penelitian.....	4
BAB II. TINJAUAN PUSTAKA	
2.1 Penelitian Terdahulu.....	5
2.2 <i>Cross Tabulation</i>	6
2.3 <i>Classification and Regression Tress (CART)</i>	6
2.3.1 Pembentukan Pohon Klasifikasi Maksimal.....	7
2.3.2 Penghentian Pembentukan Pohon Klasifikasi.....	10
2.3.3 Pemangkasan Pohon Klasifikasi.....	10
2.3.4 Penentuan Pohon Klasifikasi Optimal.....	10
2.3.5 Ketepatan Klasifikasi.....	11
2.4 Kegiatan Partisipasi Ekonomi Perempuan.....	11
BAB III. METODOLOGI PENELITIAN	
3.1 Sumber Data.....	15
3.2 Identifikasi Variabel.....	15
3.3 Metode Analisis Data.....	16

BAB IV. ANALISIS DAN PEMBAHASAN

4.1 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan di Jawa Timur	21
4.1.1 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota	21
4.1.2 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur	22
4.1.3 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan.....	23
4.1.4 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga	23
4.1.5 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi.....	24
4.2 <i>Classification and Regression Trees</i> (CART)	26
4.2.1 Klasifikasi Pohon di Kabupaten Propinsi Jawa Timur	27
4.2.2 Klasifikasi Pohon di Kotamadya Propinsi Jawa Timur	33
4.2.3 Ketepatan Klasifikasi	37

BAB V. KESIMPULAN DAN SARAN

5.1 Kesimpulan	39
5.2 Saran.....	40

DAFTAR PUSTAKA	41
-----------------------------	----

LAMPIRAN	43
-----------------------	----

BIODATA PENULIS	69
------------------------------	----

DAFTAR GAMBAR

Gambar Judul	Halaman
2.1 Pembentukan Pohon Klasifikasi.....	7
2.2 Konsep Penduduk, Usia Kerja, Angkatan Kerja, Bekerja dan Pengangguran	13
3.1 Algoritma CART.....	18
3.2 Diagram Alir Penelitian.....	19
4.1 Klasifikasi Desa dan Kota di Jawa Timur.	21
4.2 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota di Kabupaten.	22
4.3 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota di Kotamadya.	22
4.4 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur di Kabupaten.....	22
4.5 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur di Kotamadya.....	22
4.6 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan di Kabupaten	23
4.7 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan di Kotamadya	23
4.8 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga di Kabupaten.....	24
4.9 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga di Kotamadya.....	24
4.10 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi di Kabupaten	25
4.11 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi di Kotamadya	25
4.12 Pohon Klasifikasi Maksimal di Kabupaten	28
4.13 Plot <i>Relative Cost</i> dengan Jumlah Node di Kabupaten	28

4.14 Pohon Klasifikasi Optimal di Kabupaten	29
4.15 Pohon Klasifikasi Maksimal di Kotamadya	34
4.16 Plot <i>Relative Cost</i> dengan Jumlah Node di Kotamadya	34
4.17 Pohon Klasifikasi Optimal di Kotamadya	35

DAFTAR TABEL

Tabel Judul	Halaman
2.1 <i>Cross Tabulation</i> rxc	6
2.2 Hasil Klasifikasi	11
3.1 Variabel Penelitian	15
4.1 Hasil <i>Crosstab</i> Variabel Respon Dengan Prediktor.	26
4.2 Perbandingan Ketepatan Klasifikasi Antar Kombinasi Data	27
4.3 Skor Variabel Prediktor Pada Pohon Maksimal di Kabupaten.....	28
4.4 <i>Tree Sequence</i> Pohon di Kabupaten.....	29
4.5 Skor Variabel Prediktor Pada Pohon Optimal di Kabupaten	30
4.6 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kabupaten.....	30
4.7 Skor Variabel Prediktor Pada Pohon Maksimal di Kotamadya.....	33
4.4 <i>Tree Sequence</i> Pohon di Kotamadya.....	35
4.9 Skor Variabel Prediktor Pada Pohon Optimal di Kotamadya	35
4.10 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kotamadya.....	36
4.11 Ketepatan Klasifikasi Data Testing Pohon Optimal di Kabupaten.....	37
4.12 Ketepatan Klasifikasi Data Testing Pohon Optimal di Kotamadya.....	37

BAB I

PENDAHULUAN

Bab pendahuluan ini terdiri dari latar belakang, lalu rumusan masalah, tujuan penelitian, manfaat serta batasan penelitian. Adapun berikut uraiannya.

1.1 Latar Belakang

Sumber daya manusia merupakan faktor produksi yang penting. Hal ini dikarenakan manusia sebagai pelaku sekaligus tujuan dari pembangunan. Seperti yang telah diketahui bahwa Indonesia menduduki peringkat empat di dunia dengan penduduk terbanyak (Biro Statistik, 2012). Seiring dengan bertambah pesatnya pertumbuhan penduduk di Indonesia, peran dan partisipasi seluruh masyarakat Indonesia sangat dibutuhkan demi terwujudnya suatu pembangunan yang berkesinambungan. Menyikapi hal tersebut, peran perempuan sangat diperlukan untuk mempercepat pembangunan di era globalisasi ini.

Namun, partisipasi perempuan dalam kegiatan perekonomian dapat dikatakan masih kurang optimal, karena berbagai sebab. Kajian yang dilakukan oleh Badan Penelitian dan Pengembangan Hak Asasi Manusia pada tahun 2003 tentang perjalanan karir perempuan dalam jabatan eksekutif dan legislatif menunjukkan bahwa terdapat sejumlah kendala yang membatasi aktifitas perempuan untuk berkarir.

Pada tahun 2008 hasil evaluasi menunjukkan adanya perubahan kearah yang lebih baik meskipun hal tersebut belum bisa dikatakan sebagai suatu hasil yang memuaskan. Pemahaman baru dalam penafsiran agama, membuka peluang bagi perempuan untuk meniti karir sampai puncak. Begitu pula dengan tingkat pendidikan. Kesempatan pendidikan bagi laki-laki dan perempuan tidak terlalu terlihat kesenjangannya. Kegiatan perekonomian tentulah dijalankan oleh suatu tenaga kerja. Nilai tenaga kerja dicerminkan dari upah yang didapatkan. Hal inilah yang akan berdampak langsung terhadap suatu pembangunan

Indikator partisipasi perempuan dapat diukur dengan Tingkat Partisipasi Angkatan Kerja (TPAK) perempuan, yaitu rasio antara jumlah angkatan kerja dengan jumlah penduduk usia kerja. BPS Jatim (2012) menyatakan bahwa TPAK kabupaten/kotamadya di Jawa Timur tahun 2011 menurut jenis kelamin, masih menunjukkan pola yang serupa dengan kondisi sebelumnya. TPAK untuk laki-laki (84,70) terpaut cukup jauh dengan TPAK perempuan (55,01).

Sementara itu data jumlah penduduk hasil proyeksi, sebesar 37.687.662 jiwa, dimana Kota Surabaya memiliki jumlah penduduk terbesar yaitu 2.781.047 jiwa, diikuti Kabupaten Malang 2.459.982 jiwa dan Kabupaten Jember 2.345.851 jiwa. BPS Jatim juga menyebutkan bahwa kepadatan penduduk di kotamadya lebih tinggi dibanding dengan di kabupaten.

Mengacu pada jumlah penduduk di Jawa Timur yang relatif tinggi serta masih terpaut cukup jauhnya nilai TPAK perempuan dan laki-laki, maka peneliti mengklasifikasikan penduduk perempuan usia kerja di Jawa Timur berdasarkan kegiatan partisipasi ekonominya, guna mengetahui pengelompokan kegiatan partisipasi ekonomi perempuan sesuai dengan variabel yang mempengaruhinya. Sehingga diharapkan mampu memberi informasi kepada pemerintah untuk mencapai pembangunan yang lebih terarah.

Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor dapat digunakan kurva regresi. Salah satunya adalah regresi nonparametrik, regresi ini mendapat perhatian yang sangat luas dari para peneliti, karena memiliki fleksibilitas yang tinggi dalam mengestimasi kurva regresi. Contoh dari regresi nonparametrik adalah *Classification And Regression Trees* (CART). Metode CART bertujuan untuk mengklasifikasikan suatu kelompok observasi atau sebuah observasi ke dalam suatu sub kelompok dari suatu kelas-kelas yang diketahui. Dibandingkan dengan metode regresi biasa, CART lebih mudah untuk diinterpretasikan karena hasil analisis berupa topologi pohon atau berupa grafis (Lewis dan Roger,

2000), lebih akurat dan lebih cepat penghitungannya. Metode ini juga merupakan metode yang bisa diterapkan untuk himpunan data dalam jumlah besar, variabel yang sangat banyak dan dengan skala variabel campuran melalui prosedur pemilahan biner.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang telah diuraikan pada latar belakang, maka dalam penelitian ini fokus permasalahan yang akan diselidiki oleh peneliti adalah sebagai berikut.

1. Bagaimana karakteristik kegiatan partisipasi ekonomi perempuan di Jawa Timur?
2. Bagaimana klasifikasi kegiatan partisipasi ekonomi perempuan di Jawa Timur?

1.3 Tujuan Penelitian

Adapun tujuan penelitian berdasarkan rumusan masalah diatas adalah sebagai berikut.

1. Mendeskripsikan karakteristik kegiatan partisipasi ekonomi perempuan di Jawa Timur.
2. Mengklasifikasikan kegiatan partisipasi ekonomi perempuan di Jawa Timur.

1.4 Manfaat

Manfaat yang diperoleh dalam penelitian ini adalah

1. Dapat memberi gambaran kepada pemerintah mengenai partisipasi perempuan dalam memberikan kontribusi terhadap pembangunan ekonomi sehingga mampu melahirkan suatu pembangunan yang terarah dan berorientasi tinggi.
2. Menambah pengetahuan penerapan metode statistik dalam aplikasi di bidang sosial pada peningkatan partisipasi perempuan dalam kegiatan ekonomi khususnya metode klasifikasi pohon.

1.5 Batasan Penelitian

Batasan penelitian ini adalah mengklasifikasikan kegiatan partisipasi ekonomi perempuan di Jawa Timur berdasarkan variabel prediktor yang telah ditentukan. Mengacu pada penelitian yang telah dilakukan oleh Kelly dan Evans (2002) bahwa tingkat partisipasi angkatan kerja usia 65 tahun sampai 69 tahun tidak terlalu mempengaruhi keterlibatan angkatan kerja, maka data yang digunakan dibatasi khusus perempuan usia kerja kurang dari 65 tahun. Sehingga, responden dalam penelitian ini adalah perempuan usia kerja khususnya usia 15-64 tahun, baik di kabupaten maupun kotamadya Propinsi Jawa Timur.

BAB II

TINJAUAN PUSTAKA

Pada tinjauan pustaka ini akan diuraikan mengenai penelitian terdahulu serta metode analisis yang digunakan dalam penelitian ini, yaitu tabulasi silang atau *cross tabulation* untuk mengetahui karakteristik variabel respon terhadap prediktor, kemudian klasifikasi pohon dalam *Classification and Regression Trees* (CART).

2.1 Penelitian Terdahulu

Beberapa kajian yang berkaitan dengan partisipasi perempuan telah dilakukan oleh Rini (2010) tentang faktor-faktor yang mempengaruhi partisipasi perempuan kawin dalam kegiatan ekonomi Jawa Timur yang dianalisis dengan regresi tobit. Pada penelitian ini dibedakan antara pedesaan dan perkotaan di Jawa Timur agar tidak terjadi bias pada hasil yang didapatkan. Hal ini dikarenakan pola hidup dan pola konsumsi masyarakat pedesaan dan perkotaan berbeda. Sementara itu Rantau (2013) juga melakukan analisis faktor yang sama melainkan dengan metode regresi *double hurdle* tanpa dibedakan antara daerah pedesaan dan perkotaan dengan hasil estimasi parameter menunjukkan bahwa keputusan untuk berpartisipasi dalam ekonomi perempuan kawin dipengaruhi oleh umur, pendidikan akhir, jumlah anak balita, dan pendidikan akhir suami.

Sayyida (2012) menganalisis perempuan Jawa Timur tanpa dibatasi dengan status kawin, menggunakan regresi logistik biner dan hasilnya semakin muda umur, tingkat pendidikan semakin tinggi, berasal dari keluarga tidak miskin (bukan penerima BLT), tinggal di kota dan belum menikah, akan meningkatkan kecenderungan perempuan untuk bekerja di sektor non pertanian dengan jam kerja 35 jam ke atas perminggu.

Metode klasifikasi pohon pada CART itu sendiri sebelumnya telah dilakukan oleh Hardiyanti (2011) dalam mengklasifikasikan kesejahteraan rumah tangga perkotaan di Propinsi Jawa Timur dengan ketepatan klasifikasi sebesar 73%.

2.2 Cross Tabulation

Tabulasi silang atau *cross tabulation* merupakan analisis yang dilakukan untuk mengetahui ada atau tidaknya ketergantungan atau hubungan antara dua variabel yang telah ditetapkan. Berikut ini tabel *cross tabulation* rxc

Tabel 2.1 *Cross Tabulation* rxc

Baris	Lajur			
	1	2	...	c
1	n_{11}	n_{12}	...	n_{1c}
2	n_{21}	n_{22}	...	n_{2c}
⋮	⋮	⋮	⋮	⋮
r	n_{r1}	n_{r2}		n_{rc}

Hipotesis

H_0 : tidak ada hubungan antara dua variabel

H_1 : ada hubungan antara dua variabel

Statistik uji yang digunakan adalah

$$\chi_{hit}^2 = \sum_{i,j} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (2.1)$$

Dimana $e_{ij} = \frac{n_{i.} \times n_{.j}}{n_{..}}$

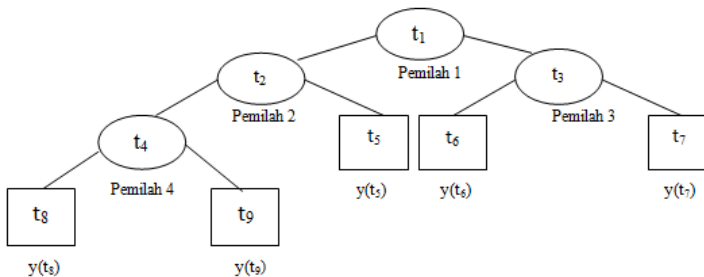
Apabila nilai $\chi_{hit}^2 > \chi_{tabel}^2$ atau $P_{value} < \alpha$ maka keputusan adalah tolak H_0 dengan kata lain terdapat hubungan antara variabel satu dan lainnya.

2.3 Classification and Regression Trees (CART)

CART merupakan salah satu metodologi statistik nonparametrik, sehingga tidak membutuhkan asumsi-asumsi di dalamnya. Metode ini dikembangkan untuk topik analisis klasifikasi, baik untuk variabel respon kategorik maupun kontinu. Jika variabel responnya berbentuk kategorik maka CART yang dihasilkan adalah pohon klasifikasi, namun jika variabel responnya berbentuk kontinu maka yang dihasilkan adalah pohon regresi. Pembentukan pohon klasifikasi ini bertujuan untuk mendapatkan suatu kelompok data yang akurat sebagai penciri

dari suatu pengklasifikasian serta menentukan prediksi struktur data (Breiman, Friedman, Olshen, dan Stone. 1984)

Pembentukan pohon klasifikasi pada dasarnya hampir sama dengan pembentukan pohon regresi, dimana simpul utama (*root node*) terdapat beberapa kelompok data didalamnya dan penentuan simpul utama berdasarkan variabel terpenting, selanjutnya simpul utama dipilah menjadi simpul anak kiri (*left child node*) dan simpul anak kanan (*right child node*), demikian dilakukan seterusnya hingga didapatkan suatu simpul akhir (*terminal node*). Berikut adalah ilustrasi pembentukan pohon klasifikasi.



Gambar 2.1 Pembentukan Pohon Klasifikasi

Notasi t_1 merupakan simpul utama, sedangkan simpul dalam (*internal nodes*) dinotasikan t_2 , t_3 dan t_4 . Simpul akhir atau simpul terminal adalah t_5 , t_6 , t_7 , t_8 , dan t_9 . Penghitungan *depth* atau kedalaman pohon dimulai dari simpul utama t_1 yang berada pada kedalaman 1, sedangkan t_2 dan t_3 berada pada kedalaman 2 begitu seterusnya sampai pada t_8 dan t_9 . Notasi $y(t)$ merupakan hasil prediksi nilai konstan. Berikut merupakan langkah-langkah analisis klasifikasi pohon.

2.3.1 Pembentukan Pohon Klasifikasi Maksimal

Pembentukan pohon klasifikasi ini terdiri dari tiga tahapan, meliputi, pemilihan pemilah, penentuan simpul terminal dan penandaan label kelas.

a. Pemilahan Pemilah

Pembentukan pohon klasifikasi membutuhkan *learning sample* yang terdiri atas N amatan sampel dimana learning sampel tersebut bersifat heterogen sehingga dilakukan pemilahan untuk mengurangi keheterogenan pada simpul utama dan memaksimumkan ukuran keheterogenan pada simpul anak. Dimana metode pemilahan yang dapat digunakan adalah Gini, Twoing, Entropy (Breiman dkk, 1984) dan Informasi (Zhang, 1999). Pemilah yang menghasilkan nilai keheterogenan tertinggi adalah pemilah yang terbaik. Sementara itu, fungsi keheterogenan yang sangat mudah dan sesuai diterapkan dalam berbagai kasus yaitu Indeks Gini (Breiman dkk, 1984) sebagai berikut.

$$i(t) = \sum_{i \neq j} (p(i, t) | p(t)) (p(j, t) | p(t)) \quad (2.2)$$

keterangan :

$p(i, t) / p(t)$ adalah proporsi kelas i simpul t

$p(j, t) / p(t)$ adalah proporsi kelas j simpul t

Menurut Breiman, dkk (1984) pertama, mencari kemungkinan pemilah pada variabel independen, sesuai dengan aturan berikut.

- i. Tiap pemilahan hanya bergantung pada nilai dari satu variabel independen.
- ii. Jika variabel independen (x_j) kontinu, pemilahan yang diijinkan adalah $x_j \leq c_i$ dan $x_j > c_i$, dimana $i = 1, 2, \dots, n-1$ sementara c_i adalah nilai tengah dari dua amatan sampel yang berbeda variabel x_j . Apabila suatu sampel berukuran n dan terdapat N nilai amatan sampel yang berbeda pada variabel x_j , maka terdapat $n-1$ kemungkinan pemilahan.
- iii. Jika variabel independen (x_j) kategori, maka pemilahan berasal dari semua kemungkinan pemilahan menurut terbentuknya 2 simpul yang saling lepas. Jika variabel x_j adalah kategori nominal bertaraf L maka akan diperoleh pemilahan sebanyak $2^{L-1}-1$. Tetapi, jika variabel x_j kategori ordinal maka akan diperoleh pemilahan sebanyak $L-1$.

Kedua, mencari pemilah terbaik dari setiap variabel independen. Pemilah yang terpilih akan membentuk suatu himpunan kelas yang disebut simpul satu, selanjutnya tiap pemilahan akan dilakukan pada tiap simpul sampai diperoleh simpul akhir dan menghasilkan 2 kelas yaitu simpul anak. Apabila terdapat pengamatan hilang (*missing value*) pada pemilah, maka pemilahan akan diganti variabel lain yang disebut pemilah pengganti (*surrogate splitter*). Pohon dikembangkan hingga ditemukan pemilah s^* yang memberikan nilai penurunan keheterogenan tertinggi, sebagai berikut.

$$\Delta i(s^*, t_1) = \max_{s \in S} \Delta i(s, t_1) \quad (2.3)$$

b. Penentuan Simpul Terminal

Simpul t dapat menjadi simpul terminal jika tidak, terdapat penurunan keheterogenan secara berarti atau adanya batasan minimum n hanya terdapat satu pengamatan pada tiap simpul anak.

c. Penandaan Label Kelas

Tahapan ini dilakukan berdasarkan aturan jumlah terbanyak, yaitu

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

Dimana

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

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

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

Proses pembentukan pohon klasifikasi berhenti saat terdapat hanya satu pengamatan dalam tiap-tiap simpul anak atau adanya batasan minimum n , semua pengamatan dalam tiap simpul anak identik, dan adanya batasan jumlah level/kedalaman pohon maksimal. Setelah pohon maksimal terbentuk, selanjutnya adalah pemangkasan pohon.

2.3.2 Penghentian Pembentukan Pohon Klasifikasi

Penghentian pembentukan pohon klasifikasi dilakukan jika tidak memungkinkan lagi untuk dilanjutkan, Lewis dan Roger (2000) menyatakan bahwa penghentian tersebut berhenti pada saat hanya ada satu pengamatan ($n=1$) dalam tiap simpul anak, semua pengamatan dalam tiap simpul anak mempunyai variabel independen yang identik sehingga tidak mungkin terjadi pemilahan, dan adanya batasan jumlah level atau tingkat kedalaman pohon maksimal.

2.3.3 Pemangkasan Pohon Klasifikasi

Pemangkasan ini dilakukan hanya untuk mencegah terbentuknya pohon klasifikasi yang sangat besar dan kompleks, tanpa mengurangi kebaikan ketepatan melalui pengurangan simpul pohon. Pohon dipangkas secara iteratif menjadi sekuen pohon yang semakin kecil dan tersarang. Berdasarkan *cost complexity pruning*, *resubstitution* suatu pohon T pada kompleksitas α adalah

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

Keterangan

$R(T)$ = *resubstitution* estimate pohon T

α = *complexity* parameter bagi penambahan satu simpul akhir pada pohon T

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

Cost complexity pruning menentukan suatu pohon bagian $T(\alpha)$ yang meminimumkan $R_{\alpha}(t)$ pada seluruh pohon bagian.

2.3.4 Penentuan Pohon Klasifikasi Optimal

Penentuan pohon klasifikasi optimal dilakukan untuk memperoleh pohon optimal yang berukuran sederhana tetapi memberikan nilai kesalahan prediksi yang cukup kecil. Terdapat dua jenis penduga pengganti, yaitu penduga sampel uji (*test sample estimate*) dan penduga validasi silang lipat V (*cross validation V-fold estimate*). Breiman, dkk (1984) mengatakan bahwa penggunaan penduga validasi silang lipat V dilakukan untuk menghitung biaya pengganti relatif pada sampel kurang

dari 900, sedangkan penggunaan penduga sampel uji untuk sampel lebih besar dari 900.

2.3.5 Ketepatan Klasifikasi

Ukuran ketepatan klasifikasi dapat diketahui dengan berbagai cara, diantaranya dengan nilai *sensitivity*, *specificity*, *error rate* dan *total accuracy rate*. Dengan perhitungan sebagai berikut.

Tabel 2.2 Hasil Klasifikasi

Actual	Predicted Class		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 observasi 1 yang tepat diprediksi sebagai 1

n_{12} = jumlah observasi 1 yang tepat diprediksi sebagai 2

n_{21} = jumlah observasi 2 yang tepat diprediksi sebagai 1

n_{22} = jumlah observasi 2 yang tepat diprediksi sebagai 2

$N_{1.}$ = jumlah observasi dari 1

$N_{2.}$ = jumlah observasi dari 2

N = jumlah total observasi

Sensitivity adalah ukuran ketepatan dari suatu kejadian yang diinginkan. *Specificity* merupakan suatu ukuran yang menyatakan kejadian yang tidak diinginkan. *Error rate* merupakan proporsi observasi yang diprediksi secara tidak benar oleh fungsi klasifikasi. Kemudian *total accuracy rate* adalah proporsi observasi yang diprediksi secara benar oleh fungsi klasifikasi.

Adapun perhitungannya sebagai berikut.

$$\text{Sensitivity} = n_{11} / N_{1.}$$

$$\text{Specificity} = n_{22} / N_{2.}$$

$$\text{Error Rate} = (n_{21} + n_{12}) / N$$

$$\text{Total Accuracy Rate} = (n_{11} + n_{22}) / N$$

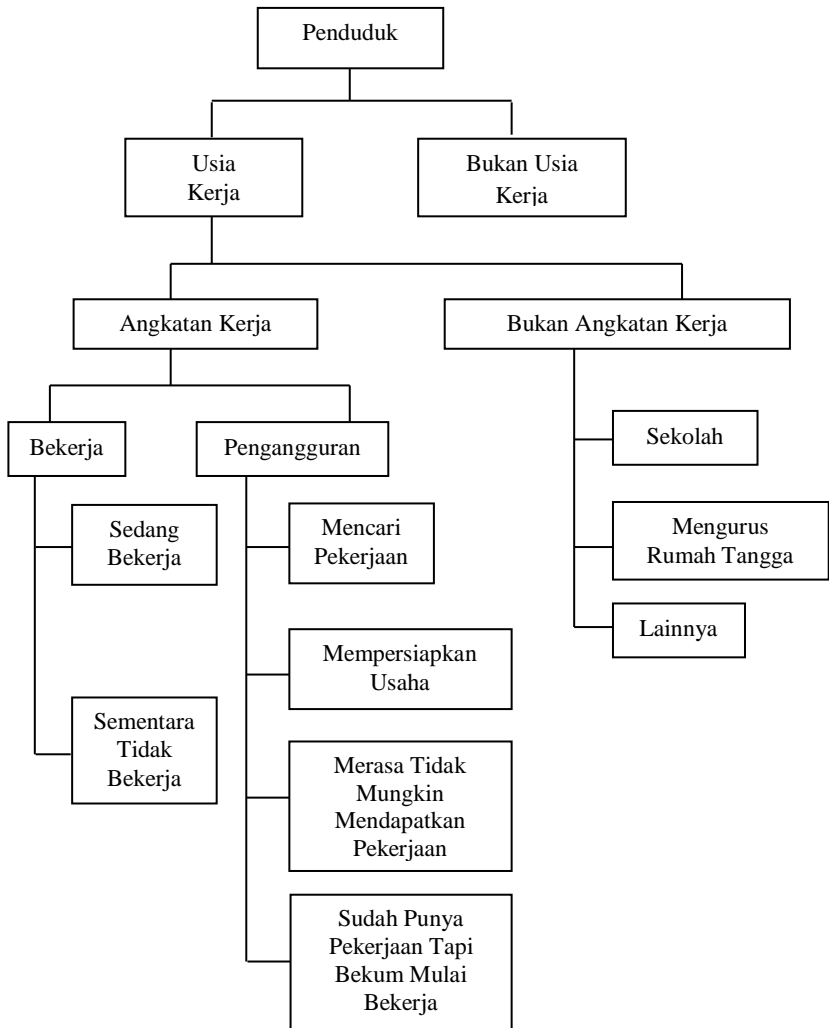
2.4 Kegiatan Partisipasi Ekonomi Perempuan

Todaro dan Smith (2011) dalam bukunya, mengemukakan bahwa perempuan mendominasi para migran dari desa ke kota.

Meningkatnya jumlah migran perempuan yang belum menikah juga menyebabkan meningkatnya proporsi rumah tangga yang dikepalai perempuan di kawasan perkotaan, yang cenderung lebih miskin. Karena anggota rumah tangga yang dikepalai perempuan ini umumnya melakukan pekerjaan di sektor informal yang produktivitasnya rendah dan memikul beban ketergantungan yang lebih tinggi. Tingkat putus sekolah di kalangan anak-anak dari rumah tangga yang dikepalai perempuan ini jauh lebih tinggi karena anak-anak tersebut kemungkinan besar harus bekerja untuk menambah penghasilan keluarga.

Legalisasi dan dorongan ekonomi dalam sektor informal perkotaan, yang sebagian besar tenaga kerjanya adalah perempuan, akan dapat meningkatkan fleksibilitas keuangan dan produktivitas usaha mereka. Agar perempuan dapat meraih manfaat ini, maka pemerintah harus meniadakan peraturan perundang-undangan yang membatasi hak-hak perempuan, selain itu juga harus dihilangkan hambatan bagi perempuan untuk ikut serta dalam program pelatihan dan layanan teknis yang disediakan pemerintah. Terakhir, penyediaan layanan perawatan anak dan keluarga berencana akan meringankan beban reproduksi yang harus dipikul perempuan dan memungkinkan mereka untuk meraih partisipasi yang lebih besar dalam perekonomian.

Bentuk kegiatan partisipasi ekonomi perempuan pada penelitian ini dibedakan menjadi dua kategori yaitu bekerja dan tidak bekerja, dimana bekerja didefinisikan sebagai kegiatan ekonomi yang dilakukan seseorang dengan maksud memperoleh atau membantu memperoleh pendapatan atau keuntungan selama paling sedikit 1 (satu) jam secara tidak terputus selama seminggu yang lalu. Kegiatan bekerja ini mencakup, baik yang sedang bekerja maupun yang punya pekerjaan tetapi dalam seminggu yang lalu sementara tidak bekerja, misalnya karena cuti, sakit dan sejenisnya. Definisi ketenagakerjaan menurut BPS (2011) dirangkum dalam Gambar 2.2.



Gambar 2.2 Konsep Penduduk, Usia Kerja, Angkatan Kerja, Bekerja dan Pengangguran

“Halaman Ini Sengaja Dikosongkan”

BAB III

METODOLOGI PENELITIAN

Pada bab ini akan diuraikan sumber data, variabel penelitian beserta definisi operasionalnya serta metode atau langkah analisis yang dilakukan dalam penelitian.

3.1 Sumber Data

Sumber data yang digunakan dalam penelitian ini merupakan data sekunder dari SUSENAS tahun 2011, dengan unit penelitian individu, perempuan usia kerja, sesuai dengan batasan penelitian yakni 15-64 tahun di Propinsi Jawa Timur, terdiri atas 28.472 jiwa di kabupaten dan 6.784 jiwa di kotamadya.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel respon (variabel dependen) dan variabel prediktor (variabel independen), antara lain sebagai berikut.

Tabel 3.1 Variabel Penelitian

Variabel	Deskripsi	Keterangan
Y	Kegiatan partisipasi ekonomi perempuan	1 : bekerja 2 : tidak bekerja
X_1	Klasifikasi desa/kelurahan	1 : perkotaan 2 : pedesaan
X_2	Umur	1 : 15-19 tahun 2 : 20-54 tahun 3 : 55-64 tahun
X_3	Status perkawinan	1 : belum kawin 2 : kawin 3 : pernah kawin
X_4	Hubungan dengan kepala rumah tangga	1 : kepala rumah tangga 2 : istri 3 : anak 4 : lainnya

Lanjutan Tabel 3.1 Variabel Penelitian

X ₅	Ijazah tertinggi yang dimiliki	1 : tidak punya ijazah 2 : SD-SMA/ sederajat 3 : D1/D2/D3 4 : D4/S1 5 : S2/S3
----------------	--------------------------------	---

Y: kegiatan partisipasi ekonomi perempuan usia kerja kurang dari 65 tahun, dibedakan dengan bekerja atau tidak bekerja. Dimana tidak bekerja meliputi pengangguran, sekolah, mengurus rumah tangga dan lainnya.

X₇: klasifikasi desa/kelurahan, dibedakan desa dan kota.

X₂: umur perempuan/responden usia 15-64 tahun dengan kategori 1-3 dikelompokkan menurut Sleumer (1996) yaitu, kurang produktif penuh, golongan produktif, tidak produktif penuh.

X₃: status perkawinan, kategori 1 yaitu perempuan dengan status belum kawin, kategori 2 perempuan dengan status kawin, dan kategori 3 untuk cerai mati ataupun cerai hidup.

X₄: hubungan dengan kepala rumah tangga, kategori 4 meliputi menantu, cucu, orangtua/mertua, famili lain, pembantu rumah tangga dan lainnya.

X₅: ijazah tertinggi yang dimiliki merupakan jenjang pendidikan terakhir dan mempunyai ijazah pada jenjang tersebut.

3.3 Metode Analisis

Langkah analisis yang akan dilakukan untuk mencapai tujuan penelitian adalah sebagai berikut.

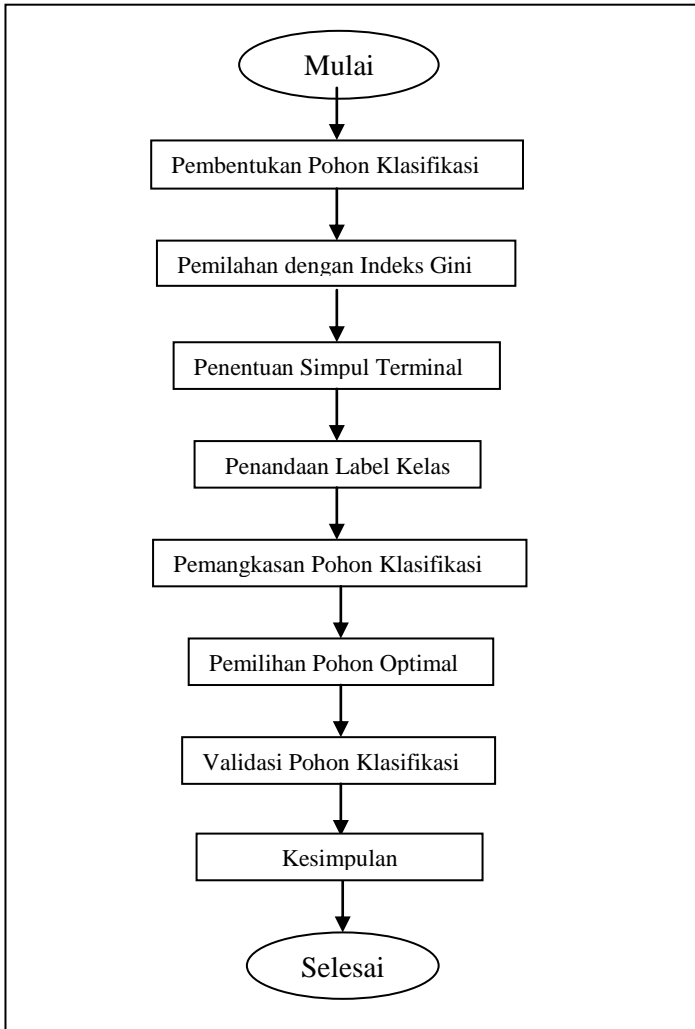
1. Mengidentifikasi karakteristik dengan menggunakan *crosstab*, tabel kontingensi kemudian ditampilkan dalam bentuk chart.
2. Mencari model klasifikasi pohon terhadap partisipasi perempuan di Jawa Timur menurut data SUSENAS 2011 dengan *software* CART 4.0 sesuai tahapan berikut.
 - a. Eksplorasi data yakni pembagian data menjadi dua yaitu data *learning* dan data *testing*, data *learning* digunakan

untuk proses pembentukan model pohon klasifikasi, sedangkan data *testing* digunakan untuk melakukan validasi model yaitu melihat kemampuan model yang terbentuk dalam menduga atau memprediksi data baru. Mencari ketepatan klasifikasi tertinggi dari setiap kombinasi data pada data kabupaten dan kotamadya.

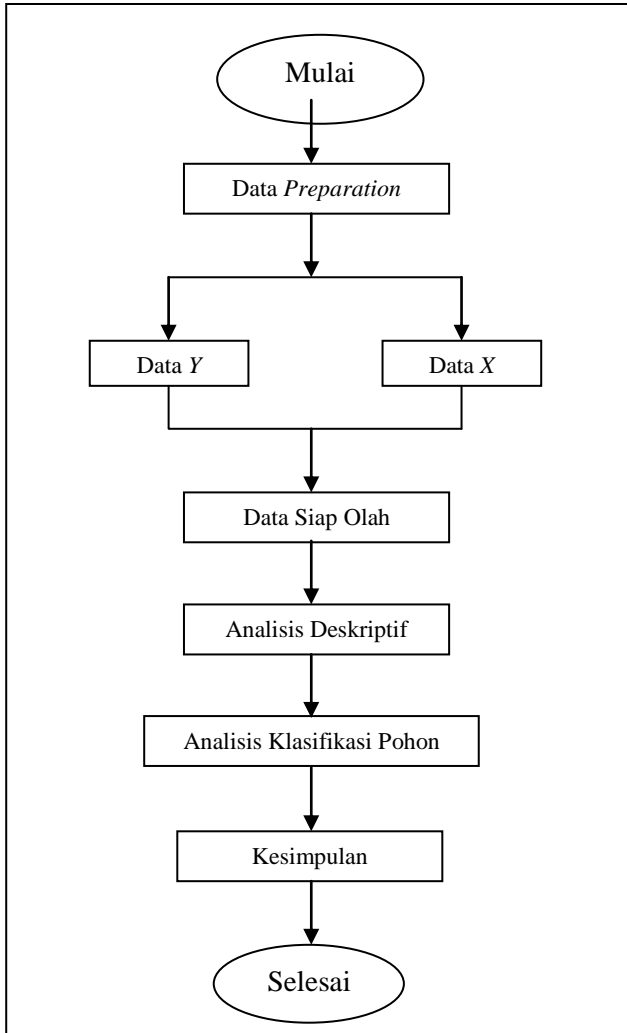
- b. Pembentukan pohon klasifikasi dengan tahapan berikut.
 - Pemilihan pemilah terbaik, dihitung berdasarkan selisih terbesar rataan kuadrat sisaan antara simpul induk dan kedua simpul anak pemisahannya.
 - Penentuan simpul terminal, yang merupakan simpul akhir dari klasifikasi pohon, diperoleh dari ketika tidak memungkinkan lagi melakukan pemilahan pada suatu simpul atau dengan kata lain tidak terdapat lagi penurunan keheterogenan yang berarti.
- c. Penghentian pembentukan pohon klasifikasi.
- d. Pemangkasan pohon klasifikasi dengan menggunakan kriteria kompleksitas kesalahan (*cost complexity*) yang minimum dan nilai kesalahan relatif penggantian (*resubstitution relative error*) yang minimum. Dari hasil proses pemangkasan ini diperoleh deretan pohon (*tree sequence*) dengan ukuran yang semakin mengecil dan tingkat kesalahan prediksi yang semakin membesar.
- e. Pemilihan pohon klasifikasi optimal adalah dengan penduga sampel uji atau kesalahan relatif penduga validasi silang lipat yang minimum.

Validasi model untuk menguji ketepatan klasifikasi dengan memasukkan data *testing* pada pohon klasifikasi optimal.

Langkah algoritma CART dan langkah analisis secara umum dapat digambarkan sebagai berikut.



Gambar 3.1 Algoritma CART



Gambar 3.2 Diagram Alir Penelitian

“Halaman Ini Sengaja Dikosongkan”

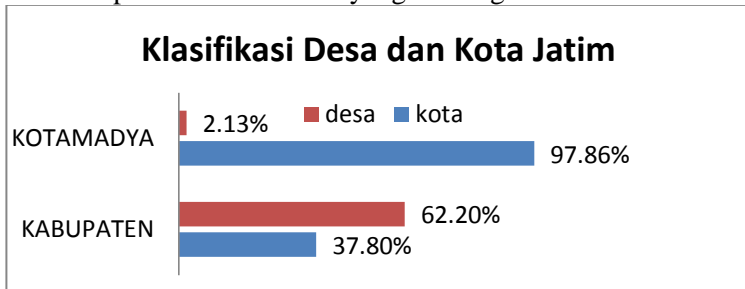
BAB 4 ANALISIS DAN PEMBAHASAN

Pada analisis dan pembahasan akan diuraikan jawaban dari rumusan masalah, yaitu karakteristik serta hasil klasifikasi dari kegiatan partisipasi ekonomi perempuan di Jawa Timur.

4.1 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan di Jawa Timur

4.1.1 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota

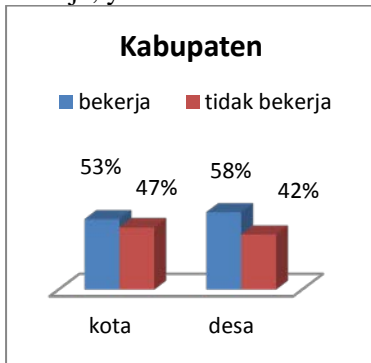
Gambar 4.1 menunjukkan bahwa kabupaten di Jawa Timur didominasi oleh pedesaan sementara di kotamadya mayoritas perkotaan dengan masih terdapat desa sebanyak 2,13%. Daerah itu adalah Kotamadya Probolinggo dan Kotamadya Batu. Hal ini dapat disebabkan oleh banyak hal, misalnya desa-desa tersebut belum memenuhi syarat untuk menjadi sebuah kota atau kelurahan pada daerah-daerah yang bersangkutan.



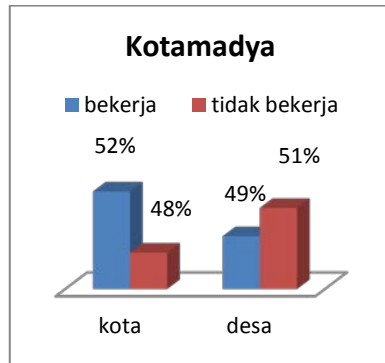
Gambar 4.1 Klasifikasi Desa dan Kota Jawa Timur

Kemudian, dari identifikasi klasifikasi desa dan kota di Jawa Timur, tampak Gambar 4.2 adalah persentase kegiatan partisipasi ekonomi perempuan berdasarkan klasifikasi desa dan kota di masing-masing kabupaten dan kotamadya di Jawa Timur. Hasilnya, persentase bekerja lebih besar pada tiap desa maupun kota, hanya pada desa di kotamadya, memiliki persentase tidak

bekerja yang lebih tinggi dibandingkan dengan persentase bekerja, yaitu sebesar 51%.



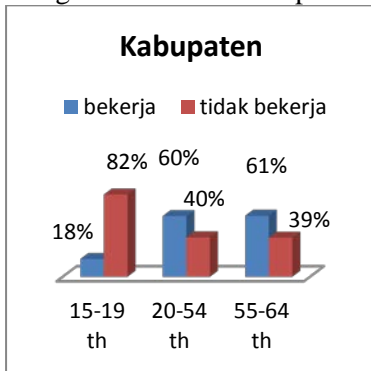
Gambar 4.2 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota di Kabupaten



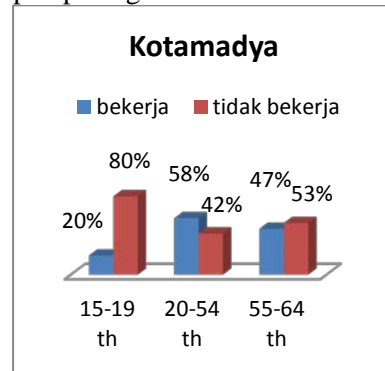
Gambar 4.3 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Klasifikasi Desa dan Kota di Kotamadya

4.1.2 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur

Karakteristik partisipasi ekonomi perempuan di Jawa Timur menghasilkan analisis seperti tampak pada gambar berikut.



Gambar 4.4 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur di Kabupaten

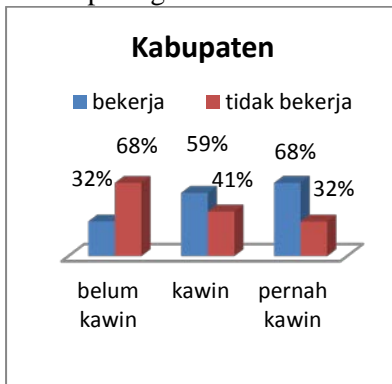


Gambar 4.5 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Kelompok Umur di Kotamadya

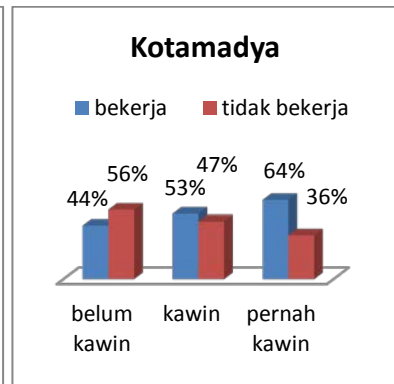
Bahwa pada kelompok umur 15-19 tahun dan 20-54 tahun di kabupaten dan kotamadya tidak terdapat perbedaan yang cukup signifikan. Hanya pada kelompok umur 55-64 tahun di kabupaten, lebih banyak yang bekerja, berbanding terbalik dengan di kotamadya, lebih banyak yang tidak bekerja.

4.1.3 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan

Selanjutnya persentase berdasarkan status perkawinan dapat dilihat pada gambar berikut.



Gambar 4.6 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan di Kabupaten

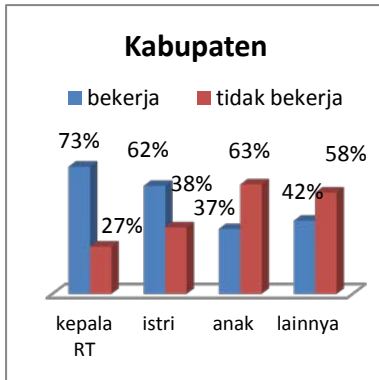


Gambar 4.7 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Status Perkawinan di Kotamadya

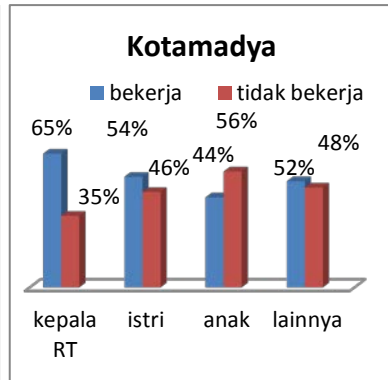
Gambar 4.6 dan 4.7 menunjukkan bahwa perempuan dengan status kawin dan pernah kawin yaitu cerai hidup maupun cerai mati, cenderung memiliki persentase bekerja yang lebih tinggi dibandingkan dengan persentase tidak bekerja.

4.1.4 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga

Pada Gambar 4.8 dan 4.9 akan ditampilkan persentase kegiatan partisipasi ekonomi perempuan berdasarkan hubungan dengan kepala rumah tangga.



Gambar 4.8 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga di Kabupaten

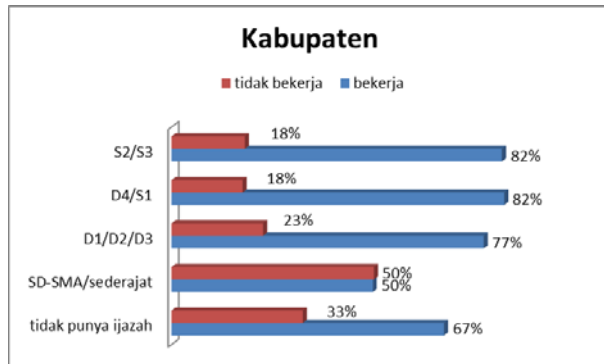


Gambar 4.9 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Hubungan Dengan Kepala Rumah Tangga di Kotamadya

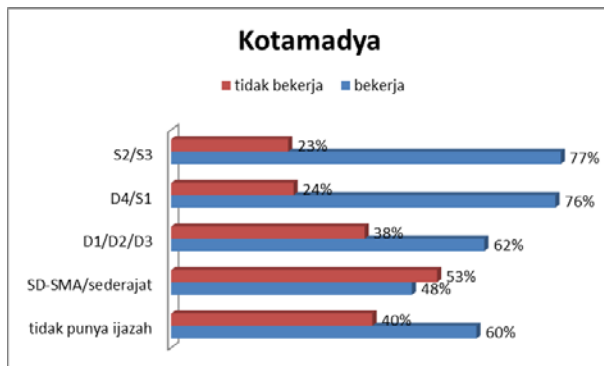
Keduanya memiliki selisih persentase kegiatan partisipasi ekonomi yang sama, hanya pada status lainnya, meliputi menantu, cucu, orangtua/mertua, famili lain, pembantu rumah tangga dan lainnya di kotamadya persentase bekerja dan tidak bekerja yang hampir sama yakni 52% dan 48%.

4.1.5 Karakteristik Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi

Persentase partisipasi perempuan berdasarkan ijazah tertinggi di kabupaten ditunjukkan pada Gambar 4.10, dimana kegiatan partisipasi ekonomi kategori bekerja didominasi oleh perempuan dengan ijazah D4/S1 dan S2/S3. Ijazah tertinggi untuk kegiatan bekerja di kabupaten tersebut memiliki persentase yang sama, yakni 82% untuk ijazah D4/S1 dan 82% pula untuk ijazah S2/S3. Kemudian diikuti oleh ijazah D1/D2/D3 sebesar 77%.



Gambar 4.10 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi di Kabupaten
Begitu juga dengan di kotamadya, yang ditunjukkan pada Gambar 4.11.



Gambar 4.11 Persentase Kegiatan Partisipasi Ekonomi Perempuan Berdasarkan Ijazah Tertinggi di Kotamadya
Kegiatan partisipasi ekonomi kategori bekerja di kotamadya didominasi oleh perempuan dengan ijazah D4/S1 dan S2/S3.

Analisis karakteristik antar variabel dalam penelitian ini, dilakukan dengan cara tabulasi silang antara variabel respon dan variabel prediktor yang telah ditampilkan dalam bentuk Gambar 4.2 sampai dengan 4.11. Selanjutnya untuk mengetahui hubungan

variabel respon dan prediktor dilakukan pengujian dengan hipotesis sebagai berikut.

H_0 : Tidak ada hubungan antara variabel respon dengan prediktor.

H_1 : Ada hubungan antara variabel respon dengan prediktor.

Tabel 4.1 Hasil *Crosstab* Variabel Respon Dengan Prediktor

df	Gambar No.	Kabupaten		Gambar No.	Kotamadya	
		χ_{hit}^2	χ_{tabel}^2		χ_{hit}^2	χ_{tabel}^2
1	4.2	75,765	3,841	4.3	0,620	3,841
2	4.4	1890	5,991	4.5	411	5,991
2	4.6	1115	5,991	4.7	86,228	5,991
3	4.8	1601	7,815	4.9	92,615	7,815
4	4.10	1050	0,711	4.11	2,22	0,711

Dengan nilai α sebesar 0,05 maka telah diperoleh nilai χ_{tabel}^2 seperti pada Tabel 4.1, dengan keputusan tolak H_0 jika $\chi_{hit}^2 > \chi_{tabel}^2$, sehingga kesimpulannya adalah seluruh variabel tolak H_0 kecuali Gambar 4.3, yakni variabel klasifikasi desa dan kota, gagal tolak H_0 yang artinya tidak terdapat hubungan antara variabel tersebut dengan kegiatan partisipasi ekonomi perempuan di kotamadya propinsi Jawa Timur. Oleh karena itu, variabel tersebut tidak diikutsertakan dalam analisis selanjutnya.

4.2 Classification And Regression Trees (CART)

Berdasarkan hasil penguraian tentang karakteristik kabupaten dan kotamadya di Jawa Timur, dimana terdapat perbedaan di dalamnya, maka untuk selanjutnya klasifikasi akan dilakukan dengan memisahkan antara kabupaten dan kotamadya.

Pada penerapan metode CART untuk klasifikasi penelitian ini, terlebih dahulu dilakukan pembagian data *learning* dan *testing*, dimana tidak ada aturan khusus dalam pembagiannya, untuk itu berikut persentase ketepatan klasifikasi pada berbagai kombinasi yang telah dilakukan oleh peneliti.

Tabel 4.2 Perbandingan Ketepatan Klasifikasi Antar Kombinasi Data

KABUPATEN			KOTAMADYA		
Kombinasi Proporsi (%)	Ketepatan Klasifikasi (%)		Kombinasi Proporsi (%)	Ketepatan Klasifikasi (%)	
	<i>Learning</i>	<i>Testing</i>		<i>Learning</i>	<i>Testing</i>
95 : 5	64,4	64,3	95 : 5	60	59,5
*90 : 10	64,4	64,6	90 : 10	60,1	59,5
85 : 15	64,5	64,4	*85 : 15	60	60,1
80 : 20	64,5	64,1	80 : 20	60,3	59
75 : 25	64,4	64,5	75 : 25	61,2	59,1

(*) Keterangan : kombinasi data yang digunakan untuk analisis

Tabel 4.2 menunjukkan bahwa yang digunakan untuk analisis data kabupaten adalah kombinasi 90% *learning* dan 10% *testing*. Sementara untuk data kotamadya yang digunakan adalah kombinasi 85% *learning* dan 15% *testing*. Ketepatan klasifikasi ini dijadikan dasar karena dapat menggambarkan kebaikan pohon klasifikasi yang terbentuk.

Berikut adalah perhitungan kemungkinan pemilah dari masing-masing variabel prediktor.

- Klasifikasi desa/kota memiliki $2^{2-1} - 1 = 1$ pemilahan
- Umur memiliki $3 - 1 = 2$ pemilahan
- Status perkawinan memiliki $2^{3-1} - 1 = 3$ pemilahan
- Hubungan dengan kepala RT memiliki $2^{4-1} - 1 = 7$ pemilahan
- Ijazah tertinggi memiliki $5 - 1 = 4$ pemilahan

4.2.1 Klasifikasi Pohon di Kabupaten Propinsi Jawa Timur

1. Pembentukan Pohon Klasifikasi Maksimal

Pohon klasifikasi maksimal merupakan pohon dengan jumlah simpul terminal terbanyak, hasil pemilahan simpul utama maupun simpul pemilah dengan indeks gini. Tujuan pemilahan adalah mengurangi keheterogenan pada simpul utama dan memaksimumkan kehomogenan simpul anak.

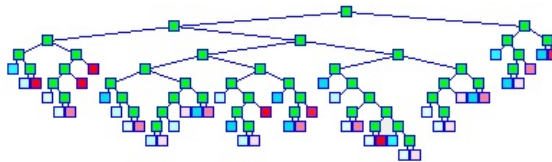
Pemilah terbaik adalah pemilah yang menghasilkan penurunan heterogenitas tertinggi. Pada tabel 4.3 diketahui skor variabel terpenting dalam klasifikasi pohon maksimal.

Tabel 4.3 Skor Variabel Prediktor Pada Pohon Maksimal di Kabupaten

Variabel	Skor
Umur	100,00
Status perkawinan	73,97
Hubungan dengan kepala rumah tangga	38,48
Ijazah tertinggi	36,48
Klasifikasi desa/kota	7,89

Berdasarkan tabel tersebut dapat diketahui bahwa umur memiliki skor tertinggi diikuti dengan status perkawinan sebesar 100 dan 73,97.

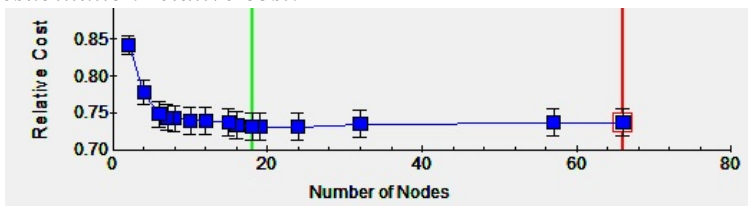
Bentuk pohon maksimal yang dihasilkan sebagai berikut.



Gambar 4.12 Pohon Klasifikasi Maksimal di Kabupaten

2. Pemangkasan Pohon Klasifikasi

Pemangkasan ini dilakukan untuk mencegah terbentuknya pohon klasifikasi yang sangat besar dan kompleks, tanpa mengurangi kebaikan ketepatan melalui pengurangan simpul pohon. Pemangkasan dilakukan dengan menggunakan *restitution relative cost*.



Gambar 4.13 Plot *Relative Cost* dengan Jumlah Node di Kabupaten

Berdasarkan plot Gambar 4.13 diketahui bahwa jumlah node pada pohon optimal (garis warna hijau) jauh lebih sedikit dibandingkan dengan pohon maksimal (garis warna merah). Nilai *resubstitution relative cost* di kabupaten pada pohon maksimal sebesar 0,729 dan pohon optimal sebesar 0,731. Selengkapnya di Tabel 4.4.

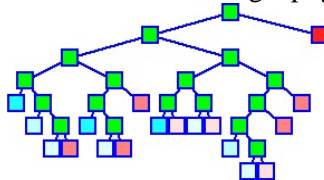
Tabel 4.4 *Tree Sequence* Pohon di Kabupaten

Dependent variable: Y				
Terminal Tree Nodes	Test Set Relative Cost	Resubstitution Relative Cost	Complexity Parameter	
1	66	0.737 +/- 0.018	0.729	0.000
6**	18	0.731 +/- 0.018	0.731	.693545E-04
7	16	0.733 +/- 0.018	0.732	.355258E-03
8	15	0.737 +/- 0.018	0.733	.548752E-03
9	12	0.739 +/- 0.018	0.737	.571520E-03
10	10	0.739 +/- 0.018	0.741	0.001
11	8	0.742 +/- 0.018	0.746	0.001
12	7	0.744 +/- 0.018	0.750	0.002
13	6	0.748 +/- 0.018	0.757	0.004
14	4	0.777 +/- 0.016	0.783	0.007
15	2	0.842 +/- 0.012	0.845	0.015
16	1	1.000 +/- 0.000	1.000	0.078

Initial misclassification cost = 0.500
Initial class assignment = 1

3. Penentuan Pohon Klasifikasi Optimal

Pohon klasifikasi optimal terbentuk melalui pemangkasan yang telah dilakukan sebelumnya. Terdapat 18 simpul terminal dengan 8 kedalaman. Pohon detail selengkapnya di Lampiran A4.



Gambar 4.14 Pohon Klasifikasi Optimal di Kabupaten

Perbedaan label kelas ditunjukkan oleh perbedaan warna pada setiap pohon klasifikasi. Simpul induk dan simpul anak sebagai pemilah ditunjukkan oleh kotak hijau, selanjutnya akan terbentuk simpul terminal berbeda warna kelas/kategori sesuai dengan skor variabel yang dibentuk oleh pohon optimal.

Tabel 4.5 Skor Variabel Prediktor Pada Pohon Optimal di Kabupaten

Variabel	Skor
Umur	100,00
Status perkawinan	72,88
Ijazah tertinggi	36,83
Hubungan dengan kepala rumah tangga	35,54
Klasifikasi desa/kota	5,03

Berdasarkan nilai skor variabel tersebut yang menjadi pemilah utama adalah variabel umur, kemudian tertinggi kedua adalah status perkawinan.

Berikut adalah karakteristik simpul terminal yang terbentuk pada pohon klasifikasi optimal di kabupaten.

Tabel 4.6 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kabupaten

Kategori Y = 1 (bekerja)			
No	Simpul ke-	N	Variabel Prediktor Pada Simpul Terminal
1	1	1079	<ul style="list-style-type: none"> • Ijazah D1/D2/D3;D4/S1 dan S2/S3 • Hubungan KRT; istri; dan anak • Umur 20-54 dan 55-64 tahun
2	2	6078	<ul style="list-style-type: none"> • Hubungan KRT; istri • Tidak punya ijazah • Umur 20-54 dan 55-64 tahun
3	3	210	<ul style="list-style-type: none"> • Status kawin dan lainnya • Hubungan anak • Tidak punya ijazah • Umur 20-54 dan 55-64 tahun
4	5	91	<ul style="list-style-type: none"> • Ijazah D1/D2/D3;D4/S1 • Umur 20-54 tahun • Hubungan lainnya
5	6	233	<ul style="list-style-type: none"> • Status kawin dan lainnya • Tidak punya ijazah dan ijazah S2/S3 • Umur 20-54 tahun • Hubungan lainnya

Lanjutan Tabel 4.6 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kabupaten

Kategori Y = 1 (bekerja)			
No	Simpul ke-	N	Keterangan Simpul Terminal
6	9	500	<ul style="list-style-type: none"> • Umur 20-54 • Status belum kawin dan lainnya • Hubungan KRT dan istri • Ijazah SD-SMA/ sederajat
7	11	6292	<ul style="list-style-type: none"> • Tinggal di pedesaan • Status kawin • Hubungan KRT dan istri • Ijazah SD-SMA/ sederajat • Umur 20-54 dan 55-64 tahun
8	13	282	<ul style="list-style-type: none"> • Status lainnya • Umur 20-54 tahun • Hubungan anak dan lainnya • Ijazah SD-SMA/ sederajat
9	14	480	<ul style="list-style-type: none"> • Tinggal di perkotaan • Status belum kawin • Umur 20-54 tahun • Hubungan anak dan lainnya • Ijazah SD-SMA/ lainnya
Kategori Y = 2 (tidak bekerja)			
No	Simpul ke-	N	Keterangan Simpul Terminal
1	4	96	<ul style="list-style-type: none"> • Status belum kawin • Hubungan anak • Tidak punya ijazah • Umur 20-54 dan 55-64 tahun
2	7	51	<ul style="list-style-type: none"> • Status belum kawin • Tidak punya ijazah dan S2/S3 • Umur 20-54 tahun • Hubungan lainnya

Lanjutan Tabel 4.6 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kabupaten

Kategori Y = 2 (tidak bekerja)			
No	Simpul ke-	N	Keterangan Simpul Terminal
3	8	430	<ul style="list-style-type: none"> • Umur 55-64 tahun • Hubungan lainnya • Tidak punya ijazah; D1/D2/D3; D4/S1; S2/S3
4	10	208	<ul style="list-style-type: none"> • Umur 55-64 tahun • Status belum kawin dan lainnya • Hubungan KRT dan istri • Ijazah SD-SMA/ sederajat
5	12	4208	<ul style="list-style-type: none"> • Tinggal di perkotaan • Status kawin • Hubungan KRT dan istri • Ijazah SD-SMA/ sederajat • Umur 20-54 dan 55-64 tahun
6	15	378	<ul style="list-style-type: none"> • Tinggal di pedesaan • Status belum kawin • Umur 20-54 tahun • Hubungan anak dan lainnya • Ijazah SD-SMA/ sederajat
7	16	98	<ul style="list-style-type: none"> • Umur 20-54 tahun • Status belum kawin dan lainnya • Hubungan anak dan lainnya • Ijazah SD-SMA/ sederajat
8	17	2375	<ul style="list-style-type: none"> • Status kawin • Hubungan anak dan lainnya • Ijazah SD-SMA/ sederajat • Umur 20-54 dan 55-64 tahun
9	18	2536	<ul style="list-style-type: none"> • Umur 15-19 tahun

Berdasarkan Tabel 4.6 diketahui bahwa variabel Y kategori 1 menghasilkan 9 simpul terminal dan variabel Y kategori 2 menghasilkan 9 simpul terminal. Jumlah N terbanyak pada

kategori 1, terletak pada simpul terminal ke-11. Sebanyak 6292 perempuan di kabupaten propinsi Jawa Timur yang bertempat tinggal di pedesaan, berstatus kawin, memiliki hubungan dengan kepala rumah tangga yaitu kepala rumah tangga itu sendiri dan istri, ijazah tertinggi yang dimiliki adalah SD-SMA/ sederajat serta dalam kategori umur 20-54 dan 55-64, memutuskan untuk bekerja atau berpartisipasi dalam kegiatan ekonomi.

Sementara itu pada kategori 2, jumlah N terbanyak terletak di simpul terminal ke-12. Sebanyak 4208 perempuan di kabupaten propinsi Jawa Timur yang bertempat tinggal di perkotaan dengan status kawin, hubungan dengan kepala rumah tangga adalah kepala rumah tangga dan istri, ijazah tertinggi SD-SMA/ sederajat dan dalam kategori umur 20-54 dan 55-64, memutuskan untuk tidak bekerja.

4.2.2 Klasifikasi Pohon di Kotamadya Propinsi Jawa Timur

1. Pembentukan Pohon Klasifikasi Maksimal

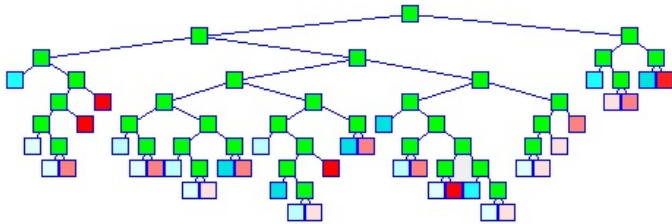
Seperti yang telah dijelaskan sebelumnya, bahwa pemilah terbaik adalah pemilah yang menghasilkan penurunan keheterogenan tertinggi. Pada tabel 4.7 diketahui skor variabel terpenting dalam klasifikasi pohon maksimal.

Tabel 4.7 Skor Variabel Prediktor Pada Pohon Maksimal di Kotamadya

Variabel	Skor
Umur	100,00
Ijazah tertinggi	37,42
Status perkawinan	33,74
Hubungan dengan kepala rumah tangga	18,92

Sama halnya dengan di kabupaten, skor variabel prediktor tertinggi pada pohon maksimal di kotamadya adalah umur. Berbeda dengan kabupaten, skor tertinggi selanjutnya adalah ijazah tertinggi kemudian status perkawinan.

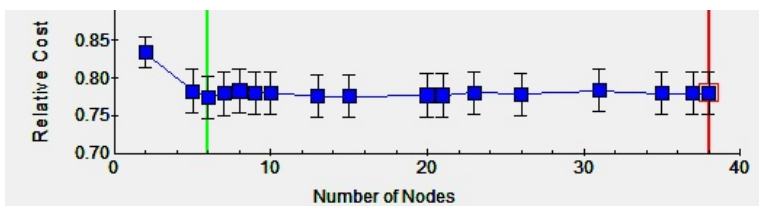
Bentuk pohon maksimal yang dihasilkan sebagai berikut.



Gambar 4.15 Pohon Klasifikasi Maksimal di Kotamadya

2. Pemangkasan Pohon Klasifikasi

Dengan langkah yang sama seperti yang sebelumnya telah dilakukan pada data kabupaten, pemangkasan pohon klasifikasi dilakukan dengan menggunakan *resubstitution relative cost*.



Gambar 4.16 Plot *Relative Cost* dengan Jumlah Node di Kotamadya

Berdasarkan Gambar 4.16, jumlah node pada pohon optimal (garis warna hijau) jauh lebih sedikit dibandingkan dengan pohon maksimal (garis warna merah). Nilai *resubstitution relative cost* di kabupaten pada pohon maksimal sebesar 0,769 dan pohon optimal sebesar 0,788. Selengkapnya dapat dilihat pada Tabel 4.10.

Tabel 4.8 *Tree Sequence* Pohon di Kotamadya

Dependent variable: Y

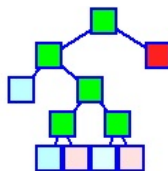
Terminal Tree Nodes	Test Set Relative Cost	Resubstitution Relative Cost	Complexity Parameter
1	38	0.779 +/- 0.028	0.769
9	15	0.775 +/- 0.028	0.774
10	13	0.775 +/- 0.028	0.775
11	10	0.780 +/- 0.028	0.778
12	9	0.780 +/- 0.028	0.779
13	8	0.782 +/- 0.029	0.781
14	7	0.779 +/- 0.029	0.783
15**	6	0.774 +/- 0.028	0.788
16	5	0.782 +/- 0.029	0.793
17	2	0.834 +/- 0.021	0.851
18	1	1.000 +/- 0.000	1.000

Initial misclassification cost = 0.500

Initial class assignment = 1

3. Penentuan Pohon Klasifikasi Optimal

Pohon klasifikasi optimal terbentuk melalui pemangkasan yang telah dilakukan sebelumnya. Terdapat 6 simpul terminal dengan 5 kedalaman.

**Gambar 4.17** Pohon Klasifikasi Optimal di Kotamadya

Gambar detail pohon optimal dapat dilihat di Lampiran B4. Kemudian berikut ini adalah skor variabel prediktor yang terbentuk pada pohon optimal.

Tabel 4.9 Skor Variabel Prediktor Pada Pohon Optimal di Kotamadya

Variabel	Skor
Umur	100,00
Ijazah Tertinggi	33,52
Status perkawinan	30,35
Hubungan dengan kepala rumah tangga	6,53

Berdasarkan nilai skor variabel tersebut, umur tetap menjadi variabel dengan skor tertinggi.

Berikut adalah karakteristik simpul terminal yang terbentuk pada pohon klasifikasi optimal di kotamadya.

Tabel 4.10 Karakteristik Simpul Terminal Pada Pohon Klasifikasi Optimal di Kotamadya

Kategori Y = 1 (bekerja)			
No	Simpul ke-	N	Variabel Prediktor Pada Simpul Terminal
1	1	559	<ul style="list-style-type: none"> • Ijazah D4/S1 dan S2/S3 • Umur 20-54 dan 55-64 tahun
2	2	840	<ul style="list-style-type: none"> • Umur 20-54 tahun • Status belum kawin dan lainnya • Tidak punya ijazah; SD-SMA/ sederajat; dan D1/D2/D3
3	4	575	<ul style="list-style-type: none"> • Tidak punya ijazah; SD-SMA/ sederajat • Status kawin • Umur 20-54 dan 55-64 tahun
Kategori Y = 2 (tidak bekerja)			
No	Simpul ke-	N	Keterangan Simpul Terminal
1	3	243	<ul style="list-style-type: none"> • Umur 55-64 tahun • Status belum kawin dan lainnya • Tidak punya ijazah; SD-SMA/ sederajat; dan D1/D2/D3
2	5	2873	<ul style="list-style-type: none"> • Ijazah SD-SMA/ sederajat • Status kawin • Umur 20-54 dan 55-64 tahun
3	6	665	<ul style="list-style-type: none"> • Umur 15-19 tahun

Pada kotamadya di Jawa Timur, terdapat 3 simpul terminal untuk variabel Y kategori 1 dan 3 simpul terminal untuk variabel Y kategori 2. Jumlah N terbanyak pada kategori 1 terletak pada simpul kedua, sebanyak 840 perempuan umur 20-54 tahun dengan status belum kawin dan lainnya dengan ijazah tertinggi SD-SMA/ sederajat, D1/D2/D3 dan tidak punya ijazah,

memutuskan untuk berpartisipasi dalam kegiatan ekonomi dengan bekerja.

Kemudian, jumlah N terbanyak pada kategori 2 terdapat pada simpul terminal ke-5, sebanyak 2873 perempuan dengan ijazah SD-SMA/ sederajat, berstatus kawin serta berumur 20-54 dan 55-64 tahun, tidak bekerja.

4.2.3 Ketepatan Klasifikasi

Langkah selanjutnya adalah validasi pohon klasifikasi optimal yang telah terbentuk dengan menggunakan data testing, pada data kabupaten maupun kotamadya di Jawa Timur. Berikut ini adalah tabel ketepatan klasifikasinya.

Tabel 4.11 Ketepatan Klasifikasi Data Testing Pohon Optimal di Kabupaten

Kelas Aktual	Kelas Prediksi		Ketepatan Klasifikasi
	Bekerja	Tidak Bekerja	
Bekerja	1159	429	73,0%
Tidak Bekerja	580	679	53,9%
Total Ketepatan Klasifikasi			64,6%

Tabel 4.12 Ketepatan Klasifikasi Data Testing Pohon Optimal di Kotamadya

Kelas Aktual	Kelas Prediksi		Ketepatan Klasifikasi
	Bekerja	Tidak Bekerja	
Bekerja	244	308	44,2%
Tidak Bekerja	103	374	78,4%
Total Ketepatan Klasifikasi			60,1%

Ketepatan klasifikasi total pada data *testing* di Kabupaten sebesar 64,6% sedangkan di kotamadya, total ketepatan klasifikasi sebesar 60,1%, artinya model klasifikasi pohon optimal yang terbentuk sudah cukup baik dalam menggambarkan model klasifikasi pohon maksimal.

“Halaman Ini Sengaja Dikosongkan”

BAB V

KESIMPULAN DAN SARAN

Berikut ini merupakan kesimpulan dari hasil analisis dan pembahasan, sekaligus saran dari penulis.

5.1 Kesimpulan

1. Secara keseluruhan gambaran kegiatan partisipasi ekonomi perempuan di kabupaten maupun di kotamadya menunjukkan persentase bekerja yang lebih tinggi daripada tidak bekerja, dengan selisih yang tidak terpaut jauh. Berdasarkan hasil penguraian karakteristik diperoleh perbedaan antara perempuan di kabupaten dan kotamadya pada variabel klasifikasi desa dan kota, kelompok umur, juga hubungan dengan kepala rumah tangga.

Hasil pengujian *chi-square* tabulasi silang menunjukkan bahwa variabel klasifikasi desa dan kota di kotamadya tidak terdapat hubungan dengan variabel respon, hal ini dikarenakan kotamadya mayoritas perkotaan dengan masih terdapat beberapa desa. Sehingga tidak mampu menggambarkan pola hubungan terhadap variabel respon.

2. Kombinasi data *learning* dan *testing* yang digunakan pada data kabupaten adalah 90%;10% sementara untuk kotamadya 85%;15%. Pohon optimal kabupaten menghasilkan 18 simpul terminal dan 8 kedalaman dengan umur sebagai skor tertinggi. Sebanyak 6292 perempuan di kabupaten propinsi Jawa Timur yang bertempat tinggal di pedesaan, berstatus kawin, memiliki hubungan dengan kepala rumah tangga yaitu kepala rumah tangga itu sendiri dan istri, ijazah tertinggi yang dimiliki adalah SD-SMA/ sederajat serta dalam kategori umur 20-54 dan 55-64, memutuskan untuk bekerja atau berpartisipasi dalam kegiatan ekonomi. Pada kategori 2, sebanyak 4208 perempuan di kabupaten propinsi Jawa Timur yang bertempat tinggal di perkotaan dengan status kawin, hubungan dengan kepala rumah tangga adalah kepala rumah tangga dan istri, ijazah tertinggi SD-SMA/ sederajat dan dalam kategori umur 20-54 dan 55-64,

memutuskan untuk tidak bekerja. Total ketepatan klasifikasi data *testing* sebesar 64,6%.

Pohon optimal di kotamadya menghasilkan 6 simpul terminal dan 5 kedalaman dengan variabel umur memiliki skor tertinggi. sebanyak 840 perempuan umur 20-54 tahun dengan status belum kawin dan lainnya dengan ijazah tertinggi SD-SMA/ sederajat, D1/D2/D3 dan tidak punya ijazah, memutuskan untuk berpartisipasi dalam kegiatan ekonomi dengan bekerja.

Kemudian, sebanyak 2873 perempuan dengan ijazah SD-SMA/ sederajat, berstatus kawin serta berumur 20-54 dan 55-64 tahun, tidak bekerja. Total ketepatan klasifikasi data *testing* sebesar 60,1%.

5.2 Saran

Penelitian selanjutnya disarankan untuk menggunakan data dengan memisahkan antara pedesaan dan perkotaan, dengan variabel independen yang signifikan dalam pembentukan model untuk mendapatkan hasil keakuratan prediksi atau ketepatan klasifikasi yang lebih baik. Selain itu, mengingat kelemahan metode CART yang tidak terdapat model didalamnya, maka perlu metode lain untuk mengetahui pengaruh antara variabel respon dan prediktor.

DAFTAR PUSTAKA

- Balitbang HAM. 2013. www.balitbangham.go.id diakses 27 September 2013
- Biro Statistik. 2012. *Jumlah Penduduk Masing-Masing Negara*.http://statistik.ptkpt.net/a.php?a=penduduk_ratio&info1=3 diakses 27 Agustus 2013.
- BPS. 2011. *Buku Konsep dan Definisi*. www.bps.go.id diakses 1 Oktober 2013
- BPS Jatim. 2012. *Keadaan Angkatan kerja di Provinsi Jawa Timur*. Surabaya : Badan Pusat Statistik Provinsi Jawa Timur
- Breiman, L., Friedman, J., Olshen, R. and Stone, C. 1984. *”Classification and Regression Trees”* Chapman Hall, New York – London.
- Evans, M.D.R., and Kelley, J. 2007, *Trends in Women’s Labor Force Participation in Australia: 1984 – 2002*, Social Science Research.
- Hardiyanti, E. 2011. Skripsi. *Model Kesejahteraan Rumah Tangga Perkotaan Dengan Menggunakan Klasifikasi Pohon*. Surabaya: ITS
- Lewis, M. D. dan Roger, J. 2000. An Introduction to Classification and Regression Tree (CART) Analysis. *Presented at the 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Fransisco, California*.<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf>
- Rantau, D.C.M. 2013. Skripsi. *Analisis Regresi Double Hurdle Terhadap Faktor-Faktor yang Mempengaruhi Partisipasi Perempuan Kawin Dalam Kegiatan Ekonomi di Jawa Timur*. Surabaya: ITS
- Rini, M.P. 2010. Skripsi. *Analisis Regresi Tobit Pada Faktor-Faktor yang Mempengaruhi Partisipasi Perempuan Kawin Dalam Kegiatan Ekonomi di Jawa Timur*. Surabaya: ITS

- Sayyida. 2012. Tesis. *Analisis Partisipasi Ekonomi Perempuan Dengan Metode Regresi Logistik Biner Bivariat di Propinsi Jawa Timur*. Surabaya: ITS
- Sleumer D. W. 1996. Ilmu Dasar Kependudukan. Jakarta: Balai Pustaka
- Todaro, M. P dan Smith, S. C. 2011. *Pembangunan Ekonomi Edisi Kesebelas Jilid 1*. Jakarta: Erlangga
- Walpole, E. R. 1995. *Pengantar Statistika-Edisi ke-3*. Jakarta: PT.Gramedia Pustaka Utama.
- Zhang, H. and Singer, B. 1999. *Recursive Partitioning in the Health Sciences*. Springer-Verlag, NewYork.
<http://c2s2.yale.edu/resources/lectures/LectureNotesOnRP.pdf>

BIODATA PENULIS



Penulis dilahirkan sebagai anak pertama dari tiga bersaudara di Pasuruan pada tanggal 15 Februari 1990 dengan nama lengkap SHARFINA WIDYANDINI. Penulis mulai menyelesaikan pendidikan formal di SDN Pekuncen I Pasuruan, SMP Negeri 2 Pasuruan dan SMA Negeri 1 Pasuruan. Kemudian penulis mengikuti ujian masuk diploma ITS-Surabaya dan diterima di Jurusan D3 Statistika Fakultas MIPA dengan NRP 1308 030 049 pada tahun 2008.

Setelah selesai menempuh pendidikan D3, penulis melanjutkan program studi ke S1 Lintas Jalur Statistika FMIPA ITS dengan NRP 1311 106 002. Penulis yang akrab dipanggil Fina ini memiliki hobi mencoba hal-hal baru di dunia hiburan seperti kuliner, fashion, fotografi, dll disamping kegiatannya dalam menuntut ilmu setinggi-tingginya. Karena baginya, keseimbangan dalam hidup itu menjadi prioritas utama. Apabila ingin berdiskusi mengenai Tugas Akhir ini, dapat menghubungi alamat email sharfinawidyandini@yahoo.com.

DAFTAR LAMPIRAN

Lampiran Judul	Halaman
A Kabupaten.....	43
A.1 Tabulasi Silang Variabel Respon dan Prediktor	43
A.2 Ketepatan Klasifikasi <i>Learning-Testing</i>	46
A.3 <i>Node Information</i> Pohon Optimal	49
A.4 <i>Tree Detail</i> Pohon Optimal	57
B Kotamadya.....	59
B.1 Tabulasi Silang Variabel Respon dan Prediktor	59
B.2 Ketepatan Klasifikasi <i>Learning-Testing</i>	62
B.3 <i>Node Information</i> Pohon Optimal	65
B.4 <i>Tree Detail</i> Pohon Optimal.....	67

LAMPIRAN

A. KABUPATEN

A.1 Tabulasi Silang Variabel Respon dan Prediktor

- Y dengan X_1

		klasifikasi		Total
		kota	desa	
Y	bekerja	5699	10313	16012
	tidak bekerja	5063	7397	12460
Total		10762	17710	28472

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	75.765 ^a	1	.000
Continuity Correction ^b	75.551	1	.000
Likelihood Ratio	75.645	1	.000
Linear-by-Linear Association	75.763	1	.000
N of Valid Cases ^b	28472		

- Y dengan X_2

		umur			Total
		15-19 th	20-54 th	55-64 th	
Y	bekerja	507	13191	2314	16012
	tidak bekerja	2334	8629	1497	12460
Total		2841	21820	3811	28472

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.890E3 ^a	2	.000
Likelihood Ratio	1.966E3	2	.000
Linear-by-Linear Association	976.873	1	.000
N of Valid Cases	28472		

• **Y dengan X₃**

		status_perkawinan		
		belum kawin	kawin	lainnya
Y	bekerja	1201	12624	2187
	tidak bekerja	2538	8890	1032
Total		3739	21514	3219

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.115E3 ^a	2	.000
Likelihood Ratio	1.120E3	2	.000
Linear-by-Linear Association	955.705	1	.000
N of Valid Cases	28472		

- **Y dengan X₄**

		hubungan				Total
		kepala RT	istri	anak	lainnya	
Y	bekerja	2054	10740	2108	1110	16012
	tidak bekerja	761	6605	3557	1537	12460
Total		2815	17345	5665	2647	28472

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.601E3 ^a	3	.000
Likelihood Ratio	1.612E3	3	.000
Linear-by-Linear Association	1.291E3	1	.000
N of Valid Cases	28472		

- **Y dengan X₅**

		ijazah_tertinggi				
		tidak punya ijazah	SD-SMA/ sederajat	D1/D2/D3	D4/S1	S2/S3
Y	bekerja	5383	9573	290	730	36
	tidak bekerja	2594	9617	85	156	8
Total		7977	19190	375	886	44

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.050E3 ^a	4	.000
Likelihood Ratio	1.093E3	4	.000
Linear-by-Linear Association	37.444	1	.000
N of Valid Cases	28472		

A.2 Ketepatan Klasifikasi *Learning-Testing*

- Kombinasi 95% dan 5%

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class
Class      1          2          Actual
Total
-----
1          10820.00   4365.00   15185.00
2           5255.00   6598.00   11853.00
-----
PRED. TOT.      16075.00   10963.00   27038.00
CORRECT         0.713     0.557
SUCCESS IND.    0.151     0.118
TOT. CORRECT    0.644
SENSITIVITY:    0.713     SPECIFICITY:    0.557
FALSE REFERENCE: 0.327   FALSE RESPONSE: 0.398
REFERENCE = "1", RESPONSE = "2"
-----

```

```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class
Class      1          2          Actual
Total
-----
1           597.00    230.00    827.00
2           282.00    325.00    607.00
-----
PRED. TOT.      879.00    555.00    1434.00
CORRECT         0.722     0.535
SUCCESS IND.    0.145     0.112
TOT. CORRECT    0.643
SENSITIVITY:    0.722     SPECIFICITY:    0.535
FALSE REFERENCE: 0.321   FALSE RESPONSE: 0.414

```

REFERENCE = "1", RESPONSE = "2"

• **Kombinasi 90% dan 10%**

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	10277.00	4147.00	14424.00
2	4968.00	6233.00	11201.00
PRED. TOT.	15245.00	10380.00	25625.00
CORRECT	0.712	0.556	
SUCCESS IND.	0.150	0.119	
TOT. CORRECT	0.644		
SENSITIVITY:	0.712	SPECIFICITY:	0.556
FALSE REFERENCE:	0.326	FALSE RESPONSE:	0.400
REFERENCE = "1", RESPONSE = "2"			

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	1159.00	429.00	1588.00
2	580.00	679.00	1259.00
PRED. TOT.	1739.00	1108.00	2847.00
CORRECT	0.730	0.539	
SUCCESS IND.	0.172	0.097	
TOT. CORRECT	0.646		
SENSITIVITY:	0.730	SPECIFICITY:	0.539
FALSE REFERENCE:	0.334	FALSE RESPONSE:	0.387
REFERENCE = "1", RESPONSE = "2"			

• **Kombinasi 85% dan 15%**

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	9711.00	3895.00	13606.00
2	4705.00	5886.00	10591.00
PRED. TOT.	14416.00	9781.00	24197.00
CORRECT	0.714	0.556	
SUCCESS IND.	0.151	0.118	
TOT. CORRECT	0.645		
SENSITIVITY:	0.714	SPECIFICITY:	0.556

FALSE REFERENCE: 0.326 FALSE RESPONSE: 0.398
 REFERENCE = "1", RESPONSE = "2"

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	1725.00	681.00	2406.00
2	843.00	1026.00	1869.00

PRED. TOT.	2568.00	1707.00	4275.00
CORRECT	0.717	0.549	
SUCCESS IND.	0.154	0.112	
TOT. CORRECT	0.644		
SENSITIVITY:	0.717	SPECIFICITY:	0.549
FALSE REFERENCE:	0.328	FALSE RESPONSE:	0.399
REFERENCE = "1", RESPONSE = "2"			

• **Kombinasi 80% dan 20%**

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	9138.00	3668.00	12806.00
2	4404.00	5547.00	9951.00

PRED. TOT.	13542.00	9215.00	22757.00
CORRECT	0.714	0.557	
SUCCESS IND.	0.151	0.120	
TOT. CORRECT	0.645		
SENSITIVITY:	0.714	SPECIFICITY:	0.557
FALSE REFERENCE:	0.325	FALSE RESPONSE:	0.398
REFERENCE = "1", RESPONSE = "2"			

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	1	2	
1	2298.00	908.00	3206.00
2	1144.00	1365.00	2509.00

PRED. TOT.	3442.00	2273.00	5715.00
CORRECT	0.717	0.544	
SUCCESS IND.	0.156	0.105	
TOT. CORRECT	0.641		
SENSITIVITY:	0.717	SPECIFICITY:	0.544
FALSE REFERENCE:	0.332	FALSE RESPONSE:	0.399
REFERENCE = "1", RESPONSE = "2"			

- **Kombinasi 75% dan 25%**

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      1          2          Total
-----
1           8591.00    3455.00    12046.00
2           4159.00    5190.00    9349.00
-----
PRED. TOT.    12750.00    8645.00    21395.00
CORRECT      0.713      0.555
SUCCESS IND. 0.150      0.118
TOT. CORRECT 0.644

      SENSITIVITY:      0.713      SPECIFICITY:      0.555
FALSE REFERENCE:      0.326  FALSE RESPONSE:      0.400
REFERENCE = "1", RESPONSE = "2"
-----
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      1          2          Total
-----
1           2845.00    1121.00    3966.00
2           1389.00    1722.00    3111.00
-----
PRED. TOT.    4234.00    2843.00    7077.00
CORRECT      0.717      0.554
SUCCESS IND. 0.157      0.114
TOT. CORRECT 0.645

      SENSITIVITY:      0.717      SPECIFICITY:      0.554
FALSE REFERENCE:      0.328  FALSE RESPONSE:      0.394
REFERENCE = "1", RESPONSE = "2"
-----

```

A.3 Node Information Pohon Optimal

```

=====
NODE INFORMATION
=====
*****
*           Node 1: UMUR           *
*           N: 25625              *
*****

*****
*           Node 2           *   = Terminal Node 18   =
*           N: 23089         *   =           N: 2536       =
*****

Node 1 was split on UMUR
A case goes left if UMUR = (2,3)
Improvement = 0.031          Complexity Threshold = 0.078

```

Node	Cases	Wgt Counts	Cost	Class
1	25625	25625.00	0.500	1
2	23089	23089.00	0.456	1
-18	2536	2536.00	0.143	2

Weighted Counts

Class	Top	Left	Right
1	14424.00	13976.00	448.00
2	11201.00	9113.00	2088.00

Within Node Probabilities

Class	Top	Left	Right
1	0.500	0.544	0.143
2	0.500	0.456	0.857

Surrogate Split

Surrogate	Assoc.	Improve.
1 STATUS_P s 2,3	0.375	0.017

Competitor Split

Competitor	Improve.
1 HUBUNGAN 1,2	0.026
2 IJAZAH_T 1,3,4,5	0.018
3 STATUS_P 2,3	0.017
4 KLASIFIK 2	0.001

 * Node 2: IJAZAH_T *
 * N: 23089 *

 * Node 3 * * Node 10 *
 * N: 8268 * * N: 14821 *

Node 2 was split on IJAZAH_T
 A case goes left if IJAZAH_T = (1,3,4,5)
 Improvement = 0.010 Complexity Threshold = 0.015

Node	Cases	Wgt Counts	Cost	Class
2	23089	23089.00	0.456	1
3	8268	8268.00	0.355	1
10	14821	14821.00	0.489	2

Weighted Counts

Class	Top	Left	Right
1	13976.00	5794.00	8182.00
2	9113.00	2474.00	6639.00

Within Node Probabilities

Class	Top	Left	Right
1	0.544	0.645	0.489
2	0.456	0.355	0.511

Surrogate Split

Surrogate	Assoc.	Improve.
1 UMUR s 3	0.158	.185406E-05
2 STATUS_P s 3	0.098	0.002
3 HUBUNGAN s 1	0.047	0.004

Competitor Split

Competitor	Improve.
1 HUBUNGAN 1,2	0.007
2 STATUS_P 3	0.002
3 KLASIFIK 2	0.001
4 UMUR 3	.185406E-05

 * Node 3: HUBUNGAN *
 * N: 8268 *

 * Node 4 * * Node 7 *
 * N: 7463 * * N: 805 *

Node 3 was split on HUBUNGAN
 A case goes left if HUBUNGAN = (1,2,3)

```

Improvement = 0.003          Complexity Threshold = 0.003

Node   Cases   Wgt Counts   Cost Class
  3     8268     8268.00     0.355 1
  4     7463     7463.00     0.332 1
  7      805      805.00     0.445 2

                          Weighted Counts
Class   Top      Left      Right
  1     5794.00   5385.00   409.00
  2     2474.00   2078.00   396.00

                          Within Node Probabilities
Class   Top      Left      Right
  1         0.645   0.668   0.445
  2         0.355   0.332   0.555

Competitor      Split
  1 IJAZAH_T     3,4,5
  2 UMUR          2
  3 STATUS_P     2
  4 KLASIFIK     2

*****
*           Node 4: IJAZAH_T           *
*           N: 7463                     *
*****

=====
=           Terminal Node 1           = *           Node 5           *
=           N: 1079                   = *           N: 6384           *
=====
*****

```

```

Node 4 was split on IJAZAH_T
A case goes left if IJAZAH_T = (3,4,5)
Improvement = 0.001          Complexity Threshold = .564786E-03

```

```

Node   Cases   Wgt Counts   Cost Class
  4     7463     7463.00     0.332 1
 -1     1079     1079.00     0.227 1
  5     6384     6384.00     0.349 1

                          Weighted Counts
Class   Top      Left      Right
  1     5385.00   879.00    4506.00
  2     2078.00   200.00   1878.00

                          Within Node Probabilities
Class   Top      Left      Right
  1         0.668   0.773   0.651
  2         0.332   0.227   0.349

Surrogate      Split      Assoc.   Improve.
  1 STATUS_P s 1           0.013   .170463E-03

Competitor      Split      Improve.
  1 HUBUNGAN     1           .407702E-03
  2 UMUR         2           .232817E-03
  3 STATUS_P     3           .200135E-03
  4 KLASIFIK     2           .170626E-03

*****
*           Node 5: HUBUNGAN           *
*           N: 6384                     *
*****

=====
=           Terminal Node 2           = *           Node 6           *
=           N: 6078                   = *           N: 306           *
=====
*****

```

```

Node 5 was split on HUBUNGAN
A case goes left if HUBUNGAN = (1,2)
Improvement = .992480E-03   Complexity Threshold = 0.001

Node   Cases   Wgt Counts   Cost Class

```

5	6384	6384.00	0.349	1		
-2	6078	6078.00	0.339	1		
6	306	306.00	0.453	2		

Weighted Counts

Class	Top	Left	Right
1	4506.00	4348.00	158.00
2	1878.00	1730.00	148.00

Within Node Probabilities

Class	Top	Left	Right
1	0.651	0.661	0.453
2	0.349	0.339	0.547

Surrogate Split

1 STATUS_P	s 2,3	Assoc.	Improve.
Competitor	Split	0.236	.897390E-03
1 STATUS_P	2,3		Improve.
2 KLASIFIK	2		.897390E-03
3 UMUR	2		.514613E-03
			.600137E-04

* Node 6: STATUS_P *
* N: 306 *

```

=====
= Terminal Node 3 = = Terminal Node 4 =
= N: 210 = = N: 96 =
=====

```

Node 6 was split on STATUS_P
A case goes left if STATUS_P = (2,3)
Improvement = .316827E-03 Complexity Threshold = .538752E-03

Node	Cases	Wgt Counts	Cost	Class
6	306	306.00	0.453	2
-3	210	210.00	0.467	1
-4	96	96.00	0.289	2

Weighted Counts

Class	Top	Left	Right
1	158.00	125.00	33.00
2	148.00	85.00	63.00

Within Node Probabilities

Class	Top	Left	Right
1	0.453	0.533	0.289
2	0.547	0.467	0.711

Competitor Split

1 KLASIFIK	2	Improve.
2 UMUR	3	.303488E-04
		.519477E-06

* Node 7: UMUR *
* N: 805 *

```

=====
* Node 8 * = Terminal Node 8 =
* N: 375 * = N: 430 =
=====

```

Node 7 was split on UMUR
A case goes left if UMUR = (2)
Improvement = .766591E-03 Complexity Threshold = 0.002

Node	Cases	Wgt Counts	Cost	Class
7	805	805.00	0.445	2
8	375	375.00	0.434	1
-8	430	430.00	0.345	2

Weighted Counts

Class	Top	Left	Right
1	409.00	235.00	174.00
2	396.00	140.00	256.00

Within Node Probabilities

Class	Top	Left	Right
1	0.445	0.566	0.345
2	0.555	0.434	0.655

Surrogate	Split	Assoc.	Improve.
1 STATUS_P s 1,2		0.346	.239741E-03
2 IJAZAH_T s 3,4,5		0.228	.617318E-03

Competitor	Split	Improve.
1 IJAZAH_T 3,4		.630477E-03
2 STATUS_P 2		.314740E-03
3 KLASIFIK 2		.742103E-05

```

*****
*           Node 8: IJAZAH_T           *
*           N: 375                       *
*****

```

```

=====
=           Terminal Node 5           = *           Node 9           *
=           N: 91                     = *           N: 284           *
=====
*****

```

Node 8 was split on IJAZAH_T
A case goes left if IJAZAH_T = (3,4)
Improvement = .323643E-03 Complexity Threshold = .345258E-03

Node	Cases	Wgt Counts	Cost	Class
8	375	375.00	0.434	1
-5	91	91.00	0.241	1
9	284	284.00	0.492	1

Class	Top	Left	Right
1	235.00	73.00	162.00
2	140.00	18.00	122.00

Class	Top	Left	Right
1	0.566	0.759	0.508
2	0.434	0.241	0.492

Competitor	Split	Improve.
1 STATUS_P 2,3		.824530E-04
2 KLASIFIK 1		.120597E-04

```

*****
*           Node 9: STATUS_P           *
*           N: 284                       *
*****

```

```

=====
=           Terminal Node 6           = =           Terminal Node 7           =
=           N: 233                     = =           N: 51                     =
=====

```

Node 9 was split on STATUS_P
A case goes left if STATUS_P = (2,3)
Improvement = .154657E-03 Complexity Threshold = .690516E-03

Node	Cases	Wgt Counts	Cost	Class
9	284	284.00	0.492	1
-6	233	233.00	0.452	1
-7	51	51.00	0.334	2

Class	Top	Left	Right
1	162.00	142.00	20.00
2	122.00	91.00	31.00

Class	Top	Left	Right
1	0.508	0.548	0.334
2	0.492	0.452	0.666

Competitor	Split	Improve.
------------	-------	----------

```

1 KLASIFIK      2                               .258460E-04
2 IJAZAH_T     1                               .795324E-06

*****
*           Node 10: HUBUNGAN                    *
*           N: 14821                             *
*****

*****
*           Node 11                             *
*           N: 11208                             *
*****

*****
*           Node 14                             *
*           N: 3613                              *
*****

Node 10 was split on HUBUNGAN
A case goes left if HUBUNGAN = (1,2)
Improvement = 0.003      Complexity Threshold = 0.018

Node   Cases   Wgt Counts   Cost Class
  10    14821    14821.00     0.489 2
  11    11208    11208.00     0.479 1
  14     3613     3613.00     0.394 2

Weighted Counts
Class      Top      Left      Right
  1         8182.00   6534.00   1648.00
  2         6639.00   4674.00   1965.00

Within Node Probabilities
Class      Top      Left      Right
  1         0.489    0.521    0.394
  2         0.511    0.479    0.606

Surrogate      Split      Assoc.   Improve.
  1 STATUS_P s 2,3      0.223   .397988E-04
Competitor      Split      Improve.
  1 STATUS_P      3              0.002
  2 KLASIFIK      2              .430420E-03
  3 UMUR          2              .164710E-04

*****
*           Node 11: STATUS_P                    *
*           N: 11208                             *
*****

*****
*           Node 12                             *
*           N: 708                               *
*****

*****
*           Node 13                             *
*           N: 10500                             *
*****

Node 11 was split on STATUS_P
A case goes left if STATUS_P = (1,3)
Improvement = 0.001      Complexity Threshold = 0.007

Node   Cases   Wgt Counts   Cost Class
  11    11208    11208.00     0.479 1
  12     708     708.00       0.322 1
  13   10500    10500.00     0.490 1

Weighted Counts
Class      Top      Left      Right
  1         6534.00   517.00    6017.00
  2         4674.00   191.00   4483.00

Within Node Probabilities
Class      Top      Left      Right
  1         0.521    0.678    0.510
  2         0.479    0.322    0.490

Surrogate      Split      Assoc.   Improve.
  1 HUBUNGAN s 1      0.503   .965998E-03
Competitor      Split      Improve.
  1 KLASIFIK      2              0.001
  2 HUBUNGAN      1              .965998E-03
  3 UMUR          2              .505755E-04

*****
*           Node 12: UMUR                       *
*           N: 708                               *
*****

```

```

=====
=      Terminal Node 9      =      Terminal Node 10      =
=      N: 500              =      N: 208                =
=====

Node 12 was split on UMUR
A case goes left if UMUR = (2)
Improvement = 0.001      Complexity Threshold = .561520E-03
Node      Cases      Wgt Counts      Cost Class
   12      708      708.00      0.322 1
   -9      500      500.00      0.227 1
  -10      208      208.00      0.466 2
                          Weighted Counts
Class      Top      Left      Right
1          517.00    407.00    110.00
2          191.00     93.00     98.00
                          Within Node Probabilities
Class      Top      Left      Right
1          0.678     0.773     0.466
2          0.322     0.227     0.534
Competitor      Split      Improve.
1 KLASIFIK      2          .332202E-04
2 STATUS_P      1          .161082E-04
*****
*      Node 13: KLASIFIK      *
*      N: 10500              *
*****

=====
=      Terminal Node 11      =      Terminal Node 12      =
=      N: 6292              =      N: 4208                =
=====

Node 13 was split on KLASIFIK
A case goes left if KLASIFIK = (2)
Improvement = 0.001      Complexity Threshold = 0.013
Node      Cases      Wgt Counts      Cost Class
   13     10500     10500.00     0.490 1
  -11     6292     6292.00     0.455 1
  -12     4208     4208.00     0.460 2
                          Weighted Counts
Class      Top      Left      Right
1          6017.00    3815.00    2202.00
2          4483.00    2477.00    2006.00
                          Within Node Probabilities
Class      Top      Left      Right
1          0.510     0.545     0.460
2          0.490     0.455     0.540
Competitor      Split      Improve.
1 HUBUNGAN      1          .346624E-05
2 UMUR          2          .255212E-05
*****
*      Node 14: STATUS_P      *
*      N: 3613              *
*****
*****
*      Node 15      *      Terminal Node 17      =
*      N: 1238      *      N: 2375                =
*****

Node 14 was split on STATUS_P
A case goes left if STATUS_P = (1,3)
Improvement = 0.002      Complexity Threshold = 0.001
Node      Cases      Wgt Counts      Cost Class
   14      3613     3613.00     0.394 2
   15      1238     1238.00     0.488 1
  -17      2375     2375.00     0.336 2
                          Weighted Counts
Class      Top      Left      Right
1          1648.00    712.00     936.00

```

```

2          1965.00      526.00      1439.00
                Within Node Probabilities
Class      Top      Left      Right
1          0.394      0.512      0.336
2          0.606      0.488      0.664

Surrogate
1 UMUR      s 3
  Competitor Split
1 KLASIFIK  1
2 HUBUNGAN  3
3 UMUR      2

```

```

Assoc. Improve.
0.062 .404459E-04
                Improve.
.365260E-03
.890098E-04
.404459E-04

```

```

*****
*          Node 15: UMUR          *
*          N: 1238                *
*****

```

```

*****
*          Node 16          *      =      Terminal Node 16      =
*          N: 1140          *      =      N: 98                  =
*****

```

```

Node 15 was split on UMUR
A case goes left if UMUR = (2)
Improvement = .323010E-03      Complexity Threshold = 0.001

```

```

Node   Cases   Wgt Counts   Cost Class
  15     1238     1238.00     0.488 1
  16     1140     1140.00     0.470 1
 -16      98       98.00       0.320 2

```

```

Weighted Counts
Class      Top      Left      Right
1          712.00     675.00     37.00
2          526.00     465.00     61.00

```

```

Within Node Probabilities
Class      Top      Left      Right
1          0.512      0.530      0.320
2          0.488      0.470      0.680

```

```

Competitor Split
1 KLASIFIK  1
2 STATUS_P  3
3 HUBUNGAN  3
                Improve.
.360007E-04
.247960E-04
.167930E-04

```

```

*****
*          Node 16: STATUS_P      *
*          N: 1140                *
*****

```

```

=====
=      Terminal Node 13      =      *      Node 17          *
=          N: 282            =      *      N: 858            *
=====

```

```

Node 16 was split on STATUS_P
A case goes left if STATUS_P = (3)
Improvement = .208852E-03      Complexity Threshold = 0.001

```

```

Node   Cases   Wgt Counts   Cost Class
  16     1140     1140.00     0.470 1
 -13     282     282.00     0.384 1
  17     858     858.00     0.498 1

```

```

Weighted Counts
Class      Top      Left      Right
1          675.00     190.00     485.00
2          465.00     92.00     373.00

```

```

Within Node Probabilities
Class      Top      Left      Right
1          0.530      0.616      0.502
2          0.470      0.384      0.498

```

```

Competitor Split
1 KLASIFIK  1
2 HUBUNGAN  4
                Improve.
.104740E-03
.485003E-04

```

```

*****
*          Node 17: KLASIFIK      *
*****

```



```

*                               N: 858                               *
*****
=====
=      Terminal Node 14      =      Terminal Node 15      =
=      N: 480                =      N: 378                =
=====
Node 17 was split on KLASIFIK
A case goes left if KLASIFIK = (1)
Improvement = .270584E-03      Complexity Threshold = 0.002
Node   Cases   Wgt Counts      Cost Class
  17     858     858.00      0.498 1
 -14     480     480.00      0.440 1
 -15     378     378.00      0.432 2
                               Weighted Counts
Class   Top      Left      Right
  1     485.00   298.00   187.00
  2     373.00   182.00   191.00
                               Within Node Probabilities
Class   Top      Left      Right
  1     0.502    0.560    0.432
  2     0.498    0.440    0.568
Competitor      Split
  1 HUBUNGAN      4
                                           Improve.
                                           .350451E-04

```

A.4 Tree Detail Pohon Optimal

- Gambar -

B. KOTAMADYA

B.1 Tabulasi Silang Variabel Respon dan Prediktor

- Y dengan X_1

		klasifikasi		Total
		kota	desa	
Y	bekerja	3470	71	3541
	tidak bekerja	3169	74	3243
Total		6639	145	6784

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.620 ^a	1	.431
Continuity Correction ^b	.495	1	.482
Likelihood Ratio	.619	1	.431
Linear-by-Linear Association	.620	1	.431
N of Valid Cases ^b	6784		

- Y dengan X_2

		umur		
		15-19 th	20-54 th	55-64 th
Y	bekerja	155	3021	365
	tidak bekerja	634	2198	411
Total		789	5219	776

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4.110E2 ^a	2	.000
Likelihood Ratio	431.936	2	.000
Linear-by-Linear Association	120.334	1	.000
N of Valid Cases	6784		

- **Y dengan X₃**

		status_perkawinan		
		belum kawin	kawin	lainnya
Y	bekerja	650	2426	465
	tidak bekerja	837	2147	259
Total		1487	4573	724

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	86.228 ^a	2	.000
Likelihood Ratio	86.948	2	.000
Linear-by-Linear Association	85.753	1	.000
N of Valid Cases	6784		

- **Y dengan X₄**

		hubungan				Total
		kepala RT	istri	anak	lainnya	
Y	bekerja	420	2058	698	365	3541
	tidak bekerja	226	1774	900	343	3243
Total		646	3832	1598	708	6784

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	92.615 ^a	3	.000
Likelihood Ratio	93.428	3	.000
Linear-by-Linear Association	49.024	1	.000
N of Valid Cases	6784		

- **Y dengan X₅**

		ijazah_tertinggi				
		tidak punya ijazah	SD-SMA/sederajat	D1/D2/D3	D4/S1	S2/S3
Y	bekerja	475	2432	136	468	30
	tidak bekerja	313	2688	84	149	9
Total		788	5120	220	617	39

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.220E2 ^a	4	.000
Likelihood Ratio	230.789	4	.000
Linear-by-Linear Association	83.060	1	.000
N of Valid Cases	6784		

B.2 Ketepatan Klasifikasi *Learning-Testing*

- Kombinasi 95% dan 5%

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      1          2          Total
-----
1          1498.00    1861.00    3359.00
2          712.00    2365.00    3077.00
-----
PRED. TOT.      2210.00    4226.00    6436.00
CORRECT        0.446      0.769
SUCCESS IND.   -0.076     0.291
TOT. CORRECT   0.600
SENSITIVITY:   0.446      SPECIFICITY: 0.769
FALSE REFERENCE: 0.322  FALSE RESPONSE: 0.440
REFERENCE = "1", RESPONSE = "2"
-----

```

```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      1          2          Total
-----
1          76.00     106.00     182.00
2          35.00     131.00     166.00
-----
PRED. TOT.      111.00     237.00     348.00
CORRECT        0.418      0.789
SUCCESS IND.   -0.105     0.312
TOT. CORRECT   0.595
SENSITIVITY:   0.418      SPECIFICITY: 0.789
FALSE REFERENCE: 0.315  FALSE RESPONSE: 0.447
REFERENCE = "1", RESPONSE = "2"
-----

```

- Kombinasi 90% dan 10%

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      1          2          Total
-----
1          1409.00    1750.00    3159.00
2          677.00    2240.00    2917.00
-----
PRED. TOT.      2086.00    3990.00    6076.00
CORRECT        0.446      0.768
SUCCESS IND.   -0.074     0.288
TOT. CORRECT   0.601
SENSITIVITY:   0.446      SPECIFICITY: 0.768
FALSE REFERENCE: 0.325  FALSE RESPONSE: 0.439
REFERENCE = "1", RESPONSE = "2"
-----

```

=====
 TEST SAMPLE CLASSIFICATION TABLE
 =====

Actual Class	Predicted Class		Actual Total
	1	2	
1	165.00	217.00	382.00
2	70.00	256.00	326.00
PRED. TOT.	235.00	473.00	708.00
CORRECT	0.432	0.785	
SUCCESS IND.	-0.108	0.325	
TOT. CORRECT	0.595		
SENSITIVITY:	0.432	SPECIFICITY:	0.785
FALSE REFERENCE:	0.298	FALSE RESPONSE:	0.459
REFERENCE = "1", RESPONSE = "2"			

• **Kombinasi 85% dan 15%**

=====
 LEARNING SAMPLE CLASSIFICATION TABLE
 =====

Actual Class	Predicted Class		Actual Total
	1	2	
1	1330.00	1659.00	2989.00
2	644.00	2122.00	2766.00
PRED. TOT.	1974.00	3781.00	5755.00
CORRECT	0.445	0.767	
SUCCESS IND.	-0.074	0.287	
TOT. CORRECT	0.600		
SENSITIVITY:	0.445	SPECIFICITY:	0.767
FALSE REFERENCE:	0.326	FALSE RESPONSE:	0.439
REFERENCE = "1", RESPONSE = "2"			

=====
 TEST SAMPLE CLASSIFICATION TABLE
 =====

Actual Class	Predicted Class		Actual Total
	1	2	
1	244.00	308.00	552.00
2	103.00	374.00	477.00
PRED. TOT.	347.00	682.00	1029.00
CORRECT	0.442	0.784	
SUCCESS IND.	-0.094	0.321	
TOT. CORRECT	0.601		
SENSITIVITY:	0.442	SPECIFICITY:	0.784
FALSE REFERENCE:	0.297	FALSE RESPONSE:	0.452
REFERENCE = "1", RESPONSE = "2"			

- **Kombinasi 80% dan 20%**

```
=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
```

Actual Class	Predicted Class		Actual Total
	1	2	
1	1256.00	1549.00	2805.00
2	594.00	1993.00	2587.00

```
-----
PRED. TOT.      1850.00      3542.00      5392.00
CORRECT         0.448         0.770
SUCCESS IND.    -0.072         0.291
TOT. CORRECT    0.603
SENSITIVITY:      0.448      SPECIFICITY:      0.770
FALSE REFERENCE: 0.321  FALSE RESPONSE: 0.437
REFERENCE = "1", RESPONSE = "2"
```

```
=====
TEST SAMPLE CLASSIFICATION TABLE
=====
```

Actual Class	Predicted Class		Actual Total
	1	2	
1	318.00	418.00	736.00
2	153.00	503.00	656.00

```
-----
PRED. TOT.      471.00      921.00      1392.00
CORRECT         0.432         0.767
SUCCESS IND.    -0.097         0.296
TOT. CORRECT    0.590
SENSITIVITY:      0.432      SPECIFICITY:      0.767
FALSE REFERENCE: 0.325  FALSE RESPONSE: 0.454
REFERENCE = "1", RESPONSE = "2"
```

- **Kombinasi 75% dan 25%**

```
=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
```

Actual Class	Predicted Class		Actual Total
	1	2	
1	1219.00	1432.00	2651.00
2	543.00	1911.00	2454.00

```
-----
PRED. TOT.      1762.00      3343.00      5105.00
CORRECT         0.460         0.779
SUCCESS IND.    -0.059         0.298
TOT. CORRECT    0.612
```

```
SENSITIVITY:      0.460      SPECIFICITY:      0.779
FALSE REFERENCE: 0.308  FALSE RESPONSE: 0.428
REFERENCE = "1", RESPONSE = "2"
```



```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class       1          2          Total
-----
1           376.00     514.00     890.00
2           174.00     615.00     789.00
-----
PRED. TOT.      550.00     1129.00    1679.00
CORRECT         0.422     0.779
SUCCESS IND.    -0.108     0.310
TOT. CORRECT    0.591
SENSITIVITY:    0.422     SPECIFICITY: 0.779
FALSE REFERENCE: 0.316   FALSE RESPONSE: 0.455
REFERENCE = "1", RESPONSE = "2"
-----

```

B.3 Node Information Pohon Optimal

```

*****
*                Node 1: UMUR                *
*                N: 5755                      *
*****
*                Node 2                *   =   Terminal Node 6   =
*                N: 5090                *   =   N: 665           =
*****
Node 1 was split on UMUR
A case goes left if UMUR = (2,3)
Improvement = 0.027          Complexity Threshold = 0.075
Node      Cases   Wgt Counts   Cost Class
1         5755    5755.00     0.500 1
2         5090    5090.00     0.458 1
-6        665     665.00     0.185 2

Weighted Counts
Class     Top         Left         Right
1         2989.00     2858.00     131.00
2         2766.00     2232.00     534.00
Within Node Probabilities
Class     Top         Left         Right
1         0.500     0.542     0.185
2         0.500     0.458     0.815
Surrogate      Split          Assoc.   Improve.
1 STATUS_P s 2,3          0.031    0.004
Competitor
1 IJAZAH_T    1,3,4,5          0.013
2 HUBUNGAN    1,2,4            0.004
3 STATUS_P    2,3              0.004
4 KLASIFIK    1                .209655E-05
*****
*                Node 2: IJAZAH_T          *
*                N: 5090                      *
*****
=====
=   Terminal Node 1   =   *   Node 3   *
=   N: 559           =   *   N: 4531  *
=====
Node 2 was split on IJAZAH_T
A case goes left if IJAZAH_T = (4,5)

```

```

Improvement = 0.008          Complexity Threshold = 0.010
Node  Cases  Wgt Counts  Cost Class
  2    5090    5090.00    0.458 1
 -1     559     559.00    0.262 1
  3    4531    4531.00    0.481 1

          Weighted Counts
Class      Top      Left      Right
1          2858.00   421.00   2437.00
2          2232.00   138.00   2094.00

          Within Node Probabilities
Class      Top      Left      Right
1          0.542    0.738    0.519
2          0.458    0.262    0.481

          Competitor      Split      Improve.
1 STATUS_P      1,3          0.006
2 HUBUNGAN      1,3          0.003
3 UMUR          2            0.003
4 KLASIFIK      1            .135642E-06

          *****
          *          Node 3: STATUS_P          *
          *          N: 4531                  *
          *****

          *****
          *          Node 4          *          *          Node 5          *
          *          N: 1083          *          *          N: 3448          *
          *****

Node 3 was split on STATUS_P
A case goes left if STATUS_P = (1,3)
Improvement = 0.005          Complexity Threshold = 0.014
Node  Cases  Wgt Counts  Cost Class
  3    4531    4531.00    0.481 1
  4    1083    1083.00    0.380 1
  5    3448    3448.00    0.487 2

          Weighted Counts
Class      Top      Left      Right
1          2437.00   691.00   1746.00
2          2094.00   392.00   1702.00

          Within Node Probabilities
Class      Top      Left      Right
1          0.519    0.620    0.487
2          0.481    0.380    0.513

          Surrogate      Split      Assoc.      Improve.
1 HUBUNGAN s 1,3,4          0.516          0.002
          Competitor      Split      Improve.
1 HUBUNGAN      1,3          0.003
2 IJAZAH_T      1,3          0.001
3 UMUR          2            0.001
4 KLASIFIK      2            .282184E-04

          *****
          *          Node 4: UMUR          *
          *          N: 1083                  *
          *****

=====
=          Terminal Node 2          =          Terminal Node 3          =
=          N: 840                    =          N: 243                    =
=====

Node 4 was split on UMUR
A case goes left if UMUR = (2)
Improvement = 0.002          Complexity Threshold = 0.003
Node  Cases  Wgt Counts  Cost Class
  4    1083    1083.00    0.380 1
 -2     840     840.00    0.336 1
 -3     243     243.00    0.470 2

          Weighted Counts
Class      Top      Left      Right
1          691.00    572.00   119.00

```

```

2          392.00      268.00      124.00
          Within Node Probabilities
Class      Top      Left      Right
1          0.620      0.664      0.470
2          0.380      0.336      0.530

Competitor      Split
1 HUBUNGAN      1,2,3      Improve.
2 STATUS_P      1          .195537E-03
3 IJAZAH_T      2          .166857E-03
4 KLASIFIK      2          .285775E-05
                    .117874E-06

*****
*              Node 5: IJAZAH_T      *
*              N: 3448              *
*****

=====
=      Terminal Node 4      =      Terminal Node 5      =
=              N: 575      =      N: 2873      =
=====

Node 5 was split on IJAZAH_T
A case goes left if IJAZAH_T = (1,3)
Improvement = 0.002      Complexity Threshold = 0.013

Node      Cases      Wgt Counts      Cost      Class
5          3448      3448.00      0.487      2
-4          575      575.00      0.433      1
-5          2873      2873.00      0.471      2

Weighted Counts
Class      Top      Left      Right
1          1746.00      337.00      1409.00
2          1702.00      238.00      1464.00

Within Node Probabilities
Class      Top      Left      Right
1          0.487      0.567      0.471
2          0.513      0.433      0.529

Competitor      Split
1 UMUR          2          Improve.
2 HUBUNGAN      1,2      .462205E-03
3 KLASIFIK      2          .397857E-03
                    .738480E-04

```

B.4 Tree Detail Pohon Optimal

- Gambar-