

**PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN
CLASSIFICATION AND REGRESSION TREE PADA KLASIFIKASI
STATUS KESEJAHTERAAN RUMAH TANGGA DI KOTA BATU**

SKRIPSI

**OLEH
AISA KHOIRUL UMAROH
NIM. 16610030**



**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

**PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN
CLASSIFICATION AND REGRESSION TREE PADA KLASIFIKASI
STATUS KESEJAHTERAAN RUMAH TANGGA DI KOTA BATU**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Aisa Khoirul Umaroh
NIM. 16610030**

**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

**PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN
CLASSIFICATION AND REGRESSION TREE PADA KLASIFIKASI
STATUS KESEJAHTERAAN RUMAH TANGGA DI KOTA BATU**

SKRIPSI

Oleh
Aisa Khoirul Umaroh
NIM. 16610030

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 16 Juni 2020

Pembimbing I



Ria Dhea Layla N. K., S.Si., M.Si
NIP. 19900709 20189201 2 228

Pembimbing II



Muhammad Khudzaifah, S.Si., M.Si
NIP. 19900511 20160801 1 057

Mengetahui,
Ketua Jurusan Matematika



Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

**PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN
CLASSIFICATION AND REGRESSION TREE PADA KLASIFIKASI
STATUS KESEJAHTERAAN RUMAH TANGGA DI KOTA BATU**

SKRIPSI

**Oleh
Aisa Khoirul Umaroh
NIM. 16610030**

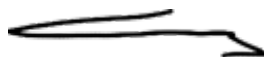
Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)

Tanggal 24 Juni 2020

Penguji Utama : Abdul Aziz, S.Si., M.Si
Ketua Penguji : Dr. Sri Harini, M.Si
Sekertaris Penguji : Ria Dhea Layla N. K., S.Si., M.Si
Anggota Penguji : Muhammad Khudzaifah, S.Si., M.Si



Mengetahui,
Ketua Jurusan Matematika



Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertandatangan di bawah ini:

Nama : Aisa Khoirul Umaroh

NIM : 16610030

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Perbandingan Metode Regresi Logistik Biner dan
Classification And Regression Tree pada Klasifikasi Status
Kesejahteraan Rumah Tangga di Kota Batu

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencamumkan sumber cuplikan atau daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 16 Juni 2020
Yang membuat pernyataan,



Aisa Khoirul Umaroh
NIM. 16610030

MOTO

“Memohon kepada Allah SWT dan Yakin bahwa
Dia akan Menjawab Doamu” – Tirmidzi

“Bismillah, Selalu Diyakini Bisa”



PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Kedua orang tua penulis, Bapak Yanto dan Ibu Ngatiani tercinta,
serta kedua kakak penulis, Anif Munandar dan Anis Safa'atul Anggraeni
yang senantiasa dengan ikhlas dan istiqomah mendoakan, memberikan kasih
sayang, nasihat, dan semangat.



KATA PENGANTAR

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah Swt atas rahmat, taufik serta hidayah-Nya, sehingga penulis mampu menyelesaikan penyusunan skripsi ini sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Dalam proses penyusunan skripsi ini, penulis banyak mendapat bimbingan dan arahan dari berbagai pihak. Untuk itu ucapan terima kasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya penulis sampaikan kepada:

1. Prof. Dr. H. Abd. Haris, M.Ag, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Usman Pagalay, M.Si, selaku ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Ria Dhea Layla N. K., S.Si., M.Si, selaku dosen pembimbing I yang telah banyak memberikan arahan, nasihat, motivasi, dan berbagi pengalaman yang berharga kepada penulis.
5. Muhaammad Khudzaifah, S.Si., M.Si., selaku dosen pembimbing II yang telah banyak memberikan arahan dan berbagi ilmunya kepada penulis.

6. Segenap sivitas akademika Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang terutama seluruh dosen, terima kasih atas segala ilmu dan bimbingannya.
7. Bapak dan ibu serta kakak-kakak tercinta yang selalu memberikan doa, semangat, serta motivasi kepada penulis sampai saat ini.
8. Sahabat-sahabat terbaik penulis, Lutfi Alwi Muzakka, Soulmate, Berdelapan, Baperan yang selalu menemani, membantu, dan memberikan dukungan sehingga penulis dapat menyelesaikan skripsi ini.
9. GenBI Korkom Malang, Bank Indonesia KPw Malang, HMJ “Integral” Matematika, DEMA F-SAINTEK, dan PMII Rayon “Pencerahan” Galileo, yang telah memberikan ilmu yang tak akan pernah saya dapat di bangku kuliah.
10. Seluruh teman-teman di Jurusan Matematika angkatan 2016 “Mathematics 2016” khususnya teman-teman peminatan statistika yang berjuang bersama-sama.

Semoga Allah SWT melimpahkan rahmat dan karunia-Nya kepada kita semua. Akhirnya penulis berharap semoga dengan rahmat dan izin-Nya mudah-mudahan skripsi ini bermanfaat bagi penulis dan bagi pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 15 Juni 2020

Penulis

DAFTAR ISI

HALAMAN JUDUL	
HALAMAN PENGAJUAN	
HALAMAN PERSETUJUAN	
HALAMAN PENGESAHAN	
HALAMAN PERNYATAAN KEASLIAN TULISAN	
HALAMAN MOTO	
HALAMAN PERSEMBAHAN	
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
ABSTRAK	xiv
ABSTRACT	xv
ملخص	xvi
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	8
1.3 Tujuan Penelitian	8
1.4 Manfaat Penelitian	9
1.5 Batasan Masalah	9
1.6 Sistematika Penulisan	10
BAB II TINJAUAN PUSTAKA	
2.1 Regresi Logistik Biner	13
2.1.1 Estimasi Parameter	15
2.1.2 Uji Signifikansi Parameter.....	23
2.1.3 Uji Multikolinieritas	25
2.1.4 Uji Kesesuaian Model.....	26
2.1.5 Interpretasi Koefisien Regresi	26
2.2 <i>Classification and Regression Tree (CART)</i>	27
2.2.1 Pembentukan Pohon Klasifikasi	30
2.2.2 Pemangkasan Pohon Klasifikasi	33
2.2.3 Penentuan Pohon Klasifikasi Optimal	35
2.3 Menghitung Ketepatan Klasifikasi	36
2.4 Kesejahteraan dan Kemiskinan Rumah Tangga	38

2.4.1	Faktor-faktor Kesejahteraan Rumah Tangga.....	42
2.5	Integrasi Al-Qur'an.....	44

BAB III METODE PENELITIAN

3.1	Pendekatan Penelitian	47
3.2	Sumber Data	47
3.3	Variabel Penelitian.....	48
3.4	Langkah Penelitian	49
3.5	Flowchart.....	51

BAB IV ANALISIS DAN PEMBAHASAN

4.1	Analisis Deskriptif.....	53
4.1.1	Deskripsi Usia Kepala Rumah Tangga.....	54
4.1.2	Deskripsi Jenis Kelamin Kepala Rumah Tangga	55
4.1.3	Deskripsi Status Perkawinan Kepala Rumah Tangga.....	56
4.1.4	Deskripsi Banyak Anggota Rumah Tangga.....	57
4.1.5	Deskripsi Tingkat Pendidikan Kepala Rumah Tangga.....	58
4.1.6	Deskripsi Status Pekerjaan Kepala Rumah Tangga.....	60
4.1.7	Deskripsi Kesehatan Kepala Rumah Tangga.....	62
4.1.8	Deskripsi Status Kepemilikan Tempat Tinggal	63
4.1.9	Deskripsi Bahan Bakar Memasak.....	64
4.1.10	Deskripsi Kepemilikan Telepon Seluler	66
4.2	Analisis Regresi Logistik Biner	67
4.2.1	Uji Multikolinieritas	67
4.2.2	Estimasi Parameter	68
4.2.3	Uji Signifikansi Parameter.....	69
4.2.4	Model Regresi Logistik Biner.....	74
4.2.5	Uji Kesesuaian Model.....	75
4.2.6	Interpretasi Koefisien Regresi Logistik	76
4.2.7	Ketepatan Klasifikasi Regresi Logistik Biner	77
4.3	Analisis <i>Classification and Regression Tree</i> (CART).....	78
4.3.1	Pembentukan Pohon Klasifikasi Maksimal	78
4.3.2	Pemangkasan Pohon Klasifikasi Maksimal.....	82
4.3.3	Pohon Klasifikasi Optimal.....	84
4.3.4	Ketepatan Klasifikasi CART	87
4.4	Perbandingan Nilai Ketepatan Klasifikasi	88

BAB V PENUTUP

5.1	Kesimpulan	92
5.2	Saran	93

DAFTAR PUSTAKA	95
-----------------------------	-----------

LAMPIRAN

RIWAYAT HIDUP

DAFTAR TABEL

Tabel 2.1	Perhitungan Ketepatan Klasifikasi.....	37
Tabel 3.1	Variabel Prediktor.....	46
Tabel 4.1	Deskripsi Pengeluaran Per Kapita Per Bulan	54
Tabel 4.2	Usia Kepala Rumah Tangga (Tahun)	54
Tabel 4.3	Jenis Kelamin KRT Berdasarkan Status Kesejahteraan	55
Tabel 4.4	Status Perkawinan Berdasarkan Status Kesejahteraan	57
Tabel 4.5	Banyak Anggota Rumah Tangga.....	58
Tabel 4.6	Status Perkawinan Berdasarkan Status Kesejahteraan	59
Tabel 4.7	Status Pekerjaan KRT Berdasarkan Status Kesejahteraan	61
Tabel 4.8	Kesehatan KRT Berdasarkan Status Kesejahteraan	62
Tabel 4.9	Status Kepemilikan Tempat Tinggal	64
Tabel 4.10	Bahan Bakar Memasak Berdasarkan Status Kesejahteraan.....	65
Tabel 4.11	Status Kepemilikan HP Berdasarkan Status Kesejahteraan	67
Tabel 4.12	Nilai VIF.....	68
Tabel 4.13	Estimasi Parameter	68
Tabel 4.14	Uji Wald Pertama.....	70
Tabel 4.15	Uji Wald Kedua	71
Tabel 4.16	Uji G	73
Tabel 4.17	Koefisien Regresi.....	74
Tabel 4.18	Uji Hosmer dan Lemeshow	76
Tabel 4.19	<i>Odds Ratio</i>	76
Tabel 4.20	Ketepatan Klasifikasi Model	77
Tabel 4.21	Perhitungan Pemilahan Setiap Variabel	79
Tabel 4.22	Urutan Skor Variabel Terpenting	79
Tabel 4.23	Nilai <i>Complexity Parameter</i>	82
Tabel 4.24	Skor Variabel Pohon Optimal.....	85
Tabel 4.25	Ketepatan Klasifikasi CART	87
Tabel 4.26	Perbandingan Variabel Signifikan	88
Tabel 4.27	Perbandingan Ketepatan Klasifikasi	89

DAFTAR GAMBAR

Gambar 2.1	Struktur Pohon Klasifikasi	28
Gambar 2.2	Indikator Kemiskinan UND	42
Gambar 3.1	<i>Flowchart</i>	49
Gambar 4.1	Status Kesejahteraan Rumah Tangga.....	53
Gambar 4.2	Jenis Kelamin KRT	55
Gambar 4.3	Status Perkawinan KRT	56
Gambar 4.4	Tingkat Pendidikan KRT	58
Gambar 4.5	Status Pekerjaan KRT	60
Gambar 4.6	Kesehatan KRT	62
Gambar 4.7	Status Kepemilikan Tempat Tinggal.....	63
Gambar 4.8	Bahan Bakar Memasak	65
Gambar 4.9	Status Kepemilikan HP	66
Gambar 4.10	Pohon Klasifikasi Maksimal	82
Gambar 4.11	Plot <i>Complexity Parameter</i>	84
Gambar 4.12	Pohon Klasifikasi Optimal	84

ABSTRAK

Umaroh, Aisa Khoirul. 2020. **Perbandingan Metode Regresi Logistik Biner dan *Classification And Regression Tree* pada Klasifikasi Status Kesejahteraan Rumah Tangga di Kota Batu.** Skripsi. Jurusan Matematika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Ria Dhea Layla N. K., S.Si., M.Si. (II) Muhammad Khudzaifah, S.Si., M.Si.

Kata kunci: Klasifikasi, Regresi Logistik Biner, CART (*Classification and Regression Tree*), Kesejahteraan Rumah Tangga

Kesejahteraan merupakan keadaan yang menggambarkan suatu rumah tangga mampu memenuhi kebutuhan hidup. Banyak indikator yang dapat menentukan faktor kesejahteraan rumah tangga, yaitu faktor kependudukan, pendidikan, ketenagakerjaan, kesehatan, perumahan, dan teknologi informasi dan komunikasi. Penelitian ini bertujuan untuk mengidentifikasi faktor atau variabel yang berpengaruh terhadap klasifikasi status kesejahteraan rumah tangga di Kota Batu, yang diklasifikasikan menjadi kategori miskin dan tidak miskin. Variabel yang digunakan adalah usia, jenis kelamin, status perkawinan, tingkat pendidikan tertinggi, status pekerjaan dan kesehatan kepala rumah tangga, selain itu variabel banyak anggota rumah tangga, status kepemilikan tempat tinggal, bahan bakar memasak, dan kepemilikan telepon seluler. Metode yang digunakan adalah regresi logistik biner dan CART (*Classification and Regression Tree*). Hasil nilai ketepatan klasifikasi dari kedua metode tersebut akan dibandingkan. Berdasarkan hasil penelitian dengan metode regresi logistik biner dan CART menunjukkan bahwa variabel yang berpengaruh terhadap model klasifikasi adalah variabel usia dan status pekerjaan kepala rumah tangga. Nilai ketepatan klasifikasi model regresi logistik biner adalah 94,69%, lebih tinggi dibandingkan model klasifikasi CART adalah 89,36%. Sehingga model terbaik untuk memprediksi faktor kesejahteraan rumah tangga di Kota Batu adalah metode regresi logistik biner.

ABSTRACT

Umaroh, Aisa Khoirul. 2020. **Comparison of Binary Logistic Regression and the Classification and Regression Tree Method in the Classification of Household Welfare Status in Batu City**. Thesis. Department of Mathematics. Faculty of Science and Technology. Maulana Malik Ibrahim State Islamic University of Malang. Advisors: (I) Ria Dhea Layla NK, S.Si., M.Si. (II) Muhammad Khudzaifah, S.Si., M.Si.

Keywords: Classification, Logistic Regression Binary, CART (Classification and Regression Tree), Household Welfare

Welfare is a condition that describes a household ability to meet the necessities of life. Many indicators can determine household welfare factors, namely population, education, employment, health, housing, and information and communication technology. This study aims to identify factors or variables that influence the classification of household welfare status in Batu City, which are classified into poor and non-poor. The variables used are age, gender, marital status, the level of the highest education, employment status and the health of the head of the household, and the additional variables are number of family members, ownership of the residence, cooking fuel, and cell phone ownership. The methods used are binary logistic regression and CART (Classification and Regression Tree). The results of the classification accuracy of these methods will be compared. The results of research with binary logistic regression and CART methods show the variables that influence the classification model are age and employment status of the head of household. The value of the classification accuracy of the binary logistic regression model was 94.69%, higher than the classification model of CART 89.36%. Consequently, the best model for predicting household welfare factors in Batu City is the binary logistic regression method.

ملخص

خير الأمرة، عائشة. 2020 . مقارنة بين طريقة الانحدار اللوجستي الثنائي وشجرة التصنيف والانحدار في تصنيف حالة رعاية الأسرة في مدينة باتو. البحث العلمي الرياضيات. كلية العلوم والتكنولوجيا. جامعة مولانا مالك إبراهيم الإسلامية الحكومية بمالانغ. المشرفة: (ا) ريبا ديلا ليلي ن. ك، الماجستير، المشرف: (اا) محمد هديفة، الماجستير.

مفتاح الكلمات: التصنيف ، ثنائي الانحدار اللوجستي، CART، رعاية الأسرة.

الرفاهية هي حالة توضح أن الأسرة قادرة على تلبية احتياجات الحياة. يمكن للعديد من المؤشرات أن تحدد عوامل رفاهية الأسرة، أي السكان والتعليم والعمالة والصحة والإسكان وتكنولوجيا المعلومات والاتصالات. تهدف هذه الدراسة إلى تحديد العوامل أو المتغيرات التي تؤثر على تصنيف حالة رفاهية الأسرة في مدينة باتو، والتي يتم تصنيفها إلى فئات فقيرة وغير فقيرة. المتغيرات المستخدمة هي العمر والجنس والحالة الاجتماعية والمستوى التعليمي، وصحة رب الأسرة، بالإضافة إلى المتغير العديد من أفراد الأسرة، وحالة ملكية الإقامة، والوقود للطبخ، و ملكية الهاتف الخليوي. الطريقة المستخدمة هي الانحدار اللوجستي الثنائي و CART شجرة التصنيف والانحدار ستنتم مقارنة نتائج دقة التصنيف للطريقتين. وبناء على نتائج البحوث مع ثنائي طرق الانحدار اللوجستي والسلة يظهر أن المتغيرات التي تؤثر في نموذج التصنيف هي العمر والمهني رئيس المتغيرات متغير. قيمة تصنيف دقة من و الانحدار اللوجستي الثنائي نموذج كانت 94,69%، وهي نسبة أعلى من أن تصنيف نموذج السلة 89,36%. بحيث يكون أفضل نموذج للتنبؤ بعوامل رفاهية الأسرة في مدينة باتو هو طريقة الانحدار اللوجستي الثنائي.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kesejahteraan merupakan salah satu aspek penting yang harus terpenuhi bagi kehidupan masyarakat. Usaha yang dapat dilakukan untuk mencapai kesejahteraan dengan membangun dan menjaga kondisi stabilitas ekonomi dan sosial. Menurut Badan Pusat Statistik (BPS) bahwa kesejahteraan adalah keadaan yang merepresentasikan dimana rumah tangga dapat memenuhi kebutuhan jasmani dan rohani sesuai dengan tingkat hidup. Kesejahteraan dapat menunjukkan tolak ukur kualitas masyarakat secara keseluruhan dari sisi pangan, pendidikan, kesehatan, lingkungan, ketenagakerjaan, dan teknologi.

Istilah kesejahteraan erat kaitannya dengan problematika kemiskinan dan kesenjangan sosial ekonomi penduduk. Kemiskinan merupakan kondisi dimana seseorang atau rumah tangga mengalami kesulitan untuk memenuhi kebutuhan dasar, sementara lingkungan pendukungnya kurang memberikan peluang untuk meningkatkan kesejahteraan secara berkesinambungan (Cahyat dkk, 2007: 2). Menurut BPS, kemiskinan dipandang sebagai keadaan ketidakmampuan penduduk dari segi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur berdasarkan sisi pengeluaran. Sehingga dapat dikatakan bahwa penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan rendah atau dibawah Garis Kemiskinan (GK). Pengeluaran per kapita adalah biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama satu bulan dibagi dengan banyaknya anggota rumah

tangga. Sementara itu, garis kemiskinan merupakan jumlah nilai pengeluaran kebutuhan minimum untuk makanan dan bukan makanan, seperti perumahan, sandang, pendidikan, dan kesehatan per kapita per bulan.

Menurut Bappenas, status kesejahteraan penduduk dapat diukur berdasarkan proporsi pengeluaran kebutuhan rumah tangga. Selain menggunakan pengeluaran rumah tangga per kapita per bulan, status kesejahteraan dapat diukur dengan indikator penciri lain yang berpengaruh. Indikator kesejahteraan rumah tangga dapat dilihat dari hubungan dengan kepala rumah tangga, yaitu tingkat pendidikan, status perkawinan, usia, beban tanggungan, pendapatan, dan jenis kelamin dari kepala rumah tangga. Terdapat indikator lainnya seperti kepemilikan aset dan perumahan, status pekerjaan, kesehatan keluarga, kepemilikan hewan ternak, konsumsi makanan, dan partisipasi politik serta akses terhadap informasi (Suryadarma, 2005: 6-7).

Permasalahan kesejahteraan dan penanganan kemiskinan penduduk menjadi pekerjaan rumah bagi pemerintah, baik pemerintah pusat maupun daerah. Tak terkecuali bagi Pemerintah Kota Batu. Kota Batu merupakan salah satu kota yang berada di wilayah Provinsi Jawa Timur yang dikenal sebagai kota wisata. Sejak berdirinya pada tahun 2001, Kota Batu telah mampu melakukan perencanaan, pelaksanaan serta evaluasi proyek-proyek pembangunan ekonomi secara mandiri dan merata sehingga penduduk Kota Batu semakin meningkat kesejahteraannya. Kota Batu mempunyai luas wilayah 199,09 km² dan mempunyai penduduk sebanyak 203.997 jiwa yang tersebar di 3 kecamatan, yaitu Kecamatan Batu, Bumiaji, dan Junrejo. Sebagai daerah yang mengedepankan sektor pariwisata, banyak penduduk Kota Batu bekerja di sektor penunjang

pariwisata, misalnya pada akomodasi dan perdagangan, baik usaha kecil maupun besar, sektor rumah makan, tempat wisata, dan perhotelan. Disisi lain bahwa Kota Batu sebagian besar wilayahnya merupakan lahan pertanian dan perkebunan sehingga hal tersebut membuat mayoritas penduduk bekerja di sektor pertanian. Selain itu, tersebar pada beberapa sektor lapangan pekerjaan lainnya seperti penduduk yang bekerja pada sektor jasa, industri, keuangan, peternakan, bangunan, dan lain-lain.

Aspek kesejahteraan penduduk dalam rumah tangga dan pemerataan ekonomi masyarakat menjadi rencana pembangunan berkelanjutan bagi Pemerintah Kota Batu. Pencapaian positif dari pembangunan pada bidang kesejahteraan masyarakat terjadi setiap tahunnya. Hasilnya bahwa penduduk miskin di Kota Batu terus mengalami penurunan. Selama kurun waktu 3 tahun terakhir terjadi penurunan sebesar 0,5 persen dari tahun 2017 sampai 2019. Pada tahun 2017, presentase kemiskinan sebesar 4,31% dengan jumlah penduduk miskin 8770 jiwa dan tahun 2018 presentase kemiskinan pada angka 3,89% dengan jumlah 7980 penduduk miskin. Kemudian, tahun 2019 penduduk miskin di Kota Batu tercatat sebanyak 7890 jiwa dengan presentase kemiskinan sebesar 3,81%. Hasil tersebut menunjukkan kecenderungan mengikuti trend baik di tingkat provinsi Jawa Timur maupun nasional.

Sama halnya dengan presentase kemiskinan, nilai garis kemiskinan Kota Batu turut menunjukkan trend positif. Pada tahun 2018 garis kemiskinan di Kota Batu mengalami kenaikan dibandingkan tahun 2017 yaitu sebesar 10,07%, dari nilai 424.354 rupiah per kapita per bulan menjadi 467.073 rupiah per kapita per

bulan. Di tahun 2019, garis kemiskinan di Kota Batu juga mengalami kenaikan menjadi 501.016 rupiah per kapita per bulan (BPS Kota Batu, 2019: 151).

Banyak indikator yang dapat menentukan status kesejahteraan rumah tangga, oleh sebab itu dapat dicari indikator berpengaruh sebagai faktor yang mempengaruhi kesejahteraan rumah tangga di Kota Batu. Penentuan faktor-faktor tersebut dapat dilakukan dengan menggunakan metode pengklasifikasian. Kesejahteraan rumah tangga tersebut diklasifikasikan dalam 2 kategori, yaitu kategori miskin dan tidak miskin. Pengklasifikasian merupakan metode statistik yang dapat digunakan untuk mengelompokkan data yang disusun dengan cara yang sistematis. Metode yang dapat untuk mengklasifikasikan status kesejahteraan rumah tangga di Kota Batu adalah regresi logistik biner dan *Classification and Regression Tree* (CART).

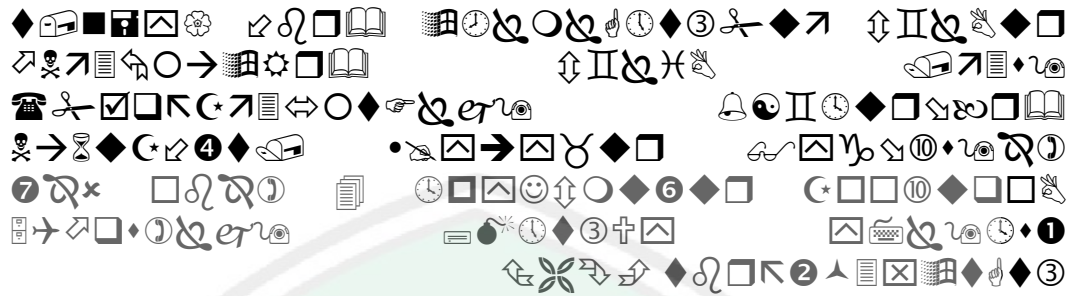
Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari pola hubungan antara variabel respon (Y) yang bersifat biner atau dikotomis dan beberapa variabel prediktor (X) yang berskala kategorik maupun rasio. Regresi logistik biner digunakan untuk pengklasifikasian sejumlah pengamatan dengan variabel respon biner dua kategori berdasarkan satu atau lebih variabel prediktor. Tujuan analisis regresi logistik biner adalah untuk mendapatkan model terbaik dan paling sederhana yang dapat menggambarkan hubungan antara variabel respon biner dengan variabel prediktor yang dipilih. Kelebihan penggunaan analisis regresi logistik adalah tidak berbias, informasi pengamatan lebih spesifik, terdapat estimasi atau pendugaan parameter dan prosesnya relatif lebih cepat.

Pengklasifikasian dapat pula dilakukan dengan menggunakan metode *Classification and Regression Tree* (CART) yaitu metode berstruktur pohon yang diperkenalkan oleh Breiman, Friedman, Olshen RA dan Stone CJ pada tahun 1984. Keempat ilmuwan tersebut memperkenalkan metode CART berupa metode pohon regresi dan pohon klasifikasi. Jika variabel respon yang diteliti berjenis kategorik maka CART akan menghasilkan pohon klasifikasi. Sedangkan, jika variabel terikat yang diteliti memiliki jenis data kontinu maka CART akan menghasilkan pohon regresi. Metode klasifikasi pada CART dilakukan dengan membuat pohon klasifikasi yang diperoleh melalui penyekatan atau pemilahan secara berulang terhadap sebuah data. Tujuan utamanya untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian. Kelebihan penggunaan metode CART adalah tidak ada asumsi distribusi pada variabel prediktornya yang harus terpenuhi, memudahkan dalam pengambilan keputusan dan menginterpretasi pada struktur data yang kompleks dan multivariabel, hasil klasifikasi akhir berbentuk sederhana dan efisien dalam mengklasifikasi data baru.

Berdasarkan penelitian sebelumnya yang berkaitan dengan kasus pengklasifikasian menggunakan metode regresi logistik biner dan CART diantaranya adalah penelitian yang dilakukan oleh Agung Waluyo, dkk (2015) dalam penelitian “Perbandingan Klasifikasi Nasabah Kredit Menggunakan Regresi Logistik Biner dan CART (*Classification and Regression Tree*)” diperoleh hasil bahwa menggunakan regresi logistik diketahui hanya pendapatan debitur yang mempengaruhi status kreditnya dengan akurasi klasifikasi mencapai 80%. Sedangkan, menggunakan CART terdapat beberapa variabel yang mempengaruhi status kredit dan ketepatan klasifikasi sebesar 80,9%. Sehingga

menunjukkan bahwa metode CART dalam mengklasifikasikan status kredit debitur lebih baik daripada regresi logistik biner. Penelitian yang dilakukan oleh Galih Mahalisa (2016) dalam jurnal “Metode Pohon Gabungan pada CART untuk Analisa Kesejahteraan Rumah Tangga di Banjarmasin” memberikan hasil bahwa ketepatan klasifikasi data sebesar 84,00% dengan hasil variabel paling berpengaruh adalah usia kepala rumah tangga, banyak anggota rumah tangga, aspek pendidikan, dan status pekerjaan. Selanjutnya, penelitian dalam jurnal “Metode *Bootstrap Aggregating* Regresi Logistik Biner untuk Ketepatan Klasifikasi Kesejahteraan Rumah Tangga di Kota Pati” oleh Ridha Ramandhani, dkk (2017) dengan hasil kajian bahwa variabel prediktor yang berpengaruh signifikan adalah jenis kelamin kepala rumah tangga, jumlah anggota rumah tangga, dan penguasaan telepon seluler dengan ketepatan klasifikasi sebesar 79,87%. Diterapkan pula oleh Marina Elviana Fobia, dkk (2020) dengan judul “Perbandingan Metode *Classification and Regression Tree* (CART) dan Metode Regresi Logistik Biner dalam Mengklasifikasi Status Wanita Bekerja di Kota Kupang” yang memberikan hasil bahwa analisis CART dan regresi logistik biner menghasilkan variabel yang berpengaruh yaitu umur dan status perkawinan. Perbedaan prediksi pada kedua metode yaitu prediksi wanita bekerja pada metode CART adalah wanita dengan umur 15-64 tahun dan memiliki status perkawinan menikah atau cerai. Sedangkan prediksi pada regresi logistik biner adalah wanita bekerja hanya wanita dengan umur 15-64 tahun. Wanita menikah atau cerai diprediksi tidak bekerja. Presentase ketepatan klasifikasi metode CART dan regresi logistik biner sebesar 65,2%.

Permasalahan kesejahteraan rumah tangga disinggung dalam Al-Qur'an, yaitu pada surah Ar-Ruum ayat 21:



Artinya: “Dan di antara tanda-tanda kekuasaan-Nya ialah Dia menciptakan untukmu isteri-isteri dari jenismu sendiri, supaya kamu cenderung dan merasa tenteram kepadanya, dan dijadikan-Nya diantaramu rasa kasih dan sayang. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda bagi kaum yang berfikir.”

Ayat tersebut dapat memberikan pengertian bahwa kesejahteraan rumah tangga dalam Islam adalah keluarga yang diliputi dengan adanya ketentraman lahir batin dan rasa kasih sayang bagi anggota keluarga. Tafsir Ibnu Katsir menjelaskan bahwa Allah menciptakan bagi kaum laki-laki, kaum wanita dari kelak mereka menjadi istri-istri, supaya cenderung dan merasa tenteram kepadanya. Semakna dengan kandungan dalam surat Al-A'raf ayat 189.

Berdasarkan uraian latar belakang, maka dilakukan penelitian mengenai klasifikasi terhadap faktor-faktor yang mempengaruhi kesejahteraan rumah tangga di Kota Batu. Menentukan faktor atau variabel yang paling berpengaruh terhadap status kesejahteraan penduduk berdasarkan hasil Survei Sosial dan Ekonomi Nasional (SUSENAS) tahun 2018 oleh BPS. Selain itu, pada penelitian ini dilakukan perbandingan nilai ketepatan klasifikasi untuk mendapatkan model yang terbaik di antara dua metode yang digunakan, yaitu analisis regresi logistik

biner dan *Classification and Regression Tree* (CART) dalam mengklasifikasikan status kesejahteraan rumah tangga di Kota Batu.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang penelitian, maka dapat diperoleh susunan rumusan masalah sebagai berikut:

1. Bagaimana model klasifikasi menggunakan metode regresi logistik biner dan *Classification and Regression Tree* (CART) pada klasifikasi status kesejahteraan rumah tangga di Kota Batu?
2. Berdasarkan model regresi logistik biner dan CART yang terbentuk, variabel apa saja yang berpengaruh terhadap model klasifikasi status kesejahteraan rumah tangga di Kota Batu?
3. Bagaimana perbandingan nilai ketepatan klasifikasi dari penggunaan metode regresi logistik biner dan *Classification and Regression Tree* (CART)?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Mengetahui model klasifikasi dari metode regresi logistik biner dan *Classification and Regression Tree* (CART) pada klasifikasi status kesejahteraan rumah tangga di Kota Batu.
2. Mengetahui variabel-variabel yang mempengaruhi model klasifikasi status status kesejahteraan rumah tangga di Kota Batu.

3. Mengetahui perbandingan nilai ketepatan klasifikasi penggunaan model regresi logistik biner dan metode *Classification and Regression Tree* (CART) sehingga dapat diperoleh model yang terbaik.

1.4 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah:

1. Manfaat Teoritis

Menambah pemahaman mengenai penerapan metode regresi logistik biner dan *Classification and Regression Tree* (CART) dalam menentukan model klasifikasi, nilai ketepatan klasifikasi, dan memperoleh faktor-faktor yang mempengaruhi kesejahteraan rumah tangga di Kota Batu.

2. Manfaat Praktis

Mengetahui penerapan model regresi logistik biner dan metode *Classification and Regression Tree* (CART) dalam contoh permasalahan sosial, perekonomian dan kependudukan, yaitu menentukan karakteristik atau faktor-faktor yang mempengaruhi kesejahteraan rumah tangga.

1.5 Batasan Masalah

Pada penelitian ini agar permasalahan yang akan diteliti menjadi lebih fokus dan lebih jelas maka permasalahan dibatasi oleh:

1. Data yang digunakan merupakan data mikro hasil Survei Sosial dan Ekonomi (SUSENAS) pada tahun 2018 oleh Badan Pusat Statistik, khusus untuk wilayah Kota Batu.

2. Variabel-variabel kesejahteraan rumah tangga yang digunakan adalah sebagai berikut:

a. Faktor kependudukan yaitu:

- Usia kepala rumah tangga,
- Jenis kelamin kepala rumah tangga,
- Status perkawinan kepala rumah tangga,
- Banyak anggota rumah tangga.

b. Faktor pendidikan yaitu:

- Tingkat pendidikan tertinggi kepala rumah tangga.

c. Faktor ketenagakerjaan:

- Status pekerjaan kepala rumah tangga.

d. Faktor Kesehatan yaitu:

- Kesehatan kepala rumah tangga dalam 1 bulan terakhir.

e. Faktor perumahan yaitu:

- Status kepemilikan tempat tinggal,
- Bahan bakar memasak.

f. Faktor teknologi informasi dan komunikasi yaitu:

- Kepemilikan telepon seluler (HP).

1.6 Sistematika Penulisan

Adapun sistematika penulisan yang digunakan dalam skripsi ini adalah sebagai berikut:

Bab I Pendahuluan

Pada bab ini akan diuraikan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

Bab II Kajian Pustaka

Pada bagian kajian pustaka menjelaskan mengenai teori yang mendasari pembahasan diantaranya teori analisis regresi logistik biner, metode estimasi parameter, pengujian multikolinieritas dan kesesuaian model, metode *Classification and Regression Tree* (CART), pembentukan pohon klasifikasi, uji hasil ketepatan klasifikasi dan definisi kesejahteraan dan kemiskinan, serta faktor-faktor yang mempengaruhi kesejahteraan rumah tangga.

Bab III Metode Penelitian

Pada bab ini akan disajikan mengenai pendekatan penelitian, sumber data, variabel yang digunakan dalam penelitian, langkah-langkah penelitian, dan *flowchart* atau diagram alir penelitian.

Bab IV Pembahasan

Bab ini merupakan bagian inti dari penulisan skripsi ini yang menjabarkan gambaran objek penelitian dengan analisis deskriptif, hasil pengolahan data dengan menggunakan metode regresi logistik biner dan *Classification and Regression Tree* (CART), dan menjelaskan perbandingan hasil ketepatan klasifikasi dari kedua metode tersebut.

Bab V Penutup

Pada bab penutup berisi tentang kesimpulan dan saran dari hasil yang diperoleh pada Bab IV yaitu analisis dan pembahasan.



BAB II

TINJAUAN PUSTAKA

2.1 Regresi Logistik Biner

Regresi logistik merupakan sebuah metode untuk membuat model yang memperkirakan ada atau tidaknya karakteristik berdasarkan hubungan antara variabel respon atau dependen (Y) dan variabel prediktor atau independen (X). Tujuan dari metode ini adalah untuk menentukan model yang terbaik, paling efektif dan model yang paling sesuai untuk menjelaskan hubungan antara variabel dependen (respon atau terikat) dan himpunan variabel independen (prediktor atau bebas). Model regresi logistik terdiri dari regresi logistik dengan respon biner, ordinal, dan multinomial (Hosmer, 2000: 1).

Regresi logistik biner adalah model yang digunakan untuk menganalisis hubungan antara satu variabel respon yang berupa kategorik dengan beberapa variabel prediktor. Variabel respon yang digunakan merupakan data yang bersifat *dichotomus* atau biner yaitu terdiri dari dua kategori, sedangkan variabel prediktornya terdiri dari satu atau lebih variabel yang dapat berupa jenis data rasio, interval, nominal, atau ordinal. Kategori variabel respon tersebut bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik. Variabel respon biner ini ditunjukkan oleh notasi variabel Y . Jika variabel respon Y bernilai 1 untuk “sukses” dan 0 untuk “gagal”. Sehingga diperoleh distribusi probabilitas dari variabel Y , untuk $y = 1$ adalah $P(Y = 1|X = x) = \pi(x)$ dan $y = 0$ adalah $P(Y = 0|X = x) = 1 - \pi(x)$ (Agresti, 2007: 68).

Jika terdapat data hasil pengamatan sebanyak p variabel prediktor, maka variabel prediktornya yaitu $x_1, x_2, x_3, \dots, x_p$. Sehingga peluang nilai sukses atau parameter $\pi(x)$ dari variabel prediktor x dapat diperoleh model umum persamaan regresi logistik sebagai berikut (Hosmer, 2000: 31-32):

$$\pi(x_i) = \frac{\exp^{g(x)}}{1 + \exp^{g(x)}} \quad (2.1)$$

dengan $g(x_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$.

Maka diperoleh bentuk umum persamaan regresi logistik dengan melibatkan beberapa variabel prediktor x dapat diformulasikan sebagai berikut:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \quad (2.2)$$

Keterangan:

$\pi(x_i)$: Peluang kejadian dengan nilai probabilitas $0 \leq \pi(x_i) \leq 1$

x_i : Variabel prediktor ke- i ($i = 1, 2, 3, \dots$)

β_0 : Konstanta

β_i : Koefisien variabel prediktor ke- i ($i = 1, 2, 3, \dots, p$)

Regresi logistik biner akan menunjukkan pola hubungan antara variabel x dengan $\pi(x_i)$. Nilai $\pi(x_i)$ adalah nilai probabilitas kejadian yang diperoleh dari kejadian yang disebabkan variabel x dengan nilai yang diperoleh dalam selang $0 \leq \pi(x_i) \leq 1$. Maka nilai harapan dan varians Y yang diperoleh masing-masing adalah $\pi(x_i)$ dan $1 - \pi(x_i)$. Model umum regresi logistik dengan fungsi $\pi(x_i)$ merupakan fungsi non linier maka untuk membuatnya menjadi fungsi linier agar dapat dilihat hubungan antara variabel respon dan variabel prediktor harus dilakukan dengan transformasi logit. Bentuk logit dari fungsi $\pi(x_i)$ diperoleh sebagai berikut:

$$\begin{aligned}
\pi(x_i) &= \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}} \\
[\pi(x_i)][1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}] &= e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\pi(x_i) + \pi(x_i) e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} &= e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\pi(x_i) &= e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} - \pi(x_i) e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\pi(x_i) &= [1 - \pi(x_i)] e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\frac{\pi(x_i)}{1 - \pi(x_i)} &= e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\ln \left[\frac{\pi(x_i)}{1 - \pi(x_i)} \right] &= \ln e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \\
\ln \left[\frac{\pi(x_i)}{1 - \pi(x_i)} \right] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \\
g(x_i) = \ln \left[\frac{\pi(x_i)}{1 - \pi(x_i)} \right] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \\
&= \beta_0 + \sum_{i=1}^p x_i \beta_i \\
&= X\beta
\end{aligned} \tag{2.3}$$

2.1.1 Estimasi Parameter

Pendekatan yang digunakan untuk estimasi parameter model regresi logistik adalah metode *Maximum Likelihood Estimation* (MLE). Metode ini digunakan karena jika menggunakan metode kuadrat terkecil (*least square*) seperti pada regresi linear biasanya maka akan menghasilkan penaksir parameter yang tidak lagi mempunyai sifat statistik unbiased dan tidak memiliki varians minimum. Terdapat asumsi homoskedastisitas yang tidak mungkin dipenuhi oleh variabel biner. Hal tersebut dikarenakan pada regresi logistik biner variabel respon (y) mengikuti sebaran distribusi *Bernoulli* sehingga nilai variansnya berubah bergantung pada nilai peluang suksesnya.

Fungsi *likelihood* yang didapat adalah fungsi peluang pengamatan untuk memperoleh estimasi parameter yang tidak diketahui. Hasil estimasi parameter yang tidak diketahui merupakan hasil dari memaksimumkan nilai peluang dari data pengamatan. Metode MLE menaksir koefisien atau parameter β dengan cara memaksimumkan fungsi *likelihood* tersebut (Hosmer, 2000: 8-9).

Misalkan y_i menyatakan jumlah sukses dalam sampel ukuran N dengan probabilitas sukses (π_i) dan probabilitas gagal ($1 - (\pi_i)$) dengan y_i mengikuti sebaran distribusi *Bernoulli*. Jika x_i dan y_i adalah pasangan variabel prediktor dan respon pada pengamatan ke- i , dengan $i = 1, 2, 3, \dots, N$. Maka fungsi probabilitasnya adalah sebagai berikut:

$$\begin{aligned}\pi(Y_i = y_i) &= \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \\ y_i &= 0, 1\end{aligned}\quad (2.4)$$

Jika y_i diasumsikan independen identik (saling bebas) maka dari persamaan (2.5) diperoleh fungsi *likelihood* untuk model regresi logistik biner sebagai berikut:

$$\begin{aligned}\ell(\beta; y_i) &= \prod_{i=1}^N \pi(Y_i = y_i) \\ &= \prod_{i=1}^N \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}\end{aligned}\quad (2.5)$$

Selanjutnya, persamaan (2.6) dimaksimumkan dengan bentuk fungsi log *likelihood* seperti berikut:

$$\begin{aligned}L(\beta; y_i) &= \ln \ell(\beta; y_i) = \ln \prod_{i=1}^N \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \\ &= \sum_{i=1}^N y_i \ln \pi(x_i) + (1 - y_i) \ln(1 - \pi(x_i)) \\ &= \sum_{i=1}^N y_i \ln \left(\frac{e^{g(x_i)}}{1 + e^{g(x_i)}} \right) + (1 - y_i) \ln \left(1 - \left(\frac{e^{g(x_i)}}{1 + e^{g(x_i)}} \right) \right) \\ &= \sum_{i=1}^N y_i [\ln(e^{g(x_i)}) - \ln(1 + e^{g(x_i)})] + (1 - y_i) \ln \left(\frac{1}{1 + e^{g(x_i)}} \right)\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^N y_i [\ln(e^{g(x_i)}) - \ln(1 + e^{g(x_i)})] + (1 - y_i) [\ln 1 - \ln(1 + e^{g(x_i)})] \\
&= \sum_{i=1}^N y_i [\ln(e^{g(x_i)}) - \ln(1 + e^{g(x_i)})] - (1 - y_i) [\ln(1 + e^{g(x_i)})] \\
&= \sum_{i=1}^N [y_i \ln(e^{g(x_i)}) - y_i \ln(1 + e^{g(x_i)}) - \ln(1 + e^{g(x_i)}) + \\
&\quad y_i \ln(1 + e^{g(x_i)})] \\
&= \sum_{i=1}^N [y_i \ln(e^{g(x_i)}) - \ln(1 + e^{g(x_i)})] \\
&= \sum_{i=1}^N (y_i g(x_i) - \ln(1 + e^{g(x_i)}))
\end{aligned}$$

Karena $g(x_i) = \ln \left[\frac{\pi(x_i)}{1-\pi(x_i)} \right] = x_i \beta$, dengan $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_p)$ dan

$x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{ip})$ merupakan nilai-nilai variabel prediktor ke-1 sampai

ke- p dari pengamatan ke- i sehingga $g(x_i) = \ln \left[\frac{\pi(x_i)}{1-\pi(x_i)} \right] = \sum_{j=1}^p x_{ij} \beta_j$, maka:

$$\begin{aligned}
L(\beta) &= \sum_{i=1}^N (y_i x_i \beta - \ln(1 + e^{x_i \beta})) \\
&= \sum_{i=1}^N y_i x_i \beta - \sum_{i=1}^N \ln(1 + e^{x_i \beta}) \\
&= \sum_{i=1}^N y_i \left(\sum_{j=1}^p x_{ij} \beta_j \right) - \sum_{i=1}^N \ln \left(1 + e^{\sum_{j=1}^p x_{ij} \beta_j} \right) \tag{2.6}
\end{aligned}$$

Selanjutnya persamaan (2.6) diturunkan terhadap $(\beta_0, \beta_1, \beta_2, \dots, \beta_p)$.

Penurunan tersebut dilakukan dengan menyamadengkan nol sehingga dapat diperoleh hasil sebagai berikut:

Turunan fungsi log *likelihood* terhadap β_0 :

$$\frac{\partial L(\beta)}{\partial \beta_0} = \frac{\partial \left(\sum_{i=1}^N y_i x_{i0} \beta_0 - \sum_{i=1}^N \ln \left(1 + e^{\sum_{j=1}^p x_{ij} \beta_j} \right) \right)}{\partial \beta_0} = 0$$

$$\sum_{i=1}^N y_i x_{i0} - \sum_{i=1}^N x_{i0} \frac{e^{\sum_{j=1}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=1}^p x_{ij} \beta_j}} = 0$$

$$\sum_{i=1}^N y_i x_{i0} - \sum_{i=1}^N x_{i0} \pi(x_i) = 0$$

$$\sum_{i=1}^N y_i - \sum_{i=1}^N \pi(x_i) = 0$$

Turunan fungsi log *likelihood* terhadap β_1 :

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1} = \frac{\partial \left(\sum_{i=1}^N y_i x_{i1} \beta_1 - \sum_{i=1}^N \ln \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right) \right)}{\partial \beta_1} = 0$$

$$\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} = 0$$

$$\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \pi(x_i) = 0$$

Turunan fungsi log *likelihood* terhadap β_2 :

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_2} = \frac{\partial \left(\sum_{i=1}^N y_i x_{i2} \beta_2 - \sum_{i=1}^N \ln \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right) \right)}{\partial \beta_2} = 0$$

$$\sum_{i=1}^N y_i x_{i2} - \sum_{i=1}^N x_{i2} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} = 0$$

$$\sum_{i=1}^N y_i x_{i2} - \sum_{i=1}^N x_{i2} \pi(x_i) = 0$$

⋮

Turunan fungsi log *likelihood* terhadap β_p :

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_p} = \frac{\partial \left(\sum_{i=1}^N y_i x_{ip} \beta_p - \sum_{i=1}^N \ln \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right) \right)}{\partial \beta_p} = 0$$

$$\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} = 0$$

$$\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \pi(x_i) = 0$$

Berdasarkan turunan pertama dari $L(\boldsymbol{\beta})$ terhadap $(\beta_0, \beta_1, \beta_2, \dots, \beta_p)$,

maka dapat diperoleh persamaan umum:

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_p} = \sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \pi(x_i)$$

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_p} = \sum_{i=1}^N x_{ip} [y_i - \pi(x_i)] \quad (2.7)$$

Jika persamaan (2.7) dinyatakan dalam bentuk matriks maka diperoleh:

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix} \begin{bmatrix} y_1 - \pi(x_1) \\ y_2 - \pi(x_2) \\ \vdots \\ y_N - \pi(x_N) \end{bmatrix} = \mathbf{X}^T (\mathbf{Y} - \boldsymbol{\pi}) \quad (2.8)$$

Selanjutnya dilakukan penurunan kedua dari fungsi log *likelihood* terhadap masing-masing parameter sebagai berikut:

Turunan kedua fungsi log *likelihood* β_0 :

$$\begin{aligned} \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \beta_0} &= \frac{\partial(\partial L(\boldsymbol{\beta}))}{\partial \beta_0 \beta_0} = \frac{\partial(\sum_{i=1}^N y_i - \sum_{i=1}^N \pi(x_i))}{\partial \beta_0} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \beta_0} &= \frac{\partial \left(\sum_{i=1}^N y_i - \sum_{i=1}^N \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right) \right)}{\partial \beta_0} \\ &= - \sum_{i=1}^N \left[\frac{e^{\sum_{j=0}^p x_{ij} \beta_j} \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right) - \left(e^{\sum_{j=0}^p x_{ij} \beta_j} \right)^2}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right)^2} \right] \\ &= - \sum_{i=1}^N \left[\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right)} - \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right)^2 \right] \\ &= - \sum_{i=1}^N \left[\pi(x_i) - \left(\pi(x_i) \right)^2 \right] \\ &= - \sum_{i=1}^N \left[\pi(x_i) (1 - \pi(x_i)) \right] \\ &\vdots \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \beta_p} &= \frac{\partial(\partial L(\boldsymbol{\beta}))}{\partial \beta_0 \beta_p} = \frac{\partial(\sum_{i=1}^N y_i - \sum_{i=1}^N \pi(x_i))}{\partial \beta_p} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_0 \beta_p} &= \frac{\partial \left(\sum_{i=1}^N y_i - \sum_{i=1}^N \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right) \right)}{\partial \beta_p} \\ &= - \sum_{i=1}^N \left[\frac{x_{ip} e^{\sum_{j=0}^p x_{ij} \beta_j} \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right) - x_{ip} \left(e^{\sum_{j=0}^p x_{ij} \beta_j} \right)^2}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j} \right)^2} \right] \end{aligned}$$

$$\begin{aligned}
&= -\sum_{i=1}^N \left[x_{ip} \frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{\left(1+e^{\sum_{j=0}^p x_{ij}\beta_j}\right)} - x_{ip} \left(\frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{1+e^{\sum_{j=0}^p x_{ij}\beta_j}} \right)^2 \right] \\
&= -\sum_{i=1}^N \left[x_{ip} \pi(x_i) - x_{ip} (\pi(x_i))^2 \right] \\
&= -\sum_{i=1}^N x_{ip} [\pi(x_i)(1 - \pi(x_i))]
\end{aligned}$$

Turunan kedua fungsi log *likelihood* β_1 :

$$\begin{aligned}
\frac{\partial^2 L(\beta)}{\partial \beta_1 \beta_0} &= \frac{\partial(\partial L(\beta))}{\partial \beta_1 \beta_0} = \frac{\partial(\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \pi(x_i))}{\partial \beta_0} \\
\frac{\partial^2 L(\beta)}{\partial \beta_1 \beta_0} &= \frac{\partial \left(\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \left(\frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{1+e^{\sum_{j=0}^p x_{ij}\beta_j}} \right) \right)}{\partial \beta_0} \\
&= -\sum_{i=1}^N \left[\frac{x_{i1} x_{i0} e^{\sum_{j=0}^p x_{ij}\beta_j} \left(1+e^{\sum_{j=0}^p x_{ij}\beta_j} \right) - x_{i1} x_{i0} \left(e^{\sum_{j=0}^p x_{ij}\beta_j} \right)^2}{\left(1+e^{\sum_{j=0}^p x_{ij}\beta_j} \right)^2} \right] \\
&= -\sum_{i=1}^N \left[x_{i1} x_{i0} \frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{\left(1+e^{\sum_{j=0}^p x_{ij}\beta_j} \right)} - x_{i1} x_{i0} \left(\frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{1+e^{\sum_{j=0}^p x_{ij}\beta_j}} \right)^2 \right] \\
&= -\sum_{i=1}^N \left[x_{i1} x_{i0} \pi(x_i) - x_{i1} x_{i0} (\pi(x_i))^2 \right] \\
&= -\sum_{i=1}^N x_{i1} x_{i0} [\pi(x_i)(1 - \pi(x_i))] \\
&\vdots \\
\frac{\partial^2 L(\beta)}{\partial \beta_1 \beta_p} &= \frac{\partial(\partial L(\beta))}{\partial \beta_1 \beta_p} = \frac{\partial(\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \pi(x_i))}{\partial \beta_p} \\
\frac{\partial^2 L(\beta)}{\partial \beta_1 \beta_p} &= \frac{\partial \left(\sum_{i=1}^N y_i x_{i1} - \sum_{i=1}^N x_{i1} \left(\frac{e^{\sum_{j=0}^p x_{ij}\beta_j}}{1+e^{\sum_{j=0}^p x_{ij}\beta_j}} \right) \right)}{\partial \beta_p} \\
&= -\sum_{i=1}^N \left[\frac{x_{i1} x_{ip} e^{\sum_{j=0}^p x_{ij}\beta_j} \left(1+e^{\sum_{j=0}^p x_{ij}\beta_j} \right) - x_{i1} x_{ip} \left(e^{\sum_{j=0}^p x_{ij}\beta_j} \right)^2}{\left(1+e^{\sum_{j=0}^p x_{ij}\beta_j} \right)^2} \right]
\end{aligned}$$

$$\begin{aligned}
&= -\sum_{i=1}^N \left[x_{i1} x_{ip} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right)} - x_{i1} x_{ip} \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right)^2 \right] \\
&= -\sum_{i=1}^N \left[x_{i1} x_{ip} \pi(x_i) - x_{i1} x_{ip} (\pi(x_i))^2 \right] \\
&= -\sum_{i=1}^N x_{i1} x_{ip} [\pi(x_i)(1 - \pi(x_i))]
\end{aligned}$$

Turunan kedua fungsi log *likelihood* β_p :

$$\begin{aligned}
\frac{\partial^2 L(\beta)}{\partial \beta_p \beta_0} &= \frac{\partial(\partial L(\beta))}{\partial \beta_p \beta_0} = \frac{\partial(\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \pi(x_i))}{\partial \beta_0} \\
\frac{\partial^2 L(\beta)}{\partial \beta_p \beta_0} &= \frac{\partial \left(\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right) \right)}{\partial \beta_0} \\
&= -\sum_{i=1}^N \left[\frac{x_{ip} x_{i0} e^{\sum_{j=0}^p x_{ij} \beta_j} \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right) - x_{ip} x_{i0} \left(e^{\sum_{j=0}^p x_{ij} \beta_j}\right)^2}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right)^2} \right] \\
&= -\sum_{i=1}^N \left[x_{ip} x_{i0} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right)} - x_{ip} x_{i0} \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right)^2 \right] \\
&= -\sum_{i=1}^N \left[x_{ip} x_{i0} \pi(x_i) - x_{ip} x_{i0} (\pi(x_i))^2 \right] \\
&= -\sum_{i=1}^N x_{ip} x_{i0} [\pi(x_i)(1 - \pi(x_i))] \\
&\vdots \\
\frac{\partial^2 L(\beta)}{\partial \beta_p \beta_p} &= \frac{\partial(\partial L(\beta))}{\partial \beta_p \beta_p} = \frac{\partial(\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \pi(x_i))}{\partial \beta_p} \\
\frac{\partial^2 L(\beta)}{\partial \beta_p \beta_p} &= \frac{\partial \left(\sum_{i=1}^N y_i x_{ip} - \sum_{i=1}^N x_{ip} \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right) \right)}{\partial \beta_p} \\
&= -\sum_{i=1}^N \left[\frac{x_{ip} x_{ip} e^{\sum_{j=0}^p x_{ij} \beta_j} \left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right) - x_{ip} x_{ip} \left(e^{\sum_{j=0}^p x_{ij} \beta_j}\right)^2}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right)^2} \right]
\end{aligned}$$

$$\begin{aligned}
&= -\sum_{i=1}^N \left[x_{ip} x_{ip} \frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{\left(1 + e^{\sum_{j=0}^p x_{ij} \beta_j}\right)} - x_{ip} x_{ip} \left(\frac{e^{\sum_{j=0}^p x_{ij} \beta_j}}{1 + e^{\sum_{j=0}^p x_{ij} \beta_j}} \right)^2 \right] \\
&= -\sum_{i=1}^N \left[x_{ip} x_{ip} \pi(x_i) - x_{ip} x_{ip} (\pi(x_i))^2 \right] \\
&= -\sum_{i=1}^N x_{ip} x_{ip} [\pi(x_i)(1 - \pi(x_i))]
\end{aligned}$$

Secara umum turunan kedua fungsi log *likelihood* dapat diperoleh dalam bentuk persamaan berikut:

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} = -\sum_{i=1}^N x_{ia} x_{ib} [\pi(x_i)(1 - \pi(x_i))]; \quad a, b = 0, 1, 2, \dots, p \quad (2.9)$$

Jika persamaan (2.9) dinyatakan dalam bentuk matriks maka diperoleh:

$$\begin{aligned}
&\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix} \cdot \begin{bmatrix} \pi(x_1)(1 - \pi(x_1)) & 0 & \dots & 0 \\ 0 & \pi(x_2)(1 - \pi(x_2)) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi(x_i)(1 - \pi(x_i)) \end{bmatrix} \\
&\begin{bmatrix} 1 & x_{11} & \dots & x_{N1} \\ 1 & x_{12} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1p} & \dots & x_{Np} \end{bmatrix} = \mathbf{X}^T \mathbf{V} \mathbf{X} \quad (2.10)
\end{aligned}$$

Estimator $\hat{\boldsymbol{\beta}}$ diperoleh dengan menyelesaikan hasil turunan fungsi log *likelihood* terhadap masing-masing parameter $\boldsymbol{\beta}$ pada persamaan (2.7) dan (2.9). Pada persamaan (2.7) dan (2.9) dalam bentuk implisit sehingga diperlukan metode numerik untuk membantu menyelesaikan. Penyelesaian dapat dilakukan dengan metode *Newton Raphson*. Metode ini menyelesaikan persamaan nonlinier secara iterasi atau algoritma perulangan untuk menentukan lokasi yang memkasimalkan suatu fungsi seperti persamaan log *likelihood* (Agresti, 2007: 88). Algoritma iterasi *Newton Raphson* untuk memperoleh nilai $\hat{\boldsymbol{\beta}}$ melalui langkah-langkah berikut:

1. Menentukan nilai estimasi awal ($\beta_{(t)}$) untuk β dengan metode kuadrat terkecil (OLS)

$$\beta_{(t)} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

2. Menentukan $X^T(Y - \pi)$ dan X^TVX , kemudian mencari invers dari X^TVX .
3. Menghitung estimasi baru pada iterasi ke- t

$$\hat{\beta}_{(t+1)} = \beta_{(t)} + (X^TVX)^{-1}(X^T(Y - \pi))$$

4. Iterasi berhenti jika telah diperoleh nilai estimasi yang konvergen

$$\hat{\beta}_{(t+1)} \cong \beta_{(t)} \text{ atau } \|\hat{\beta}_{(t+1)} - \beta_{(t)}\| < \varepsilon$$

dengan ε adalah tingkat ketelitian dalam bilangan positif.

2.1.2 Uji Signifikansi Parameter

Pengujian signifikansi terhadap parameter dilakukan untuk memeriksa peranan variabel prediktor dalam model terhadap pengaruhnya ke variabel respon. Pendekatan ini untuk menguji signifikansi koefisien variabel dalam model yang diperoleh. Uji signifikansi parameter ini dilakukan melalui dua tahap, yaitu uji signifikansi parsial dan uji signifikansi serentak. Uji parsial dilakukan dengan statistik uji Wald (W) dan uji serentak melalui uji G atau uji rasio *likelihood*.

Uji signifikansi parameter secara parsial dilakukan dengan menggunakan uji Wald. Uji Wald ini seperti uji t pada regresi linier biasa yang digunakan untuk memeriksa pengaruh koefisien regresi secara terpisah atau satu-satu pada model.

Pada uji menggunakan hipotesis sebagai berikut:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0, \quad \text{dengan } i = 1, 2, 3, \dots, p$$

Menurut Hosmer dan Lemeshow (2000: 16), rumus untuk statistik uji Wald didefinisikan sebagai berikut:

$$W_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (2.12)$$

Keterangan:

Statistik uji W mengikuti sebaran distribusi normal dengan taraf signifikansi α maka daerah pengambilan keputusan adalah jika $|W_i| > Z_{\alpha/2}$ atau $p\text{-value} < \alpha$, maka tolak H_0 .

Uji serentak dilakukan untuk mengetahui signifikansi parameter terhadap variabel respon secara serentak atau menyeluruh. Pengujian signifikansi parameter menggunakan statistik uji G dan statistik uji ini mengikuti distribusi *Chi-Square* (Hosmer, 2000: 13). Statistik uji G adalah uji rasio *likelihood* maksimum digunakan untuk menguji pengaruh variabel prediktor dalam model secara bersama-sama terhadap variabel respon. Rumus untuk statistik uji G diformulasikan sebagai berikut:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}(x_i)^{y_i} (1 - \hat{\pi}(x_i))^{1-y_i}} \right] = -2 \ln \left[\frac{L_0}{L_p} \right] \quad (2.11)$$

Keterangan:

L_0 : *Likelihood* model tidak ada variabel prediktor (hanya β_0)

L_p : *Likelihood* model terdiri dari p variabel prediktor

Hipotesis yang digunakan:

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

H_1 : Minimal terdapat satu $\beta_i \neq 0$, dengan $i = 1, 2, 3, \dots, p$

Statistik uji G mengikuti sebaran distribusi *Chi-Square* (χ^2) dengan derajat kebebasan k , dimana k merupakan banyaknya variabel prediktor. Jika asumsi taraf signifikansi α , maka daerah pengambilan keputusan adalah jika nilai $G > \chi^2_{(\alpha,k)}$ maka tolak H_0 atau menggunakan jika $p\text{-value} < \alpha$. Sebaliknya, jika nilai $G < \chi^2_{(\alpha,k)}$ maka gagal tolak H_0 atau $p\text{-value} > \alpha$.

2.1.3 Uji Multikolinieritas

Multikolinieritas adalah kondisi ketika terdapat dua atau lebih variabel bebas yang digunakan dalam regresi saling berkorelasi. Uji multikolinieritas bertujuan untuk menguji bahwa model regresi yang terbentuk terdapat atau tidaknya korelasi sempurna atau tinggi antar variabel bebas. Jika model regresi memiliki nilai korelasi tinggi atau sempurna, maka model regresi tersebut dapat dinyatakan mengandung gejala multikolinieritas (Suliyanto, 2011: 81).

Pada regresi logistik terdapat asumsi bahwa tidak boleh ada multikolinieritas. Karena multikolinieritas *standard error* dari koefisien regresi akan membesar sehingga kemungkinan uji Wald akan menghasilkan variabel prediktor yang akan tidak signifikan. Salah satu cara yang dapat digunakan untuk mendeteksi adanya multikolinieritas dapat diketahui melalui nilai VIF (*Variance Inflation Error Factor*) untuk masing-masing variabel prediktor. Jika nilai VIF yang diperoleh lebih dari atau sama dengan 10, maka hal tersebut menunjukkan adanya multikolinieritas dalam data yang digunakan. Oleh karena itu, uji multikolinieritas ini diharapkan nilai VIF kurang dari nilai 10 (Neter dkk, 1990: 408).

$$VIF_i = \frac{1}{1 - R_i^2} \quad (2.13)$$

R_i^2 adalah koefisien determinasi berganda dari variabel prediktor X_i dengan seluruh variabel prediktor lainnya.

2.1.4 Uji Kesesuaian Model

Pengujian kesesuaian model dilakukan untuk mengetahui adanya perbedaan yang signifikan antara hasil pengamatan dengan nilai hasil prediksi dari model. Pengujian kelayakan model ini dilakukan dengan menggunakan uji Hosmer dan Lemeshow. Hipotesis yang digunakan pada uji ini sebagai berikut:

H_0 : Tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model (model yang didapat sesuai).

H_1 : Terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model (model yang didapat tidak sesuai).

Daerah keputusannya adalah jika $\hat{C} > \chi_{(g-2,k)}^2$ atau $p\text{-value} < \alpha$, maka tolak

H_0 . Statistik uji yang digunakan adalah:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.14)$$

Keterangan:

O_k : Observasi pada grup ke- k

$\bar{\pi}_k$: Rata-rata taksiran peluang ke- k

n'_k : Banyaknya observasi pada grup ke- k

g : Banyaknya grup

2.1.5 Interpretasi Koefisien Regresi

Interpretasi koefisien dilakukan untuk menentukan hubungan fungsional antara variabel respon dan variabel prediktor serta mendefinisikan perubahan yang terjadi pada variabel respon yang disebabkan oleh variabel prediktor. Interpretasi

koefisien untuk model regresi logistik biner dapat dilakukan dengan menggunakan nilai *odds ratio*. *Odds ratio* merupakan nilai kecenderungan antara satu kategori dengan kategori yang lain pada variabel prediktor yang kualitatif. *Odds* merupakan rasio peluang terjadinya kejadian sukses dengan kejadian gagal dari variabel respon. Misalkan peluang sukses π dan peluang gagal $1 - \pi$, maka nilai *odds* didefinisikan sebagai berikut (Hosmer, 2000: 28):

$$Odds = \frac{\pi}{(1 - \pi)} \quad (2.15)$$

Odds ratio mengindikasikan seberapa mungkin munculnya kejadian sukses pada suatu kelompok dibandingkan dengan kelompok lainnya. Model regresi logistik dengan p variabel prediktor, maka nilai *odds ratio* dapat dihitung untuk masing-masing variabel prediktor yang didefinisikan sebagai berikut:

$$\psi = \exp(\hat{\beta}_i) \quad (2.16)$$

2.2 Classification and Regression Tree (CART)

Classification and Regression Tree (CART) adalah metode statistik baru yang digunakan untuk keperluan analisis data dengan berstruktur pohon. Metode ini pertama kali dikenalkan oleh Leo Breiman, J. H. Friedman, Richard A. Olshen, dan Charles J. Stone pada tahun 1984. CART menggunakan pohon keputusan untuk mempresentasikan hasil klasifikasi atau prediksi data. Secara otomatis mencari hubungan penting melalui serangkaian pertanyaan “ya atau tidak”. CART sering digunakan untuk mengolah data kompleks dengan banyak variabel prediktor (Steinberg, 2001: 1).

Tujuan utama penggunaan metode CART adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu proses pengklasifikasian.

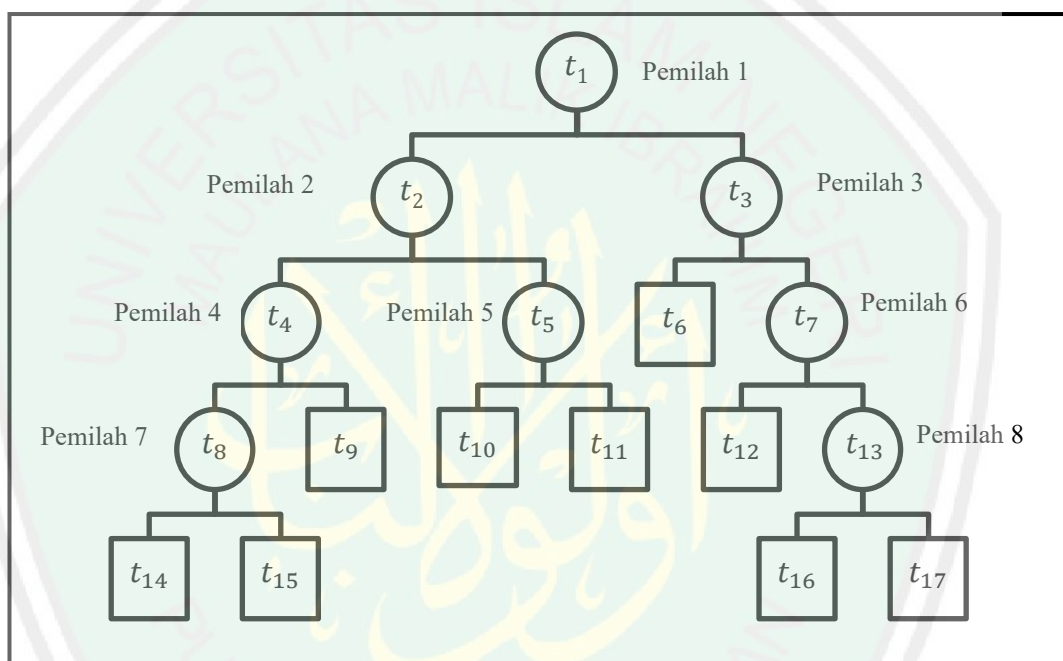
CART terdiri dari dua analisis, yaitu *classification tree* dan *regression tree*. CART akan menghasilkan *classification tree* atau pohon klasifikasi apabila variabel respon yang diteliti berupa data berskala kategorik, sedangkan apabila variabel respon berupa data berskala kontinu maka CART menghasilkan *regression tree* atau pohon regresi (Yohannes & Hoddinott, 1999: 3).

Metode CART dalam prosesnya untuk menghasilkan pohon klasifikasi dikenal dengan istilah *binary recursive partitioning* (pemilahan rekursif biner). Metode ini melakukan pemilahan atau pemecahan data secara berulang-ulang sehingga menghasilkan pohon yang tersusun atas banyak *node* (simpul), oleh karena itu disebut dengan proses rekursif. Proses pemilihan ini terus dilakukan sampai tidak ada kesempatan lagi untuk melakukan pemilahan selanjutnya. Proses CART disebut biner karena setiap *node* atau simpul induk akan mengalami pemilahan menghasilkan tepat dua simpul anak. Sedangkan, *partitioning* adalah proses pemecahan sampel menjadi bagian-bagian atau partisi-partisi yang lebih kecil dari simpul sebelumnya (Lewis, 2000: 4).

Penggunaan metode CART memberikan keuntungan pada analisis dan hasilnya, antara lain (Andriyashin, 2005: 7):

1. CART merupakan metode statistika non parametrik, sehingga tidak perlu memenuhi asumsi distribusi dan uji hioptesis tertentu.
2. Tidak memerlukan variabel untuk dipilih terlebih dahulu, artinya bahwa variabel-variabel yang diberikan pada sampel *learning* atau *training* CART akan otomatis memilih yang paling signifikan dan modelnya akan memilih pemisahan yang terbaik.

3. Sangat efisien dalam hal perhitungan, susunan CART fleksibel untuk menghitung semua data dan melakukannya dengan cepat.
4. Mampu menangani outlier dengan baik, umumnya algoritma pemisahan akan memisahkan outlier pada suatu simpul atau beberapa simpul tertentu.
5. Mampu menangani dataset dengan struktur besar dan kompleks.
6. Dapat menganalisa sembarang kombinasi jenis data, baik data kontinu atau kategorik.



Gambar 2.1 Struktur Pohon Klasifikasi

Lewis (2000: 6) menjelaskan bahwa proses CART akan membentuk pohon klasifikasi melalui empat langkah utama. Langkah pertama adalah proses pembentukan sebuah pohon klasifikasi yang terdiri dari proses pemecahan simpul induk menjadi dua simpul anak melalui aturan pemilahan tertentu yang dilakukan secara berulang-ulang. Langkah kedua adalah proses pemberhentian pembentukan pohon klasifikasi, tahap ini akan menghasilkan atau menentukan pohon klasifikasi terakhir atau maksimal (T_{max}). Langkah ketiga adalah proses pemangkasan atau pemotongan atau pemangkasan pohon menjadi pohon yang lebih kecil yang

disebut dengan proses *pruning tree*. Pada proses ini akan menghasilkan pohon klasifikasi optimal. Langkah yang terakhir adalah menentukan ketepatan atau akurasi dari prediksi pohon klasifikasi yang paling optimal.

2.2.1 Pembentukan Pohon Klasifikasi

Pembentukan pohon klasifikasi diawali dengan menentukan variabel pemilah dan nilai dari variabel tersebut (nilai *threshold*) untuk dijadikan pemilah tiap simpul. Proses pembentukan pohon klasifikasi dibutuhkan data *training* sampel L yang terdiri dari N pengamatan. Menurut Breiman dkk, pembentukan pohon klasifikasi terdiri atas tiga tahapan, sebagai berikut:

A. Pemilahan Pemilah

Pemilahan pemilah pada setiap simpul dilakukan untuk mendapatkan pemilah yang dapat menghasilkan simpul dengan tingkat kehomogenan nilai variabel respon paling maksimum. Karena pemilahan pemilah dilakukan dengan menggunakan data yang sifatnya masih heterogen. Sehingga hasil pemilahan pemilah ini untuk mengurangi keheterogenan pada simpul induk dan memaksimumkan ukuran kehomogenan pada setiap simpul induk. Proses pemilahan pemilah pada masing-masing simpul induk berdasarkan pada kriteria *goodness of split* (pemilah terbaik).

Pengukuran tingkat keheterogenan suatu kelas dari suatu simpul pada pohon klasifikasi menggunakan metode *impurity measure* $i(t)$. Metode ini dapat menemukan fungsi pemilah paling optimal. Fungsi *impurity measure* yang dapat digunakan adalah indeks Gini. Kaidah ini paling sering digunakan dalam penelitian analisis pohon klasifikasi karena

akan selalu memilah kelas dengan anggota paling besar atau kelas yang paling penting dalam simpul terlebih dahulu. Data *impurity* didefinisikan menggunakan ukuran varians. Misalkan 1 adalah semua pengamatan pada simpul t kelas ke- j dan 0 untuk lainnya. Estimasi varian contoh untuk simpul t pengamatan sebagai berikut (Breiman dkk, 1998: 103-104):

$$p(j|t)(1 - p(j|t)) \quad (2.17)$$

Indeks Gini pada simpul t kelas ke- j dapat didefinisikan sebagai berikut:

$$\begin{aligned} i(t) &= \sum_{j=1}^k p(j|t)(1 - p(j|t)) \\ &= \sum_{j=1}^k p(j|t) - p^2(j|t) \\ &= \sum_{j=1}^k p(j|t) - \sum_{j=1}^k p^2(j|t) \\ &= 1 - \sum_{j=1}^k p^2(j|t) \end{aligned} \quad (2.18)$$

Goodness of split digunakan untuk mengevaluasi pemilahan s pada simpul t . Misalkan terdapat pemilahan s yang memilah simpul t menjadi tepat dua simpul, yaitu simpul kiri t_L dengan proporsi p_L dan simpul kanan t_R dengan proporsi p_R dengan $p_R = 1 - p_L$. Maka pemilah terbaik (*goodness of split*) didefinisikan sebagai penurunan keheterogenan (*improvement*), seperti berikut (Andriyashin, 2005: 11):

$$\Delta i(s, t) = i(t) - p_L \cdot i(t_L) - p_R \cdot i(t_R) \quad (2.19)$$

bahwa $\Delta i(s, t)$ merupakan penurunan tingkat keheterogenan (*improvement*).

Pengembangan pohon dilakukan pada simpul utama t_1 , dengan cara mencari s^* yang memberikan nilai penurunan keheterogenan tertinggi, sebagai berikut:

$$\Delta i(s^*, t_1) = \max_{s \in S} \Delta i(s, t_1) \quad (2.20)$$

Sehingga simpul t_1 , dipilah menjadi t_2 dan t_3 dengan menggunakan s^* . Menggunakan cara yang sama maka dilakukan pemilahan pada t_2 dan t_3 secara terpisah dan seterusnya.

B. Penentuan Simpul Terminal

Suatu simpul t akan menjadi simpul terminal atau simpul akhir yang tidak akan dipilah lagi apabila tidak terdapat penurunan keheterogenan yang berarti. Pada simpul terminal akan hanya terdapat satu pengamatan pada tiap simpul anak atau terdapat batasan minimum sampel n dan ada batasan jumlah tingkatan kedalaman pohon maksimal. Pembentukan simpul terminal ini menjadikan pembentukan pohon klasifikasi berhenti, pemberhentian ini sesuai dengan aturan berikut ini (Tanjung & Kartiko, 2017: 80):

1. Tidak terdapat lagi variabel respon yang menunjukkan perbedaan yang signifikan terhadap variabel prediktor.
2. Jika pohon klasifikasi mencapai batas nilai maksimum dari spesifikasi, maka proses pembentukan akan berhenti. Misalkan ditetapkan batas kedalaman pertumbuhan pohon klasifikasi adalah 4, sehingga ketika pembentukan pohon telah mencapai kedalaman 4 maka pembentukan pohon klasifikasi akan berhenti.
3. Jika ukuran simpul anak kurang dari ukuran simpul anak minimum yang terspesifikasi atau berisi pengamatan dengan jumlah yang terlalu sedikit maka simpul tidak akan terpilah.

C. Penandaan Label Kelas

Breiman dkk (1998: 34) menjelaskan proses penandaan label kelas merupakan proses identifikasi tiap simpul dengan suatu kelas tertentu. Penandaan label tiap simpul akhir didasarkan aturan jumlah anggota kelas terbanyak, yaitu label simpul t dengan kelas $j = j_0$ apabila

$$p(j_0|t) = \max_j p(j|t) = \max_j \frac{N_j(t)}{N(t)} \quad (2.21)$$

Keterangan

$p(j_0|t)$: Proporsi kelas j_0 pada simpul t

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

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

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

2.2.2 Pemangkasan Pohon Klasifikasi

Menurut Breiman dkk (1998), salah satu metode yang dapat digunakan untuk menentukan pohon optimal yaitu dengan pemangkasan (*pruning*). Pada tahap ini, bagian-bagian pohon maksimal yang kurang penting akan dipangkas atau hilangkan untuk mendapatkan pohon klasifikasi yang optimal. Pohon maksimal (T_{max}) yang dihasilkan dari proses pembentukan pohon klasifikasi dapat mengakibatkan dua masalah, yaitu:

1. Pohon maksimal dapat menyebabkan pencocokan nilai yang berlebihan (*overfitting*) pada data-data baru, meskipun T_{max} memiliki nilai akurasi tinggi dengan nilai kesalahan klasifikasi terkecil. *Over fitting* ini terjadi karena terlalu banyaknya variabel independen yang digunakan untuk melakukan prediksi terhadap data baru tersebut.
2. Akibat dari T_{max} yang *over fitting* hal tersebut menjadikan T_{max} sulit dipahami atau diinterpretasikan.

Pemangkasan pohon dilakukan untuk mengatasi kedua masalah diatas dengan cara memotong atau memangkas T_{max} menjadi pohon klasifikasi (T) yang ukurannya lebih kecil dan lebih optimal. Metode yang digunakan adalah *cost complexity minimum*. Misalnya untuk sebarang pohon T yang merupakan sub pohon dari pohon klasifikasi maksimal T_{max} , maka nilai *complexity minimum* adalah (Breiman, 1998: 72):

$$R_{\alpha}(T) = R(T) + \alpha|T| \quad (2.22)$$

Keterangan:

α : Kompleksitas parameter

$|T|$: Ukuran banyaknya simpul terminal pohon T

$R_{\alpha}(T)$: Kombinasi linear nilai dan kompleksitas pohon yang dibentuk dengan menambahkan *cost penalty* bagi kompleksitas terhadap nilai kesalahan klasifikasi.

$R(T)$: Proporsi kesalahan pada klasifikasi pohon atau *resubstitution estimate*, dimana $R(T) = r(t)p(t)$.

$r(t)$ adalah simpul nilai kesalahan klasifikasi diperoleh dari $1 - \max p(j|t)$.

Sedangkan, peluang sebuah objek ($p(t)$) akan berada dalam simpul t diperoleh dengan menjumlahkan seluruh peluang kelas j pada simpul t , dimana $j = 1, 2, 3, \dots, J$. Maka $p(t)$ didefinisikan sebagai berikut:

$$\begin{aligned} p(t) &= \sum_j^J p(j, t) \\ &= p(1, t) + p(2, t) + \dots + p(J, t) \\ &= \frac{N_1(t)}{N} + \frac{N_2(t)}{N} + \dots + \frac{N_J(t)}{N} \end{aligned} \quad (2.23)$$

$$= \frac{N(t)}{N}$$

Sedangkan, peluang bahwa sebuah objek $p(j|t)$ merupakan anggota kelas j dan jika objek ini diketahui berada dalam simpul t , maka peluangnya didefinisikan sebagai berikut:

$$p(j|t) = \frac{p(j, t)}{p(t)} = \frac{\frac{N_j(t)}{N}}{\frac{N(t)}{N}} = \frac{N_j(t)}{N(t)} \quad (2.24)$$

Keterangan:

$\frac{N_j(t)}{N(t)}$: Proporsi objek-objek dalam kelas j yang berada di simpul t

$p(j, t)$: Peluang bahwa sebuah objek menjadi anggota kelas j berada dalam simpul t .

Cost complexity pruning menentukan suatu pohon bagian $T(\alpha)$ yang meminimalkan $R_\alpha(T)$ pada seluruh pohon bagian. Misalkan untuk setiap nilai α pohon bagian $T(\alpha) < T_{max}$ yang meminimalkan $R_\alpha(T)$ adalah:

$$R_\alpha(T(\alpha)) = \min_{T < T_{max}} R_\alpha(T) \quad (2.25)$$

Jika $R(T)$ digunakan untuk menentukan pohon optimal maka akan cenderung pohon terbesar adalah pohon T_1 . Karena semakin besar pohon maka akan semakin kecil nilai $R(T)$ yang diperoleh.

Proses pemangkasan pohon klasifikasi dimulai dengan mengambil simpul anak kanan t_R dan simpul anak kiri t_L yang dihasilkan dari simpul induk t . Jika $R(t) = R(t_R) + R(t_L)$, maka simpul anak t_R dan t_L dipangkas. Proses tersebut diulang-ulang sampai tidak terdapat lagi pemangkasan yang mungkin.

2.2.3 Penentuan Pohon Klasifikasi Optimal

Pohon klasifikasi yang berukuran besar akan memberikan nilai estimasi atau penduga pengganti paling kecil. Sehingga membuat pohon tersebut

cenderung dipilih sebagai penduga nilai respon. Namun, ukuran pohon yang besar dapat menyebabkan kompleksitas yang tinggi karena struktur data yang dipakai cukup kompleks. Maka, diperlukan pohon klasifikasi optimal yang sederhana tetapi memberikan nilai penduga pengganti yang cukup kecil.

Penduga pengganti yang sering digunakan adalah *cross validation estimate* atau penduga validasi silang. Penduga ini disarankan untuk menghitung nilai yang nilai pengganti relatifnya pada sampel yang kurang dari 3000. Jika ukuran sampel lebih dari 3000 sampel maka dapat menggunakan penduga sampel uji atau *test sample estimate*.

Misalkan digunakan sampel *learning* L maka penduga *cross validation* membagi pengamatan L dalam V bagian yang saling lepas dengan ukuran yang kurang lebih sama besar untuk setiap kelas. Pohon $T^{(v)}$ dibentuk dari sampel *learning* ke- v dengan $v = 1, 2, 3, \dots, V$. Misalkan $d^{(v)}(x)$ merupakan hasil pengklasifikasian, maka penduga sampel untuk $R(T_t^{(v)})$ adalah, sebagai berikut:

$$R(T_t^{(v)}) = \frac{1}{N_v} \sum_{(x_n, j_n) \in L_v}^N X(d^{(v)}(x_n) \neq j_n) \quad (2.26)$$

dengan $N_v \cong \frac{N}{V}$ yang merupakan jumlah pengamatan dalam L_v .

2.3 Menghitung Ketepatan Klasifikasi

Menghitung ketepatan klasifikasi merupakan penilaian dari prosedur klasifikasi yang telah dilakukan oleh suatu fungsi klasifikasi yang digunakan untuk mengetahui apakah data hasil klasifikasi sudah benar atau tidak. *Cross validation* merupakan salah satu metode yang dapat digunakan untuk mengukur atau menilai keakuratan sebuah model yang telah diperoleh dari suatu data. Data

yang digunakan untuk proses membuat model disebut data *training*. Sedangkan, data yang digunakan untuk uji validasi adalah data *testing*.

Ukuran yang dapat dipakai untuk melihat ketepatan hasil klasifikasi adalah APER (*Apparent Error Rate*). Menurut Johnson & Wichern (2007: 598), kriteria perbandingan prosedur klasifikasi didasarkan pada kesalahan klasifikasi sehingga APER merupakan nilai dari jumlah observasi yang salah dalam pengklasifikasian suatu fungsi klasifikasi. Nilai APER adalah nilai proporsi sampel yang telah diklasifikasikan. Menghitung APER dapat lebih mudah dengan menggunakan tabel silang antara sampel hasil observasi terhadap sampel taksiran klasifikasi seperti berikut:

Tabel 2.1 Perhitungan Ketepatan Klasifikasi

Hasil Observasi	Taksiran Klasifikasi	
	y_1	y_2
y_1	n_{11}	n_{12}
y_2	n_{21}	n_{22}

Keterangan:

n_{11} : Banyaknya subjek dari y_1 tepat klasifikasi sebagai y_1

n_{12} : Banyaknya subjek dari y_1 salah klasifikasi sebagai y_2

n_{21} : Banyaknya subjek dari y_2 salah klasifikasi sebagai y_1

n_{22} : Banyaknya subjek dari y_2 tepat klasifikasi sebagai y_2

Perhitungan nilai APER ini adalah proporsi pengamatan yang diprediksi salah oleh fungsi klasifikasi dengan n adalah total pengamatan, maka rumus akurasi APER adalah sebagai berikut:

$$\text{APER}(\%) = \left(\frac{n_{21} + n_{12}}{n} \right) \times 100\% \quad (2.27)$$

$$\text{Ketepatan Klasifikasi} = 100\% - \text{APER}(\%)$$

Setelah proses menghitung nilai APER untuk masing-masing metode maka selanjutnya dilakukan perbandingan hasil. Sehingga dapat diperoleh metode terbaik yang mempunyai nilai kesalahan terkecil dan nilai ketepatan klasifikasi tertinggi. Kemudian dibandingkan kembali kinerja pengklasifikasi tersebut menggunakan selang kepercayaan untuk melihat akurasi kotor. Selang kepercayaan seringkali berguna untuk membandingkan kinerja pengklasifikasi yang berbeda untuk menentukan pengklasifikasi mana yang bekerja lebih baik pada kumpulan data tertentu. Tujuan menggunakan selang atau interval kepercayaan adalah untuk melihat margin kesalahan sekecil mungkin. Semakin sempit interval, semakin akurat nilai statistik yang kita gunakan. Perhitungan selang kepercayaan menggunakan rumus seperti berikut:

$$\frac{2 \cdot N \cdot acc + Z_{a/2}^2 \pm Z_{a/2} \sqrt{Z_{a/2}^2 + 4Nacc - 4Nacc^2}}{2(N + Z_{a/2}^2)} \quad (2.28)$$

Keterangan:

N : Banyaknya data

acc : Nilai ketepatan klasifikasi

$Z_{\frac{a}{2}}$: Nilai tabel $Z = 1,96$

2.4 Kesejahteraan dan Kemiskinan Rumah Tangga

Kesejahteraan menjadi aspek penting yang harus diwujudkan untuk menjaga dan membina stabilitas ekonomi dan sosial dalam masyarakat. Kesejahteraan merupakan kemampuan bagi seseorang untuk memenuhi kebutuhan dasar yang bersifat material, seperti sandang, pangan, dan papan, maupun yang non material, yang mencakup aspek gizi dan kesehatan,

pengetahuan, dan kekayaan materi (Cahyat dkk, 2007: 3). Sedangkan, menurut Rambe (2008: 16), kesejahteraan adalah suatu tata kehidupan dan penghidupan sosial, material, maupun spiritual yang diliputi rasa keselamatan, kesusilaan dan ketentraman lahir dan batin yang memungkinkan setiap warga negara mampu mengadakan usaha-usaha pemenuhan jasmani, rohani, dan sosial bagi diri, keluarga serta masyarakat. Kesejahteraan menurut Badan Pusat Statistik (BPS) adalah keadaan yang merepresentasikan dimana rumah tangga dapat memenuhi kebutuhan jasmani dan rohani sesuai dengan tingkat hidup.

Menurut Bappenas, status kesejahteraan penduduk dapat diukur berdasarkan proporsi pengeluaran kebutuhan rumah tangga. Secara garis besar, kebutuhan rumah tangga dapat dikelompokkan menjadi dua jenis, yaitu kebutuhan pokok (pangan) dan kebutuhan non pokok (bukan pangan). Rumah tangga dikategorikan sejahtera apabila proporsi pengeluaran untuk kebutuhan pokok lebih dari sama dengan proporsi pengeluaran untuk kebutuhan non pokok. Sedangkan, apabila rumah tangga dikatakan berstatus kesejahteraan rendah atau miskin maka untuk proporsi pengeluaran untuk kebutuhan pokok lebih besar daripada pengeluaran kebutuhan non pokok.

Sebaliknya, kemiskinan merupakan keadaan yang menunjukkan kurang atau tidak adanya kesejahteraan. Kemiskinan seringkali ditandai dengan tingginya tingkat pengangguran dan keterbelakangan di segala aspek sehingga akan terjadi ketertinggalan masyarakat dibandingkan dengan masyarakat lain yang lebih mempunyai potensi lebih tinggi. Penduduk miskin biasanya lemah dalam kemampuan berusaha dan adanya keterbatasan akses terhadap kegiatan ekonomi (Kartasmita, 1996: 234-235). Sedangkan, menurut Cahyat dkk (2007: 2),

kemiskinan adalah kondisi dimana seseorang atau rumah tangga mengalami kesulitan untuk memenuhi kebutuhan dasar, sementara lingkungan pendukungnya kurang memberikan peluang untuk meningkatkan kesejahteraan secara berkesinambungan.

Secara umum, konsep kemiskinan dapat dibedakan menjadi 2 jenis yaitu kemiskinan absolut dan kemiskinan relatif. Kemiskinan absolut adalah kondisi ketidakmampuan seseorang untuk memenuhi kebutuhan pokok minimum seperti kebutuhan pangan, sandang, kesehatan, perumahan, dan pendidikan. Kebutuhan pokok minimum ini diartikan sebagai nilai atau ukuran finansial dalam bentuk uang. Nilai kebutuhan minimum kebutuhan pokok tersebut disebut sebagai garis kemiskinan. Sedangkan, kemiskinan relatif merupakan kondisi kemiskinan yang disebabkan pengaruh kebijakan pembangunan yang belum mampu menjangkau lapisan masyarakat, sehingga menyebabkan fenomena ketimpangan dalam masyarakat (Bappenas, 2018: 7).

Penentuan kelompok penduduk miskin menurut standar minimum pengukuran kemiskinan disusun berdasarkan kondisi hidup suatu wilayah pada waktu tertentu yang merujuk pada kelompok penduduk termiskin. Misalnya diambil sekian persen penduduk dari posisi terendah dari total penduduk yang telah diurutkan berdasarkan besaran pendapatan atau pengeluaran perkapita per bulan. Kelompok ini yang didefinisikan sebagai kelompok relatif miskin dengan ukuran kemiskinan bergantung pada kondisi pendapatan atau pengeluaran penduduk (Bappenas, 2018: 8).

BPS mengukur kemiskinan dengan menggunakan konsep pemenuhan kebutuhan dasar. BPS memandang kemiskinan sebagai keadaan ketidakmampuan

penduduk dari segi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur berdasarkan sisi pengeluaran. Sehingga dapat dikatakan bahwa penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita per bulan rendah atau dibawah Garis Kemiskinan (GK).

BPS menjelaskan bahwa pengeluaran per kapita adalah biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga. Data pengeluaran rumah tangga dapat mengungkapkan pola konsumsi rumah tangga secara umum dengan menggunakan indikator proporsi pengeluaran untuk makanan dan non makanan. Komposisi pengeluaran rumah tangga dapat dijadikan ukuran untuk menilai tingkat kesejahteraan ekonomi penduduk, makin rendah persentase pengeluaran untuk makanan terhadap total pengeluaran maka tingkat kesejahteraan semakin baik.

Metode yang digunakan BPS dalam melakukan perhitungan jumlah dan presentase penduduk miskin dengan menghitung garis kemiskinan. Garis kemiskinan merupakan jumlah nilai pengeluaran kebutuhan minimum untuk Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). GKM merupakan jumlah nilai pengeluaran kebutuhan minimum makanan yang disetarakan dengan 2.100 kilokalori perkapita per hari. Paket komoditi kebutuhan dasar makanan diwakili oleh 52 jenis komoditi seperti padi-padian, umbi-umbian, ikan, daging, telur dan susu, sayuran, kacang-kacangan, buah-buahan, minyak dan lemak, dll. Sementara GKNM merupakan nilai pengeluaran kebutuhan minimum untuk perumahan, sandang, pendidikan, dan kesehatan.

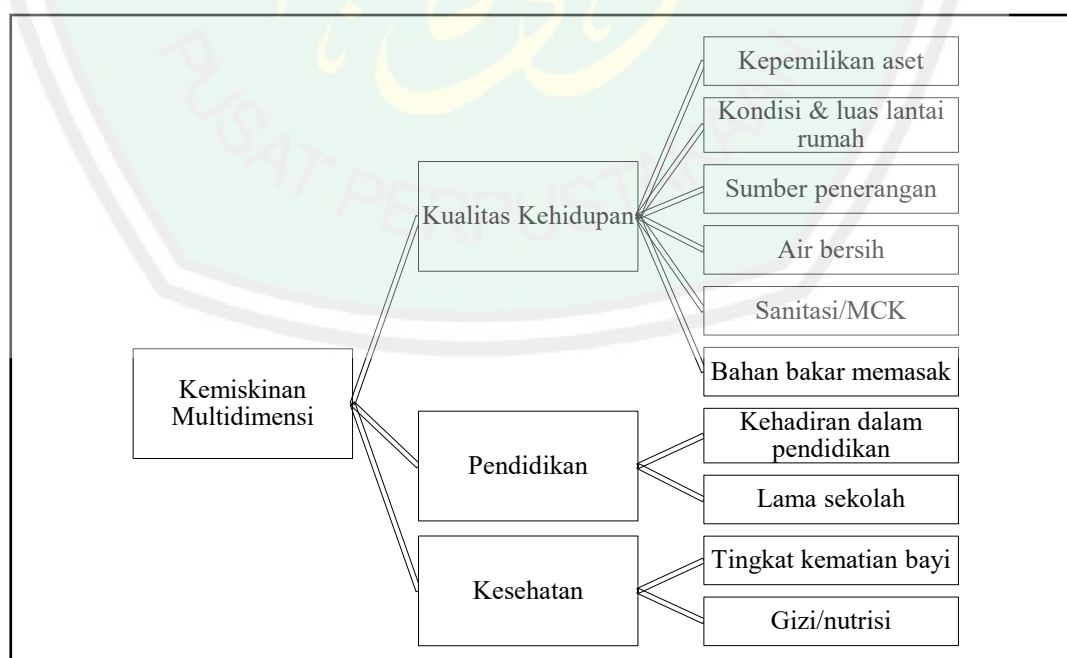
2.4.1 Faktor-faktor Kesejahteraan Rumah Tangga

Aspek rumah tangga atau keluarga sejahtera dapat dilihat berdasarkan kemampuan pemenuhan dasar, seperti kebutuhan pokok pangan, sandang, papan, kesehatan, pendidikan. Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) menggolongkan aspek keluarga sejahtera dalam 4 faktor dominan yang menjadi kebutuhan setiap keluarga. Faktor dominan tersebut terdiri dari (1) pemenuhan kebutuhan dasar, (2) kebutuhan psikologi, (3) kebutuhan pengembangan dan (4) kebutuhan aktualisasi diri untuk berkontribusi bagi masyarakat di lingkungannya. Sedangkan, menurut BKKBN bahwa keluarga prasejahtera adalah keluarga yang belum mampu memenuhi kebutuhan dasarnya secara minimal. Keluarga prasejahtera inilah yang didefinisikan sebagai keluarga miskin (Bappenas, 2010: 9-11).

Selain pengeluaran rumah tangga sebagai indikator kesejahteraan rumah tangga, faktor-faktor yang mempengaruhi status kesejahteraan rumah tangga dapat diukur dari indikator berpengaruh yang lainnya. Indikator kesejahteraan rumah tangga dapat dilihat dari variabel yang berhubungan dengan kepala rumah tangga (KRT), yaitu tingkat pendidikan atau ijazah tertinggi yang dimiliki kepala rumah tangga, status perkawinan, usia, beban tanggungan, pendapatan, dan jenis kelamin dari kepala rumah tangga. Terdapat indikator lainnya seperti kepemilikan aset dan perumahan, misalnya status kepemilikan rumah, memiliki sepeda motor, mobil, televisi, telepon seluler (HP) dan aset lainnya. Terdapat pula indikator sektor pekerjaan keluarga, kesehatan keluarga, banyaknya anggota keluarga yang bekerja, kepemilikan hewan ternak, dan indikator pengkonsumsian makanan, serta partisipasi politik dan akses terhadap informasi (Suryadarma, 2005: 6-7).

Menurut Ramandhani dkk (2017: 126), status kesejahteraan rumah tangga dapat dilihat berdasarkan variabel jenis kelamin kepala rumah tangga, usia kepala rumah tangga, status perkawinan kepala rumah tangga, ijazah tertinggi yang dimiliki, lapangan usaha atau status pekerjaan, dan beban tanggungan jumlah anggota dalam keluarga. Segi sosial ekonomi dapat ditinjau dari bahan bakar atau energi utama yang digunakan untuk memasak dan pengalaman membeli atau mendapat beras miskin (raskin). Kemudian, segi kesehatan berupa pengalaman menerima layanan kesehatan gratis. Serta, segi teknologi dan komunikasi dilihat berdasarkan kemampuan penguasaan atau kepemilikan telepon seluler/HP dan komputer/laptop.

Secara global, United Nations Development Programs (UNDP) mengembangkan pendekatan kemiskinan multidimensi yang menganalisa kemiskinan pada tingkatan rumah tangga maupun individu melalui 3 dimensi, yaitu kualitas kehidupan, pendidikan, dan kesehatan. Setiap dimensi tersebut terdiri dari beberapa indikator, antara lain (Bappenas, 2018: 10-11):



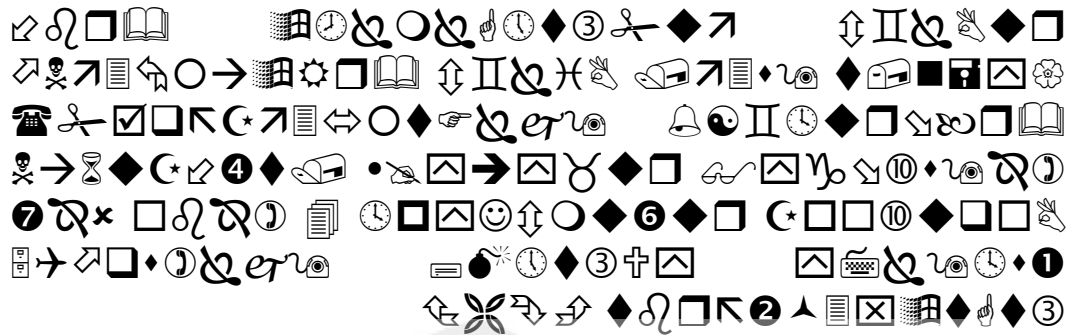
Gambar 2.2 Indikator Kemiskinan UND

2.5 Integrasi Al-Qur'an

Rumah tangga merupakan tempat bernaungnya sebuah keluarga. Pemeliharaan keluarga serta kelangsungan hidup merupakan salah satu tujuan sebuah rumah tangga. Salah satu fungsi rumah tangga bagi keluarga adalah untuk mendapatkan ikatan saudara dalam satu kesatuan sosial ekonomi yang saling mendukung dan saling menunjang. Termasuk didalamnya saling bergantung dalam pemenuhan kebutuhan ekonomi. Keluarga mempunyai tanggung jawab untuk menjaga dan mendidik anggotanya dengan memenuhi kebutuhan jasmani dan rohani.

Kesejahteraan keluarga mencakup kesejahteraan yang bersifat materiil untuk kebutuhan jasmani dan moral untuk kebutuhan rohani. Perlu adanya keseimbangan dari keduanya untuk mencapai kesejahteraan bagi sebuah keluarga. Kesejahteraan materiil dari segi pemenuhan kebutuhan sehari-hari seperti sandang, pangan, dan papan. Sedangkan, kesejahteraan moral seperti rasa kasih sayang dalam keluarga, pemenuhan kebutuhan religi, dan segi pendidikan anak. Kesejahteraan tersebut harus menyeluruh meliputi orang tua, anak serta mampu seimbang antara hak dan kewajiban. Baik hak dan kewajiban untuk diri sendiri, keluarga, masyarakat maupun terhadap Allah SWT.

Al-Qur'an merupakan sumber pokok dari ajaran agama Islam. Al-Quran digunakan sebagai petunjuk bagi manusia, sebagai sumber ilmu pengetahuan terbesar, dan sebagai peringatan dan pembelajaran bagi manusia di segala aspek kehidupan. Terdiri dari 114 surat yang didalamnya terkandung banyak kisah, pelajaran dan pengetahuan. Salah satu kandungan Al-Qur'an tentang kehidupan dalam rumah tangga yaitu pada surat Ar-Ruum ayat 21.



Artinya: “Dan di antara tanda-tanda kekuasaan-Nya ialah Dia menciptakan untukmu isteri-isteri dari jenismu sendiri, supaya kamu cenderung dan merasa tenteram kepadanya, dan dijadikan-Nya diantaramu rasa kasih dan sayang. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda bagi kaum yang berfikir.”

Menurut tafsir Ibnu Katsir, Allah menciptakan bagi kalian (kaum laki-laki) kaum wanita dari jenis kalian yang kelak mereka menjadi istri-istri kalian, supaya kamu cenderung dan merasa tenteram kepadanya. Semakna dengan kandungan dalam surat Al-A’raf ayat 189.



Artinya: “Dialah yang menciptakan kamu dari diri yang satu dan dari padanya Dia menciptakan isterinya, agar Dia merasa senang kepadanya.”

Maknanya yaitu Hawa yang diciptakan oleh Allah SWT dari tulang rusuk bagian kiri Adam. Seandainya Allah menjadikan seluruh anak Adam laki-laki dan menjadikan wanita dari jenis lainnya, seperti bangsa jin atau jenis hewan, maka pastilah tidak akan terdapat kerukunan dan rasa kasih sayang, serta tidak akan terjadi perkawinan. Bahkan sebaliknya akan terjadi pertentangan dan saling berpaling, jika berpasangan bukan dari makhluk sesama manusia. Kemudian, di antara rahmat Allah bagi manusia adalah menjadikan pasangan-pasangan atau istri-istri dari jenis mereka sendiri dan menjadikan rasa cinta dan kasih sayang di

antara mereka. Seorang laki-laki akan tetap memegang rasa cinta kepada wanita dikarenakan lahirnya anak. Sedangkan, seorang perempuan memerlukan perlindungan dan nafkah dari pasangan laki-lakinya.

Surat Ar-Ruum ayat 21 menjelaskan bahwa ada tiga kunci yang disampaikan Allah SWT terkait rumah tangga sejahtera menurut Islam, yaitu *Sakinah*, *Mawadah*, dan *Rahmah*. Menurut Quraish Shihab dalam tafsir Al-Misbah, memahami makna keluarga sejahtera adalah apabila suatu keluarga terdapat unsur: Pertama, *sakinah* diartikan sebagai ketenangan, ketentraman, damai, aman, tentram lahir dan batin antara suami istri serta anggota keluarga yang lain. Kedua, unsur *mawadah* diartikan dengan kelapangan jiwa dan kekosongan hati untuk melakukan hal-hal yang diperintahkan agama. Ketiga, *rahmah* yaitu adanya saling mengasihi, saling menyayangi, saling menghormati, saling menunjang, dan tidak saling menyakiti antara anggota keluarga.

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Pendekatan penelitian yang digunakan adalah studi literatur dan pendekatan deskriptif kuantitatif. Studi literatur merupakan metode penelitian dengan mengumpulkan bahan-bahan pustaka atau teori-teori dari berbagai referensi sebagai acuan peneliti dalam melakukan dan menyelesaikan permasalahan penelitian. Sedangkan, metode deskriptif kuantitatif adalah metode analisis data yang digunakan untuk meneliti populasi atau sampel tertentu dengan proses analisis data bersifat kuantitatif atau menggunakan ilmu statistika. Penelitian kuantitatif menekankan analisis data-data numerik, menghasilkan kesimpulan yang akan memperjelas gambaran objek, dan hubungan yang signifikan antara variabel yang diteliti.

3.2 Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder. Data sekunder merupakan data yang didapatkan secara tidak langsung oleh peneliti atau diperoleh dari sumber yang sudah ada. Data bersumber dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) oleh Badan Pusat Statistik (BPS) tahun 2018 yang diakses melalui www.silastik.bps.go.id. Data yang digunakan sebanyak 464 sampel.

3.3 Variabel Penelitian

Variabel pada penelitian ini terdiri dari variabel respon atau terikat (Y) dan variabel prediktor atau bebas (X). Variabel respon yang digunakan adalah status kesejahteraan rumah tangga yang dikategorikan menjadi 2 kategori, yaitu miskin (0) dan tidak miskin (1). Kategori tersebut ditentukan berdasarkan besarnya pengeluaran per kapita rumah tangga per bulan dan angka garis kemiskinan. Jika pengeluaran per kapita rumah tangga dibawah atau sama dengan garis kemiskinan digolongkan kategori miskin, sedangkan jika pengeluaran per kapita rumah tangga lebih dari garis kemiskinan digolongkan kategori tidak miskin. Variabel prediktor yang digunakan diuraikan sebagai berikut:

Tabel 3.1 Variabel Prediktor

No.	Variabel		Skala	Keterangan
1.	Usia kepala rumah tangga	x_1	Rasio	-
2.	Jenis kelamin kepala rumah tangga	x_2	Nominal	1: Laki-laki 2: Perempuan
3.	Status perkawinan kepala rumah tangga	x_3	Nominal	1: Belum kawin 2: Kawin 3: Cerai hidup 4: Cerai mati
4.	Banyak anggota rumah tangga	x_4	Rasio	-
5.	Tingkat pendidikan tertinggi kepala rumah tangga	x_5	Ordinal	1: SD sederajat 2: SMP sederajat 3: SMA sederajat 4: Perguruan tinggi
6.	Status pekerjaan kepala rumah tangga	x_6	Nominal	1: Berusaha sendiri 2: Berusaha dibantu buruh tidak tetap 3: Berusaha dibantu buruh tetap 4: Buruh/karyawan/ pegawai 5: Pekerjaan bebas 6: Pekerja keluarga
7.	Kesehatan KRT dalam 1 bulan terakhir	x_7	Nominal	1: Ya 2: Tidak

Lanjutan Tabel 3.1 Variabel Prediktor

8.	Status kepemilikan tempat tinggal	x_8	Nominal	1: Milik sendiri 2: Kontrak/sewa 3: Lainnya
9.	Bahan bakar memasak	x_9	Nominal	1: Listrik/gas/elpiji 2: Minyak tanah 3: Kayu 4: Lainnya
10.	Kepemilikan telepon seluler (HP)	x_{10}	Nominal	1: Ya 2: Tidak

3.4 Langkah Penelitian

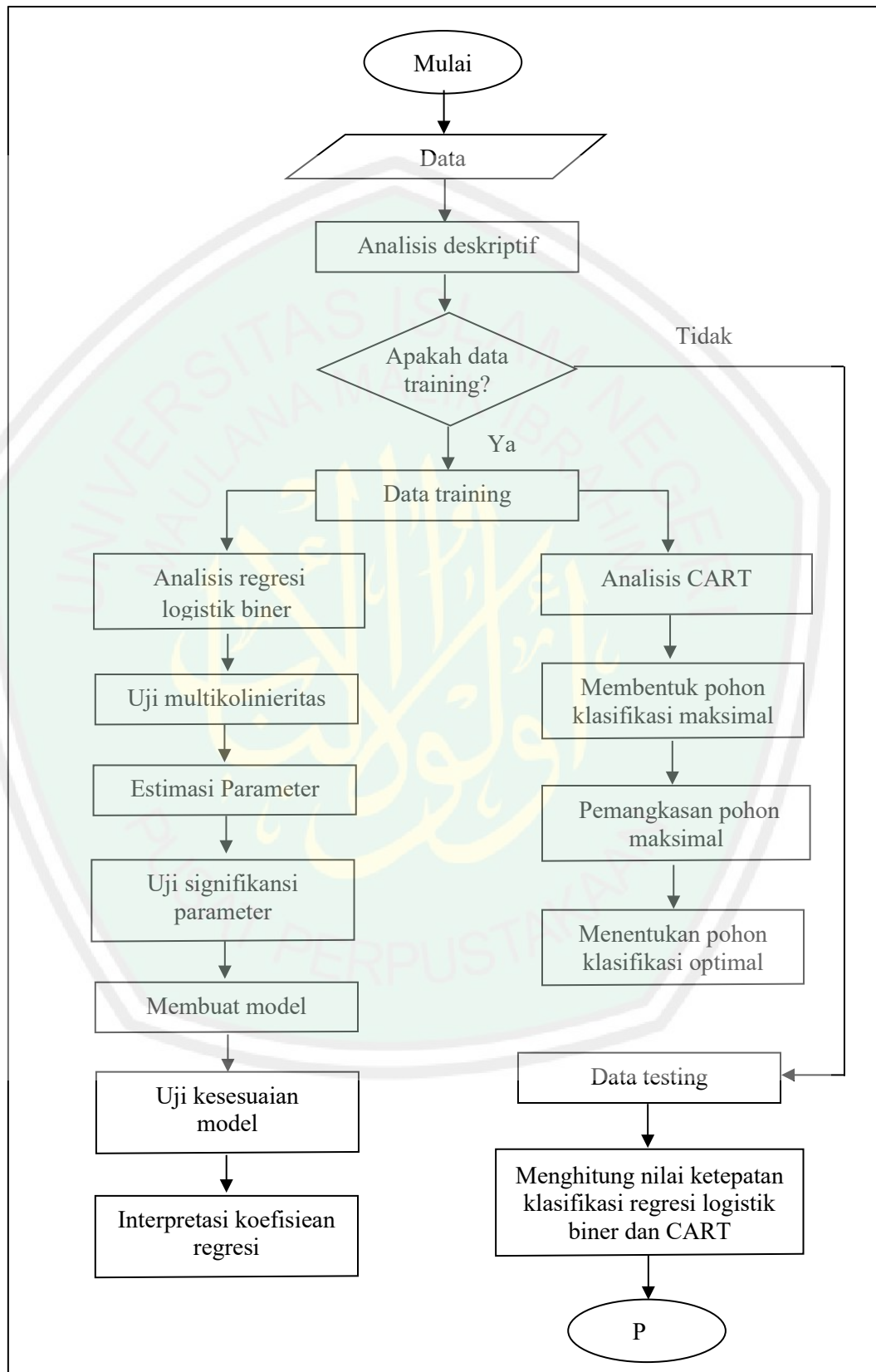
Langkah-langkah dalam penelitian ini adalah sebagai berikut:

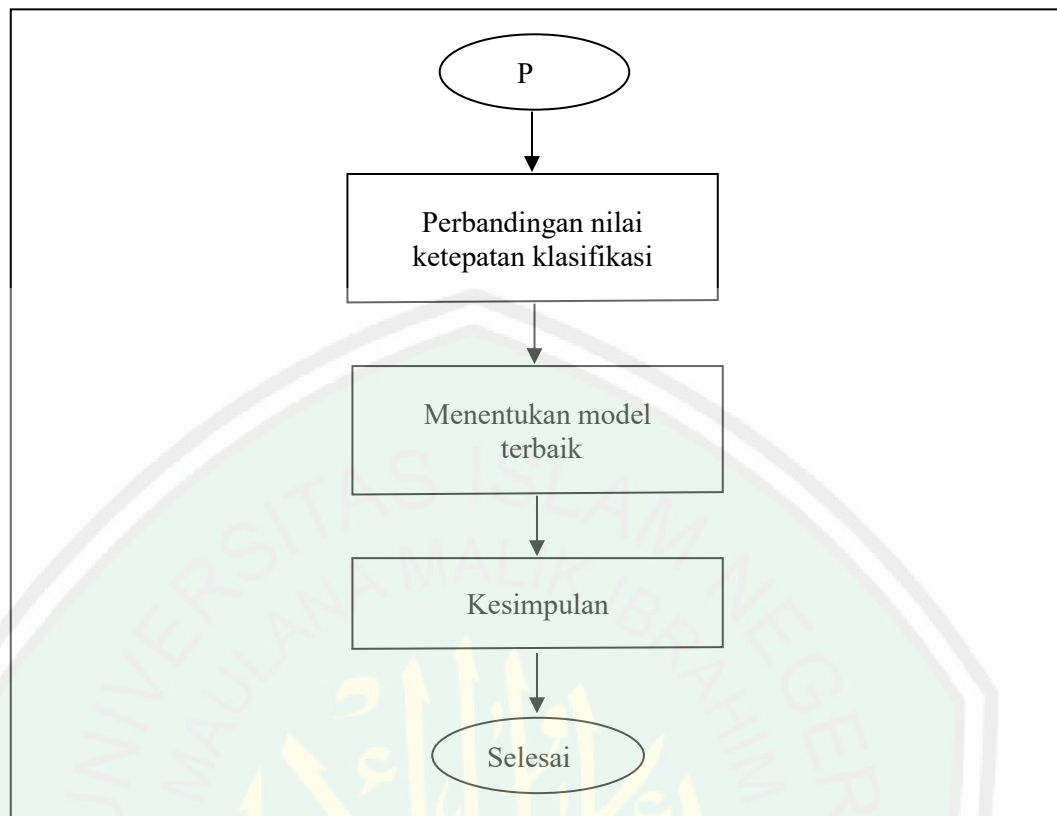
1. Analisis statistika deskriptif variabel-variabel penelitian yang digunakan.
2. Membagi data menjadi dua, yaitu data *training* dan data *testing* sebesar 80% dan 20%.
3. Pengolahan data menggunakan analisis regresi logistik biner. Langkah pada tahapan ini adalah sebagai berikut:
 - a. Melakukan pengujian multikolinieritas.
 - b. Melakukan estimasi parameter.
 - c. Melakukan pengujian signifikansi parameter secara parsial dengan menggunakan uji Wald dan serentak dengan uji G atau log *likelihood*.
 - d. Membentuk model regresi logistik biner.
 - e. Melakukan evaluasi model dengan uji kesesuaian model atau uji Hosmer dan Lemeshow untuk memperoleh model terbaik.
 - f. Menginterpretasi model menggunakan *Odds Ratio*.
 - g. Menghitung ketepatan hasil klasifikasi model.
4. Pengolahan data menggunakan analisis *Classification and Regression Tree* (CART). Langkah pada tahapan ini adalah sebagai berikut:

- a. Membentuk pohon klasifikasi maksimal dengan menggunakan data *training*.
 - i. Melakukan pemecahan atau memilah pemilah dari variabel prediktor dengan menggunakan kaidah indeks Gini.
 - ii. Mengevaluasi hasil pemilahan dengan menggunakan kriteria *goodness of split*.
 - iii. Menentukan simpul terminal.
 - iv. Melakukan penandaan label kelas pada simpul terminal.
 - b. Melakukan pemangkasan menggunakan *cost complexity minimum*.
 - c. Menentukan pohon klasifikasi optimal.
 - d. Mengukur ketepatan model pohon klasifikasi optimal yang diperoleh dengan data *training*, kemudian memasukkan data *testing* ke model tersebut sehingga akan diperoleh angka ketepatan klasifikasi.
 - e. Menginterpretasi hasil model pohon klasifikasi optimal.
5. Menentukan model terbaik dengan membandingkan hasil analisis regresi logistik biner dan analisis CART, melalui nilai ketepatan model klasifikasi dari kedua analisis tersebut.

3.5 Flowchart

Berikut adalah *flowchart* atau diagram alir penelitian yang digunakan:





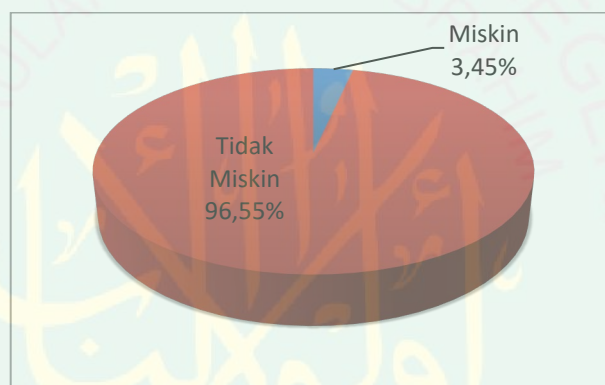
Gambar 3.1 *Flowchart*

BAB IV

ANALISIS DAN PEMBAHASAN

4.1 Analisis Deskriptif

Status kesejahteraan rumah tangga diklasifikasikan menjadi kategori miskin dan kategori tidak miskin. Gambaran secara umum mengenai karakteristik rumah tangga di Kota Batu akan dideskripsikan secara statistik berdasarkan 10 variabel prediktor yang berpengaruh. Deskripsi mengenai status kesejahteraan rumah tangga di Kota Batu disajikan dalam Gambar 4.1.



Gambar 4.1 Status Kesejahteraan Rumah Tangga

Berdasarkan Gambar 4.1 menjelaskan mengenai kesejahteraan rumah tangga di Kota Batu tahun 2018 dari 464 sampel yang diteliti, dapat diketahui bahwa kondisi rumah tangga di Kota Batu lebih banyak yang tidak miskin atau termasuk sejahtera. Sebanyak 96,55 persen atau sebanyak 448 rumah tangga dalam klasifikasi tidak miskin. Sedangkan, sebanyak 3,45 persen atau sebanyak 16 rumah tangga dalam klasifikasi rumah tangga miskin.

Status kesejahteraan rumah tangga didasarkan pada pengeluaran per kapita rumah tangga per bulan dan garis kemiskinan. Pada tahun 2018, BPS telah menetapkan angka garis kemiskinan Kota Batu yaitu Rp. 467.073. Sedangkan,

deskripsi pengeluaran per kapita rumah tangga di Kota Batu tahun 2018 ditunjukkan pada tabel berikut.

Tabel 4.1 Deskripsi Pengeluaran Per Kapita Per Bulan

	<i>Mean</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
Pengeluaran Per Kapita Per Bulan	1.473.051,136	1.179.647,32	332.607,143	13.676.910,71

Tabel 4.1 menunjukkan bahwa nilai pengeluaran per kapita rumah tangga per bulan rata-rata sebesar 1.473.051,136 rupiah dengan median atau nilai tengah sebesar 1.179.647,32 rupiah. Nilai pengeluaran per kapita rumah tangga per bulan terkecil pada angka 332.607,143 rupiah dan tertinggi pada angka 13.676.910,71 rupiah.

4.1.1 Deskripsi Usia Kepala Rumah Tangga

Kepala rumah tangga (KRT) adalah anggota rumah tangga yang bertanggung jawab atas kebutuhan dalam rumah tangga. Usia kepala rumah tangga menjadi salah satu faktor yang mempengaruhi kesejahteraan rumah tangga. Deskripsi status kesejahteraan rumah tangga berdasarkan usia kepala rumah tangga dalam satuan tahun ditunjukkan pada Tabel 4.1.

Tabel 4.2 Usia Kepala Rumah Tangga (Tahun)

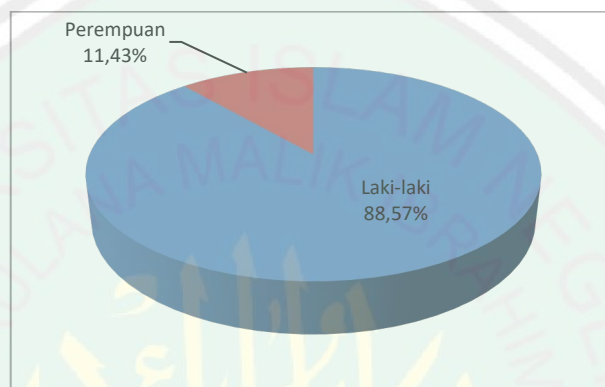
		<i>Mean</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
Y	Miskin	56,12	54	26	75
	Tidak Miskin	47,58	47	18	86

Tabel 4.2 menunjukkan bahwa rumah tangga kategori miskin kategori rata-rata kepala rumah tangga berusia 56,12 tahun dengan median atau nilai tengah pada usia 54 tahun. Sedangkan, usia minimal kepala rumah tangga 26 tahun dan usia maksimum 75 tahun. Rumah tangga yang tidak miskin rata-rata usia kepala rumah tangganya pada usia 47,58 tahun, median usia rumah tangga

tidak miskin 47 tahun, kepala rumah tangga minimal berusia 18 tahun, dan usia maksimum adalah 86 tahun.

4.1.2 Deskripsi Jenis Kelamin Kepala Rumah Tangga

Jenis kelamin kepala rumah tangga merupakan salah satu faktor yang mempengaruhi status kesejahteraan rumah tangga. Deskripsi mengenai jenis kelamin kepala rumah tangga disajikan dalam Gambar 4.2 berikut.



Gambar 4.2 Jenis Kelamin KRT

Gambar 4.2 menunjukkan bahwa jumlah kepala rumah tangga yang berjenis kelamin laki-laki lebih banyak dibandingkan dengan jumlah kepala rumah tangga yang berjenis kelamin perempuan. Sebesar 88,57 persen atau sebanyak 411 KRT berjenis kelamin laki-laki. Sedangkan, kepala rumah tangga perempuan dalam presentase 11,43 persen atau sebanyak 53 orang.

Deskripsi jenis kelamin kepala rumah tangga berdasarkan status kesejahteraan rumah tangga ditunjukkan pada Tabel 4.3.

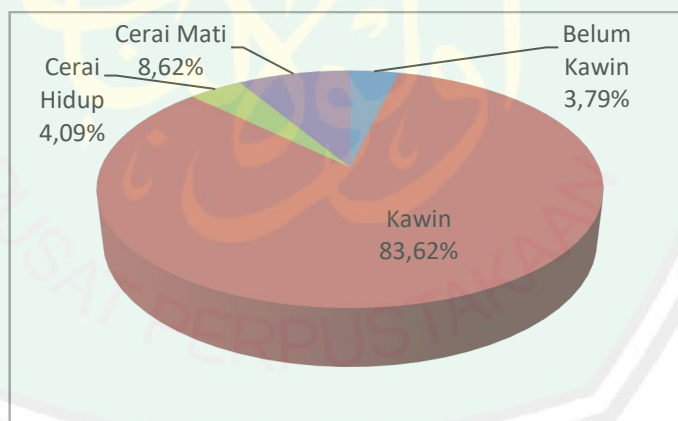
Tabel 4.3 Jenis Kelamin KRT Berdasarkan Status Kesejahteraan

		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Jenis Kelamin	Laki-laki	13 (81,25%)	398 (88,8%)	411 (88,57%)
	Perempuan	3 (18,75%)	50 (11,2%)	53 (11,43%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.3 menjelaskan mengenai jenis kelamin kepala rumah tangga berdasarkan status kesejahteraan. Sebanyak 16 KRT kategori miskin dengan 81,25 persen atau sebanyak 13 KRT berjenis kelamin laki-laki, sedangkan sisanya 18,75 persen atau 3 KRT berjenis kelamin perempuan. Sementara itu, pada rumah tangga tidak miskin dari 448 kepala rumah tangga terdapat 88,8 persen atau sebanyak 398 KRT adalah laki-laki dan sisanya sebanyak 11,2 persen atau 50 KRT adalah perempuan. Sehingga dapat disimpulkan bahwa sebagian besar kepala rumah tangga di Kota Batu berjenis kelamin laki-laki, baik pada status rumah tangga miskin maupun tidak miskin.

4.1.3 Deskripsi Status Perkawinan Kepala Rumah Tangga

Status perkawinan kepala rumah tangga merupakan salah satu faktor yang mempengaruhi status kesejahteraan rumah tangga. Deskripsi mengenai status perkawinan kepala rumah tangga disajikan dalam Gambar 4.3 berikut.



Gambar 4.3 Status Perkawinan KRT

Berdasarkan Gambar 4.3 dapat diketahui bahwa kepala rumah tangga yang paling banyak berstatus kawin dengan ditunjukkan oleh presentase sebesar 83,62 persen atau sebanyak 388 KRT. Sedangkan, status perkawinan terbanyak kedua adalah kepala rumah tangga yang berstatus cerai mati sebesar 8,62 persen atau

sebanyak 40 KRT. Sisanya sebesar 4,09 persen berstatus cerai hidup dan 3,79 persen bersatus belum kawin.

Deskripsi status perkawinan kepala rumah tangga berdasarkan status kesejahteraan rumah tangga ditunjukkan pada Tabel 4.4.

Tabel 4.4 Status Perkawinan Berdasarkan Status Kesejahteraan

		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Status Perkawinan	Belum Kawin	0 (0%)	17 (3,79%)	17 (3,79%)
	Kawin	11 (68,75%)	377 (84,15%)	388 (83,62%)
	Cerai Hidup	2 (12,50%)	17 (3,79%)	19 (4,09%)
	Cerai Mati	3 (18,75%)	37 (8,26%)	40 (8,62%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.4 menunjukkan bahwa sebanyak 16 KRT berkategori miskin, sebesar 68,75 persen atau sebanyak 11 KRT berstatus kawin, sebesar 18,75 persen atau terdapat 3 KRT berstatus cerai mati, sebesar 12,50 persen atau terdapat 2 KRT bersatus cerai hidup, dan tidak terdapat kepala rumah tangga yang berstatus belum kawin. Pada kategori tidak miskin terjadi hal yang sama, dari 448 KRT sebesar 84,15 persen atau 377 KRT berstatus kawin. Kemudian, status cerai mati menjadi terbanyak kedua sebesar 8,26 persen atau terdapat 37 KRT. Sisanya kepala rumah tangga yang berstatus belum kawin dan cerai hidup memiliki presentase sama yaitu 3,79 persen atau sebanyak 17 KRT. Sehingga dapat diketahui bahwa sebagian besar merupakan kepala rumah tangga berstatus kawin, baik kategori miskin maupun tidak miskin.

4.1.4 Deskripsi Banyak Anggota Rumah Tangga

Deskripsi banyak anggota rumah tangga yang dikelompokkan menurut status kesejahteraan disajikan dalam Tabel 4.5.

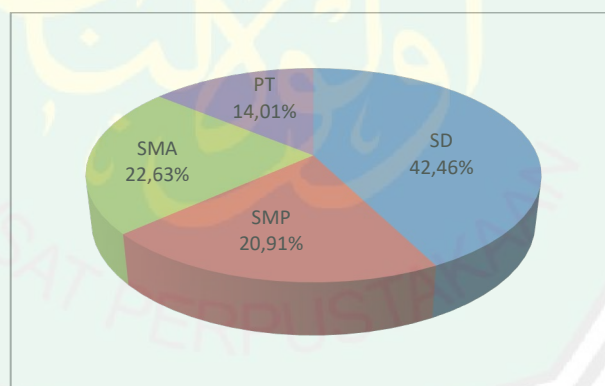
Tabel 4.5 Banyak Anggota Rumah Tangga

		<i>Mean</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
Banyak	Miskin	4	3	2	7
ART	Tidak Miskin	4	3	1	12

Tabel 4.5 menunjukkan bahwa rata-rata banyak anggota rumah tangga baik kategori miskin dan tidak miskin adalah 4 anggota. Nilai median untuk kedua kategori adalah 3. Kategori miskin memiliki anggota rumah tangga paling sedikit adalah 2 anggota, sedangkan paling banyak terdapat 7 anggota dalam rumah tangga. Sementara itu, rumah tangga kategori tidak miskin paling sedikit beranggotakan 1 orang dan rumah tangga paling banyak sebanyak 12 anggota.

4.1.5 Deskripsi Tingkat Pendidikan Kepala Rumah Tangga

Tingkat pendidikan dari kepala rumah tangga menjadi faktor pengaruh terhadap kesejahteraan rumah tangga. Deskripsi tingkat pendidikan dari rumah tangga di Kota Batu disajikan dalam Gambar 4.4.



Gambar 4.4 Tingkat Pendidikan KRT

Berdasarkan Gambar 4.4 dapat diketahui bahwa kepala rumah tangga paling banyak memiliki riwayat pendidikan tertinggi adalah SD sederejat yaitu sebesar 42,46 persen atay sebanyak 197 KRT. Kedua adalah kepala rumah tangga dengan tingkat pendidikan SMA sederajat dengan presentase 22,63 persen atau sebanyak 105 KRT. Sisanya tingkat pendidikan SMP sebanyak 97 KRT atau

20,91 persen dan tingkat pendidikan perguruan tinggi 14,01 persen atau sebanyak 65 KRT.

Deskripsi tingkat pendidikan kepala rumah tangga berdasarkan status kesejahteraan rumah tangga ditunjukkan pada Tabel 4.6.

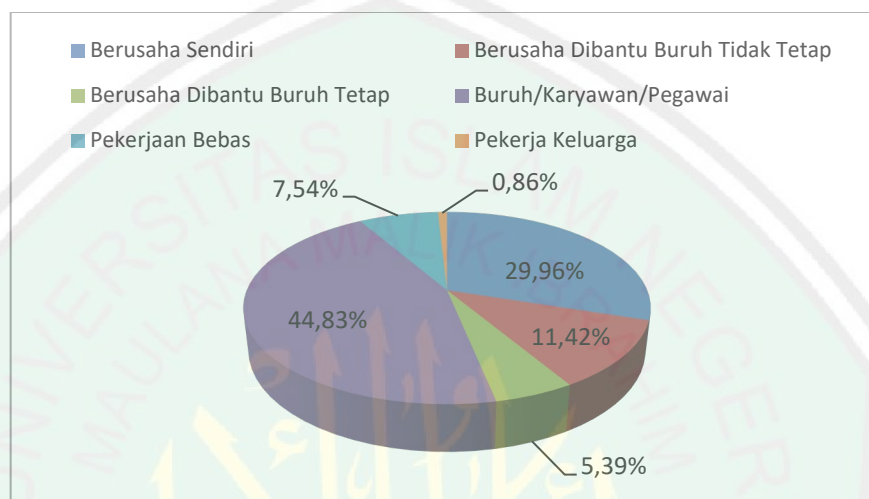
Tabel 4.6 Status Perkawinan Berdasarkan Status Kesejahteraan

		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Tingkat Pendidikan	SD Sederajat	12 (75%)	185 (41,29%)	197 (42,46%)
	SMP Sederajat	1 (6,25%)	96 (21,43%)	97 (20,91%)
	SMA Sederajat	2 (12,5%)	103 (22,99%)	105 (22,63%)
	Perguruan Tinggi	1 (6,25%)	64 (14,29%)	65 (14,01%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.6 menunjukkan bahwa sebanyak 16 KRT kategori miskin, sebesar 75 persen atau sebanyak 12 KRT berpendidikan SD sederajat, kemudian sebesar 12,5 persen atau terdapat 2 KRT merupakan lulusan SMA sederajat, dan sisanya pada tingkat pendidikan SMP dan perguruan tinggi memiliki presentase sama yaitu 6,25 persen atau hanya terdapat 1 KRT. Pada kategori tidak miskin terjadi hal yang sama, dari 448 KRT sebesar 41,29 persen atau 185 KRT memiliki tingkat pendidikan SD. Kemudian, pada tingkat pendidikan SMA menjadi terbanyak kedua sebesar 22,99 persen atau terdapat 103 KRT. Sisanya kepala rumah tangga yang mempunyai riwayat pendidikan tertinggi SMP sebesar 21,43 persen atau sebanyak 96 KRT. Sisanya sebesar 14,29 persen atau sebanyak 64 KRT pada tingkat pendidikan perguruan tinggi. Sehingga dapat disimpulkan masih lebih banyak kepala rumah tangga di Kota Batu merupakan penduduk berpendidikan sekolah dasar.

4.1.6 Deskripsi Status Pekerjaan Kepala Rumah Tangga

Status pekerjaan adalah jenis kedudukan kepala rumah tangga dalam pekerjaan utama. Status pekerjaan dari kepala rumah tangga menjadi faktor berpengaruh terhadap kesejahteraan rumah tangga. Deskripsi status pekerjaan kepala rumah tangga di Kota Batu disajikan dalam Gambar 4.5.



Gambar 4.5 Status Pekerjaan KRT

Berdasarkan Gambar 4.5 dapat diketahui bahwa kepala rumah tangga paling banyak berstatus pekerjaan sebagai buruh/karyawan/pegawai dengan presentase 44,83 persen atau sebanyak 208 KRT. Terbanyak kedua adalah pekerjaan kepala rumah tangga yang berusaha sendiri dengan presentase 29,96 persen atau sebanyak 139 KRT. Kemudian, status pekerjaan kepala rumah tangga yang berusaha sendiri dibantu buruh tidak tetap/buruh tidak dibayar sebanyak 5,39 persen atau sebanyak 53 KRT. Selanjutnya, kepala rumah tangga dengan pekerjaan bebas memiliki presentase 7,54 persen atau sebanyak 35 KRT. Sisanya adalah status pekerjaan berusaha dibantu buruh tetap sebesar 5,39 persen atau sebanyak 25 KRT dan sebagai pekerja keluarga 0,86 persen atau terdapat 4 KRT.

Deskripsi status pekerjaan kepala rumah tangga berdasarkan status kesejahteraan rumah tangga ditunjukkan pada Tabel 4.7.

Tabel 4.7 Status Pekerjaan KRT Berdasarkan Status Kesejahteraan

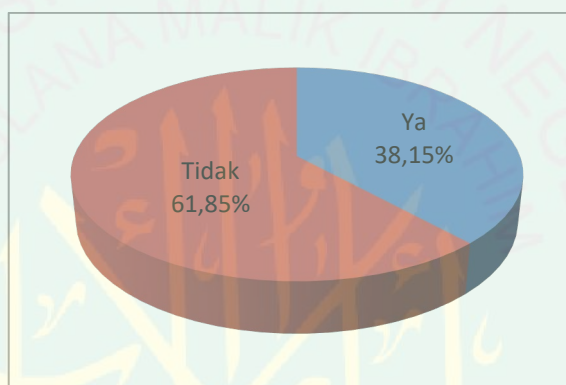
		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Status Pekerjaan	Berusaha Sendiri	7 (43,75%)	132 (29,46%)	139 (29,96%)
	Berusaha Dibantu			
	Buruh Tidak Tetap	1 (6,25%)	52 (11,61%)	53 (11,42%)
	Berusaha Dibantu			
	Buruh Tetap	0 (0%)	25 (5,58%)	25 (5,39%)
	Buruh/Karyawan/ Pegawai	4 (25%)	204 (45,54%)	208 (44,83%)
	Pekerjaan Bebas	4 (25%)	31 (6,92%)	35 (7,54%)
	Pekerja Keluarga	0 (0%)	4 (0,89%)	4 (0,86%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.7 menunjukkan bahwa kepala rumah tangga dalam kategori miskin dari 16 KRT paling banyak dalam status pekerjaan berusaha sendiri dengan presentase 43,75 persen atau sebanyak 7 KRT. Selanjutnya status pekerjaan buruh/karyawan/pegawai dan pekerja bebas memiliki presentase sama yaitu 25 persen atau terdapat 4 KRT. Sisanya sebagai pekerja berusaha dibantu buruh tidak tetap sebesar 6,25 persen atau terdapat 1 KRT. Sedangkan, tidak terdapat kepala rumah tangga dalam status pekerjaan berusaha dibantu buruh tetap dan pekerja keluarga. Pada kategori tidak miskin, kepala rumah tangga paling banyak pada pekerjaan sebagai buruh/karyawan/pegawai dengan presentase 45,54 persen atau sebanyak 204 KRT. Status pekerjaan berusaha sendiri menjadi terbanyak kedua dengan presentase 29,46 persen atau sebanyak 132 KRT. Selanjutnya status pekerjaan berusaha dibantu buruh tidak tetap sebesar 11,61 persen atau sebanyak 52 KRT. Sisanya status pekerjaan sebagai pekerja bebas sebesar 6,92 persen atau sebanyak 31 KRT dan berusaha dibantu buruh tetap sebesar 5,58 persen atau sebanyak 25 KRT. Paling sedikit adalah status pekerjaan sebagai pekerja keluarga yaitu 0,89 persen atau hanya terdapat 4 KRT. Sehingga

dapat dikatakan kepala rumah tangga paling banyak berpekerjaan sebagai buruh atau karyawan atau pegawai.

4.1.7 Deskripsi Kesehatan Kepala Rumah Tangga

Keluhan kesehatan keadaan kepala rumah tangga yang mengalami gangguan kesehatan atau kejiwaan, baik karena gangguan, penyakit, maupun kecelakaan dalam jangka waktu satu bulan terakhir sebelum pencacahan oleh tim SUSENAS. Kondisi kesehatan sebagai salah satu faktor yang mempengaruhi kesejahteraan rumah tangga yang dideskripsikan pada Gambar 4.6.



Gambar 4.6 Kesehatan KRT

Berdasarkan Gambar 4.6 dapat diperoleh informasi bahwa dari 464 sampel kepala rumah tangga sebesar 61,85 persen atau sebanyak 287 KRT tidak mengalami gangguan kesehatan selama kurun waktu satu bulan terakhir. Sedangkan, sisaya sebanyak 177 KRT atau 38,15 persen mengalami gangguan kesehatan. Deskripsi status kesehatan kepala rumah tangga yang dikelompokkan berdasarkan status kesejahteraan yang ditunjukkan pada Tabel 4.8 berikut.

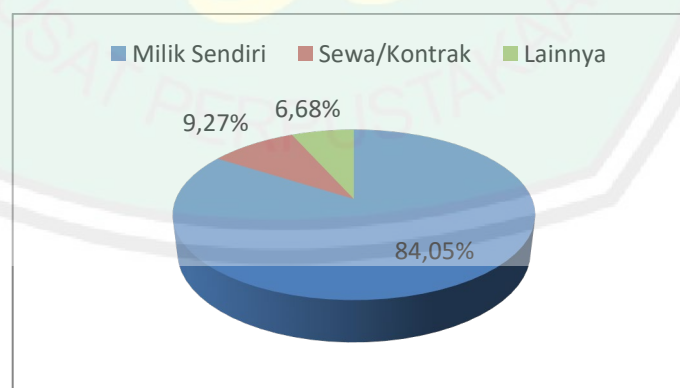
Tabel 4.8 Kesehatan KRT Berdasarkan Status Kesejahteraan

		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Kesehatan	Ya	4 (25%)	173 (38,62%)	177 (38,15%)
	Tidak	12 (75%)	275 (61,38%)	287 (61,85%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.8 menunjukkan mengenai kondisi kesehatan kepala rumah tangga berdasarkan status kesejahteraan. Sebanyak 16 KRT kategori miskin terdapat 4 KRT atau sebesar 25 persen mengalami gangguan kesehatan selama kurun waktu satu bulan terakhir. Sedangkan, sisanya 75 persen atau sebanyak 12 KRT tidak mengalami gangguan kesehatan. Sementara itu, pada rumah tangga kategori tidak miskin dari 448 kepala rumah tangga terdapat 38,62 persen atau sebanyak 173 KRT mengalami gangguan kesehatan dan sebanyak 61,38 persen atau 275 KRT tidak mengalami gangguan kesehatan. Sehingga dapat disimpulkan bahwa lebih banyak kepala rumah tangga di Kota Batu tidak mengalami gangguan kesehatan dalam kurun waktu satu bulan, baik pada status rumah tangga miskin maupun tidak miskin.

4.1.8 Deskripsi Status Kepemilikan Tempat Tinggal

Status kepemilikan tempat tinggal atau rumah yang ditempati suatu rumah tangga menjadi salah satu faktor yang dapat mengukur status kesejahteraan rumah tangga. Berikut adalah deskripsi status kepemilikan tempat tinggal rumah tangga di Kota Batu yang disajikan dalam Gambar 4.7.



Gambar 4.7 Status Kepemilikan Tempat Tinggal

Gambar 4.7 menjelaskan bahwa sebesar 84,05 persen atau sebanyak 390 rumah tangga memiliki status kepemilikan tempat tinggal dengan status milik

sendiri. Rumah tangga yang tempat tinggalnya berstatus sewa atau kontrak sebesar 9,27 persen atau sebanyak 43 rumah tangga. Sisanya 6,68 persen atau sebanyak 31 rumah tangga tempat tinggalnya berstatus lainnya. Kategori lainnya yang dimaksud mencakup tempat tinggal bebas sewa, rumah dinas, dan lainnya.

Deskripsi status kepemilikan tempat tinggal rumah tangga berdasarkan status kesejahteraan ditunjukkan dalam Tabel 4.9.

Tabel 4.9 Status Kepemilikan Tempat Tinggal

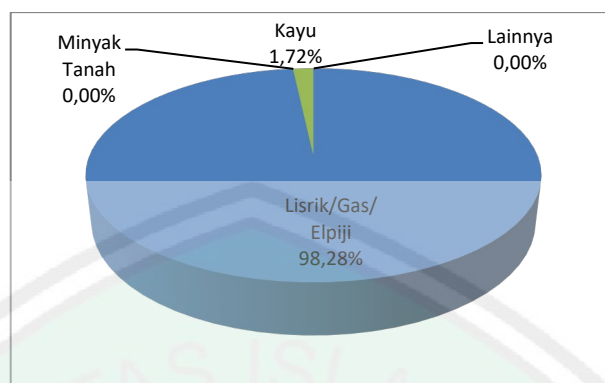
		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Status Kepemilikan	Milik Sendiri	14 (87,5%)	376 (83,93%)	390 (84,05%)
	Sewa/Kontrak	2 (12,5%)	41 (91,5%)	43 (9,27%)
	Lainnya	0 (0%)	31 (6,92%)	31 (6,68%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.9 menunjukkan bahwa rumah tangga dalam kategori miskin paling banyak memiliki status tempat tinggal milik sendiri yaitu presentase 87,5 persen atau sebanyak 14 rumah tangga. Sisanya 12,5 persen atau terdapat 2 rumah tangga memiliki status tempat tinggal sewa atau kontrak. Sedangkan, bagi rumah tangga kategori tidak miskin paling banyak juga pada status tempat tinggal milik sendiri dengan presentase sebesar 83,93 persen atau sebanyak 376 rumah tangga. Status sewa atau kontrak pada posisi kedua yaitu 91,5 persen atau sebanyak 41 rumah tangga. Status tempat tinggal lainnya seperti rumah bebas sewa, rumah dinas atau rumah adat sebesar 6,92 persen atau sebanyak 31 rumah tangga. Sehingga dapat diketahui bahwa status kepemilikan tempat tinggal terbanyak adalah milik sendiri.

4.1.9 Deskripsi Bahan Bakar Memasak

Jenis bahan bakar memasak utama yang digunakan oleh rumah tangga menjadi salah satu faktor yang mempengaruhi status kesejahteraan rumah tangga.

Deskripsi dari variabel bahan bakar memasak utama yang digunakan rumah tangga disajikan dalam Gambar 4.8.



Gambar 4.8 Bahan Bakar Memasak

Gambar 4.8 menjelaskan bahwa sebesar 98,28 persen atau sebanyak 456 rumah tangga menggunakan bahan bakar memasak utamanya adalah listrik atau gas atau elpiji. Terdapat 1,72 persen atau sebanyak 8 rumah tangga masih menggunakan bahan bakar memasak kayu. Sedangkan, tidak terdapat rumah tangga yang menggunakan bahan bakar memasak minyak tanah atau lainnya, seperti briket, batubara, arang.

Deskripsi bahan bakar memasak rumah tangga berdasarkan status kesejahteraan ditunjukkan dalam Tabel 4.10.

Tabel 4.10 Bahan Bakar Memasak Berdasarkan Status Kesejahteraan

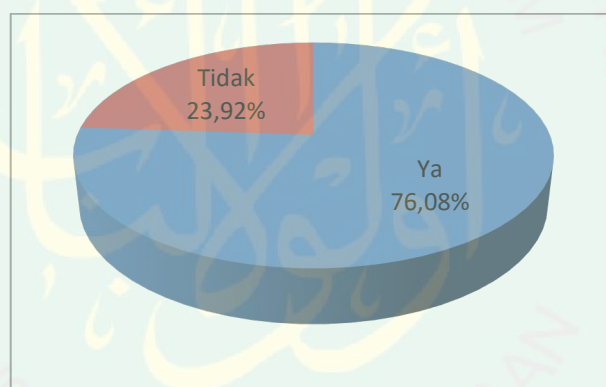
		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Bahan Bakar Memasak	Listrik/Gas/Elpiji	16 (100%)	440 (98,21%)	456 (98,28%)
	Minyak Tanah	0 (0%)	0 (0%)	0 (0%)
	Kayu	0 (0%)	8 (1,79%)	8 (1,72%)
	Lainnya	0 (0%)	0 (0%)	0 (0%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.10 menunjukkan bahwa rumah tangga kategori miskin 100 persen atau seluruhnya 16 rumah tangga menggunakan bahan bakar listik atau gas atau elpiji. Sehingga tidak ada rumah tangga menggunakan bahan bakar memasak

minyak tanah, kayu, atau lainnya. Sedangkan, rumah tangga kategori tidak miskin mayoritas juga menggunakan bahan bakar listrik atau gas atau elpiji dengan presentase 98,12 persen atau sebanyak 440 rumah tangga. Sisanya 1,79 persen atau terdapat 8 rumah tangga yang masih menggunakan kayu sebagai bahan bakar. Sehingga dapat disimpulkan bahwa sebagian besar rumah tangga menggunakan listrik atau gas atau elpiji sebagai bahan bakar memasak.

4.1.10 Deskripsi Kepemilikan Telepon Seluler

Salah satu faktor yang mempengaruhi kesejahteraan adalah faktor teknologi. Faktor ini dapat diukur dari status kepemilikan, penggunaan, dan penguasaan telepon seluler (HP). Deskripsi status kepemilikan telepon seluler (HP) disajikan dalam Gambar 4.9.



Gambar 4.9 Status Kepemilikan HP

Gambar 4.9 menyajikan data bahwa sebesar 76,08 persen atau sebanyak 353 KRT memiliki, menggunakan, dan menguasai telepon seluler (HP). Sisanya 23,92 persen atau sebanyak 111 KRT tidak memiliki, tidak menggunakan, dan tidak menguasai telepon seluler (HP).

Deskripsi status kepemilikan telepon seluler atau HP berdasarkan status kesejahteraan ditunjukkan dalam Tabel 4.11.

Tabel 4.11 Status Kepemilikan HP Berdasarkan Status Kesejahteraan

		Status Kesejahteraan		Total
		Miskin	Tidak Miskin	
Kepemilikan HP	Ya	9 (56,25%)	344 (76,79%)	353 (76,08%)
	Tidak	7 (43,75%)	104 (23,21%)	111 (23,92%)
	Total	16 (100%)	448 (100%)	464 (100%)

Tabel 4.11 menunjukkan bahwa rumah tangga kategori miskin sebesar 56,25 persen atau 9 KRT memiliki, menggunakan, dan menguasai telepon seluler (HP). Sedangkan, sisanya 53,75 persen atau 7 KRT dalam status tidak. Rumah tangga kategori tidak miskin sebesar 76,79 persen atau sebanyak 344 KRT memiliki, menggunakan, dan menguasai telepon seluler (HP). Sisanya 23,21 persen atau sebanyak 104 KRT dalam status tidak memiliki, tidak menggunakan, dan tidak menguasai telepon seluler (HP).

4.2 Analisis Regresi Logistik Biner

Sebelum dilakukan analisis, tahapan awal yang perlu dilakukan adalah membagi data menjadi dua kelompok, yaitu data *training* dan data *testing*. Adapun pembagian data *training* dan *testing* adalah 80% dan 20%. Berikut hasil analisis regresi logistik biner untuk mengetahui faktor-faktor yang mempengaruhi status kesejahteraan rumah tangga di Kota Batu.

4.2.1 Uji Multikolinieritas

Pengujian multikolinieritas dilakukan untuk mengetahui terdapat atau tidaknya korelasi antara variabel prediktor pada model regresi dengan cara melihat nilai VIF. Berikut merupakan hasil dari pengujian multikolinieritas.

Hipotesis pengujian:

H_0 : Terdapat gejala multikolineritas antara variabel prediktor ($VIF > 10$)

H_1 : Tidak terdapat gejala multikolineritas antara variabel prediktor ($VIF < 10$)

Taraf signifikansi: $\alpha = 0,05$

Daerah keputusan: Tolak H_0 jika $VIF < 10$

Tabel 4.12 Nilai VIF

Prediktor	VIF
x_1	1,478
x_2	1,970
x_3	1,381
x_4	1,131
x_5	1,098
x_6	1,092
x_7	1,161
x_8	1,092
x_9	1,000
x_{10}	1,447

Berdasarkan Tabel 4.12 dapat diketahui bahwa nilai VIF untuk masing-masing variabel prediktor kurang dari 10 maka tolak H_0 atau dapat dikatakan terima H_1 . Sehingga dapat disimpulkan bahwa pada data yang digunakan tidak terdapat gejala multikolineritas antara variabel prediktor.

4.2.2 Estimasi Parameter

Berikut merupakan hasil dari estimasi parameter pembanding yang selanjutnya akan diuji tingkat signifikansinya.

Tabel 4.13 Estimasi Parameter

Variabel	Estimasi β_i
Konstanta	23,454
x_1	-0,1026
x_2 (1)	0,213
x_3 (1)	-14,138
x_3 (2)	-18,264
x_3 (3)	-15,038
x_4	-0.067
x_5 (1)	1,382
x_5 (2)	0,316

Lanjutan Tabel 4.13

x_5 (3)	-0,922
x_6 (1)	-0,809
x_6 (2)	18,274
x_6 (3)	1,975
x_6 (4)	-2,240
x_6 (5)	18,717
x_7 (1)	-0,981
x_8 (1)	-0,676
x_8 (2)	17,186
x_9 (2)	16,556
x_{10} (1)	1,231

4.2.3 Uji Signifikansi Parameter

Pengujian signifikansi parameter dilakukan untuk melihat peranan variabel prediktor yang digunakan berpengaruh secara signifikan terhadap variabel respon. Uji signifikansi parameter meliputi uji signifikansi secara serentak dan uji signifikansi secara parsial.

1. Uji Signifikansi Parsial

Uji signifikansi parameter parsial dilakukan dengan pengujian secara individu terhadap masing-masing variabel prediktor. Uji secara parsial dilakukan untuk mengetahui pengaruh dari masing-masing variabel prediktor terhadap variabel respon. Pengujian ini menggunakan statistik uji Wald. Berikut adalah hasil pengujian signifikansi parameter secara parsial.

Hipotesis pengujian:

$$H_0 : \beta_i = 0 \text{ dengan } i = 1, 2, 3, \dots, 10$$

(Variabel prediktor (x_i) tidak memberikan pengaruh signifikan secara parsial terhadap variabel respon)

$$H_1 : \text{Minimal terdapat satu } \beta_i \neq 0 \text{ dengan } i = 1, 2, 3, \dots, 10$$

(Minimal terdapat satu variabel prediktor (x_i) yang memberikan pengaruh signifikan secara parsial terhadap variabel respon)

Taraf signifikansi: $\alpha = 0,05$

Daerah keputusan: Jika $|W_i| > Z_{\alpha/2} = 1,96$ atau $p\text{-value} < \alpha$, maka tolak H_0

Statistik uji: $W_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$

Tabel 4.14 Uji Wald Pertama

Variabel	Estimasi	SE	Wald	P-value	Keputusan
Konstanta	23,454	4629,509	0,005	0,995	
x_1	-0,1026	0,0394	-2,604	0,009*	Tolak H_0
x_2 (1)	0,213	1,544	0,138	0,890	Terima H_0
x_3 (1)	-14,138	4692,509	-0,003	0,997	Terima H_0
x_3 (2)	-18,264	4692,509	-0,004	0,996	Terima H_0
x_3 (3)	-15,038	4692,509	-0,003	0,997	Terima H_0
x_4	-0,067	0,234	-0,290	0,771	Terima H_0
x_5 (1)	1,382	1,201	1,151	0,249	Terima H_0
x_5 (2)	0,316	1,248	0,254	0,799	Terima H_0
x_5 (3)	-0,922	1,306	-0,706	0,480	Terima H_0
x_6 (1)	-0,809	1,203	0,672	0,501	Terima H_0
x_6 (2)	18,274	3621,883	0,005	0,995	Terima H_0
x_6 (3)	1,975	1,133	1,743	0,081	Terima H_0
x_6 (4)	-2,240	0,912	-2,455	0,014*	Tolak H_0
x_6 (5)	18,717	9003,787	0,002	0,998	Terima H_0
x_7 (1)	-0,981	0,796	-1,232	0,218	Terima H_0
x_8 (1)	-0,676	1,072	-0,631	0,528	Terima H_0
x_8 (2)	17,186	2789,425	0,006	0,995	Terima H_0
x_9 (2)	16,556	5521,274	0,003	0,997	Terima H_0
x_{10} (1)	1,231	0,976	1,261	0,207	Terima H_0

*: Signifikan terhadap α (0,05)

Berdasarkan Tabel 4.14 dapat diketahui dari hasil uji signifikansi secara parsial menunjukkan bahwa terdapat 2 variabel prediktor yang berpengaruh terhadap status kesejahteraan rumah tangga di Kota Batu, yaitu variabel x_1 (usia KRT) dan variabel x_6 (status pekerjaan KRT). Pada variabel x_1 menghasilkan nilai $W = |-2,604| = 2,604 > Z_{\alpha/2} = 1,96$ dan nilai $p\text{-value} = 0,009 < \alpha = 0,05$,

maka keputusannya adalah tolak H_0 . Pada variabel x_6 menghasilkan nilai yang signifikan pada variabel $x_{6(4)}$ yaitu variabel status pekerjaan KRT kategori 4 (buruh/karyawan/pegawai) dengan nilai $W = |-2,455| = 2,455 > Z_{\alpha/2} = 1,96$ dan nilai $p\text{-value} = 0,014 < \alpha = 0,05$, maka keputusannya adalah tolak H_0 . Artinya, variabel prediktor x_1 dan variabel prediktor x_6 masing-masing memberikan pengaruh signifikan secara parsial terhadap status kesejahteraan rumah tangga.

Sedangkan, variabel yang lain seperti variabel jenis kelamin KRT (x_2), status perkawinan (x_3), banyak anggota rumah tangga (x_4), tingkat pendidikan tertinggi (x_5), kesehatan KRT 1 bulan terakhir (x_7), status kepemilikan tempat tinggal (x_8), bahan bakar memasak (x_9), dan kepemilikan telepon seluler (x_{10}) tidak memberikan pengaruh secara parsial terhadap variabel status kesejahteraan rumah tangga. Karena menghasilkan nilai Wald yang lebih kecil dari nilai $Z_{\alpha/2}$ dan $p\text{-value}$ lebih besar dari nilai α .

Pada uji signifikansi parsial pertama menghasilkan 8 variabel prediktor yang tidak signifikan dan terdapat 2 variabel yang signifikan. Sehingga dilakukan uji signifikansi parsial kedua dengan hanya menggunakan 2 variabel yang signifikan yaitu variabel x_1 dan x_6 dalam pengujian. Berikut adalah hasil uji signifikansi parsial kedua.

Tabel 4.15 Uji Wald Kedua

Variabel	Estimasi	SE	Wald	P-value	Keputusan
Konstanta	6,540	1,558	4,196	0,00002	
x_1	-0,0646	0,0254	-2,543	0,011*	Tolak H_0
x_6 (1)	0,915	1,121	0,816	0,414	Terima H_0
x_6 (2)	15,522	1465,375	0,011	0,991	Terima H_0
x_6 (3)	0,976	0,869	1,123	0,261	Terima H_0
x_6 (4)	-1,456	0,742	-1,962	0,049*	Tolak H_0
x_6 (5)	16,485	3596,456	0,005	0,996	Terima H_0

*: Signifikan terhadap α (0,05)

Tabel 4.15 menunjukkan hasil uji signifikansi parameter secara parsial kedua hanya menggunakan variabel variabel x_1 dan x_6 .

- **Variabel x_1 (Usia Kepala Rumah Tangga)**

Hasil uji Wald kedua pada variabel x_1 .

$$\text{Statistik uji: } W_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{-0,0646}{0,0254} = -2,543$$

Keputusan: Tolak H_0 karena $|W_1| = |-2,543| = 2,543 > Z_{\alpha/2} = 1,96$ atau nilai $p\text{-value} = 0,11 < \alpha = 0,05$

Kesimpulan: Variabel x_1 usia kepala rumah tangga memberikan pengaruh signifikan secara parsial terhadap variabel status kesejahteraan rumah tangga.

- **Variabel x_6 (Status Pekerjaan Kepala Rumah Tangga)**

Pada uji wald kedua, variabel x_6 menghasilkan nilai yang signifikan pada variabel $x_{6(4)}$ yaitu variabel status pekerjaan kategori 4 yaitu buruh/karyawan/pegawai dengan hasil nilai:

$$\text{Statistik uji: } W_{6(4)} = \frac{\hat{\beta}_{6(4)}}{SE(\hat{\beta}_{6(4)})} = \frac{-1,456}{0,742} = -1,962$$

Keputusan: $|W_{6(4)}| = |-1,962| = 1,962 > Z_{\alpha/2} = 1,96$ dan nilai $p\text{-value} = 0,049 < \alpha = 0,05$, maka keputusannya adalah tolak H_0 .

Kesimpulan: Variabel $x_{6(4)}$ status pekerjaan kepala rumah tangga kategori buruh/karyawan/pabrik memberikan pengaruh signifikan secara parsial terhadap variabel status kesejahteraan rumah tangga.

Sedangkan, variabel x_6 kategori yang lain, yaitu status pekerjaan kategori berusaha sendiri, berusaha dibantu buruh tetap, berusaha dibantu buruh tidak tetap, dan pekerja bebas tidak memberikan pengaruh secara parsial

terhadap variabel status kesejahteraan rumah tangga. Karena menghasilkan nilai Wald yang lebih kecil dari nilai $Z_{\alpha/2}$ dan p -value lebih besar dari nilai α .

2. Uji Signifikansi Serentak

Uji serentak dilakukan untuk mengetahui signifikansi parameter terhadap variabel respon secara serentak atau menyeluruh. Pada pengujian ini digunakan uji G atau uji rasio *likelihood* maksimum. Berikut merupakan hasil pengujian signifikansi parameter secara serentak.

Hipotesis pengujian:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = 0$$

(Variabel prediktor tidak memberikan pengaruh signifikan terhadap variabel respon)

$$H_1 : \text{Minimal terdapat satu } \beta_i \neq 0 \text{ dengan } i = 1, 2, 3, \dots, 10$$

(Minimal terdapat satu variabel prediktor yang memberikan pengaruh signifikan terhadap variabel respon)

Taraf signifikansi: $\alpha = 0,05$

Daerah keputusan: Jika $G > \chi^2_{(\alpha; db)}$, maka tolak H_0

Tabel 4.16 Uji G

	G (-2LogLik)	db	$\chi^2_{(\alpha; db)}$	P-value
Model 1	70,968	20	$\chi^2_{(0,05;20)} = 31,41$	0,000
Model 2	86,768	7	$\chi^2_{(0,05;7)} = 14,067$	0,000

Berdasarkan Tabel 4.16 dapat diketahui bahwa model pertama menghasilkan nilai uji rasio *likelihood* atau nilai $G = 70,968 > \chi^2_{(0,05;20)} = 31,41$ dan model kedua menghasilkan nilai $G = 86,768 > \chi^2_{(0,05;7)} = 14,067$. Kedua

model menghasilkan nilai $p\text{-value} = 0,000 < \alpha = 0,05$. Maka diperoleh keputusan tolak H_0 . Sehingga dapat disimpulkan bahwa minimal terdapat satu variabel prediktor atau secara serentak memberikan pengaruh signifikan terhadap variabel respon status kesejahteraan rumah tangga.

4.2.4 Model Regresi Logistik Biner

Berdasarkan pengujian signifikansi parameter secara parsial pada Tabel 4.15 dapat diketahui bahwa hasil pengujian signifikansi parameter hanya signifikan terhadap variabel x_1 dan x_6 . Sehingga diperoleh nilai koefisien regresi sebagai berikut:

Tabel 4.17 Koefisien Regresi

Variabel	$\hat{\beta}$
Konstanta	6,540
x_1	-0,064
x_6 (1)	0,915
x_6 (2)	15,522
x_6 (3)	0,976
x_6 (4)	-1,456
x_6 (5)	16,485

Berikut merupakan model regresi logistik biner dalam bentuk logit.

$$g(x) = 6,540 - 0,064 x_1 + 0,915 x_{6(1)} + 15,522 x_{6(2)} + 0,976 x_{6(3)} - 1,456 x_{6(4)} + 16,486 x_{6(5)}$$

Maka diperoleh model akhir regresi logistik biner sebagai berikut:

$$\pi(x_i) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

$$\pi(x_i) = \frac{e^{(6,540 - 0,064x_1 + 0,915x_6(1) + 15,522x_6(2) + 0,976x_6(3) - 1,456x_6(4) + 16,485x_6(5))}}{1 + e^{(6,540 - 0,064x_1 + 0,915x_6(1) + 15,522x_6(2) + 0,976x_6(3) - 1,456x_6(4) + 16,485x_6(5))}}$$

Berdasarkan model yang terbentuk, dapat diketahui bahwa variabel yang berpengaruh positif terhadap kecenderungan rumah tangga menjadi miskin diantaranya status pekerjaan berusaha sendiri ($x_{6(1)}$), berusaha dibantu buruh tidak tetap ($x_{6(2)}$) maupun berusaha sendiri dibantu buruh tetap ($x_{6(3)}$) dan pekerja bebas ($x_{6(5)}$). Artinya semakin tinggi KRT yang bekerja pada status pekerjaan tersebut maka semakin tinggi pula kecenderungan rumah tangga menjadi miskin. Variabel yang berpengaruh negatif terhadap kecenderungan rumah tangga menjadi miskin diantaranya adalah variabel usia kepala rumah tangga (x_1) dan status pekerjaan sebagai buruh/karyawan/pegawai ($x_{6(4)}$). Interpretasinya adalah jika semakin tinggi usia KRT dan semakin tingginya KRT yang bekerja sebagai buruh/karyawan/pegawai, maka semakin rendah kecenderungan rumah tangga menjadi miskin dengan kata lain akan cenderung tinggi menjadi rumah tangga tidak miskin atau sejahtera. Sehingga dapat dikatakan bahwa hasil model regresi logistik biner memperoleh variabel atau faktor yang mempengaruhi status kesejahteraan rumah tangga di Kota Batu adalah variabel usia KRT dan status pekerjaan KRT sebagai buruh/karyawan/pegawai.

4.2.5 Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk mengetahui apakah terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan nilai hasil prediksi model. Pengujian ini dilakukan dengan menggunakan uji Hosmer dan Lemeshow. Berikut merupakan hasil dari uji kesesuaian model yang diperoleh.

Hipotesis pengujian:

H_0 : Tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan

kemungkinan hasil prediksi model (model yang didapat sesuai).

H_1 : Terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model (model yang didapat tidak sesuai).

Taraf signifikansi: $\alpha = 0,05$

Daerah keputusan: Jika $\hat{C} > \chi^2_{(\alpha, db)}$ atau $p\text{-value} < \alpha$, maka tolak H_0

Tabel 4.18 Uji Hosmer dan Lemeshow

$\hat{C}$	db	$\chi^2_{(0,05;8)}$	P-value
9,8139	8	15,507	0,278

Berdasarkan nilai uji Hosmer dan Lemeshow dapat diketahui bahwa nilai $\hat{C}$ sebesar 9,8139 kurang dari nilai *chi-square* tabel 15,507 dan nilai *p-value* 0,278 lebih besar daripada taraf signifikansi α (0,05), artinya gagal tolak H_0 . Sehingga dapat disimpulkan bahwa tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model atau model yang didapat sesuai.

4.2.6 Interpretasi Koefisien Regresi Logistik

Interpretasi koefisien untuk model regresi logistik biner dapat dilakukan dengan menggunakan nilai *odds ratio*. *Odds ratio* merupakan nilai kecenderungan antara satu kategori dengan kategori yang lain pada variabel prediktor yang kualitatif. Berikut adalah perhitungan nilai odds ratio untuk masing-masing variabel yang signifikan terhadap model.

$$\text{Odds Ratio} = \psi_i = e^{(\hat{\beta}_i)}$$

Tabel 4.19 Odds Ratio

Variabel	$\hat{\beta}$	Odds Ratio
x_1	-0,064	$\psi_1 = e^{(-0,064)} = 0,938$
x_6 (4)	-1,456	$\psi_{6(4)} = e^{(-1,456)} = 0,233$

Tabel 4.19 menunjukkan nilai *odds ratio* yang berguna untuk mengetahui besarnya pengaruh variabel prediktor terhadap status kesejahteraan rumah tangga. Interpretasi yang bisa diperoleh adalah sebagai berikut:

- Variabel x_1 (Usia Kepala Rumah Tangga)

Variabel usia KRT memiliki nilai *odds ratio* sebesar 0,938. Hal ini menunjukkan bahwa rumah tangga dengan usia kepala rumah tangga yang lebih muda mempunyai kecenderungan untuk menjadi rumah tangga tidak miskin 0,938 kali dibandingkan dengan kepala rumah tangga yang berusia lebih tua.

- Variabel $x_{6(4)}$ (Status Pekerjaan KRT Kategori Buruh/Karyawan/Pegawai)

Variabel status pekerjaan buruh/karyawan/pegawai memiliki nilai *odds ratio* sebesar 0,233. Hal ini menunjukkan bahwa status pekerjaan buruh/karyawan/pegawai cenderung menjadi rumah tangga tidak miskin 0,233 kali dari status pekerjaan yang lain.

4.2.7 Ketepatan Klasifikasi Regresi Logistik Biner

Ketepatan klasifikasi model dihitung menggunakan data *testing*. Berikut adalah hasil ketepatan klasifikasi dari hasil observasi dan taksiran klasifikasi dari model faktor yang mempengaruhi status kesejahteraan rumah tangga dalam bentuk tabel berikut.

Tabel 4.20 Ketepatan Klasifikasi Model

Hasil Observasi	Taksiran Klasifikasi		Total
	Miskin	Tidak Miskin	
Miskin	0	1	1
Tidak Miskin	4	89	93
Total	4	90	94

Nilai ketepatan klasifikasi menggunakan metode regresi logistik biner dihitung dengan menggunakan APER seperti berikut:

$$\text{APER}(\%) = \left(\frac{4+1}{94} \right) \times 100\% = 0,053 \times 100\% = 5,31\%$$

$$\text{Ketepatan Klasifikasi} = 100\% - 5,31\% = 94,69\%$$

Berdasarkan perhitungan nilai APER, perhitungan nilai APER ini adalah proporsi pengamatan yang diprediksi salah oleh fungsi klasifikasi dengan n sebagai total pengamatan. Diperoleh nilai APER atau nilai kesalahan klasifikasi 5,31% maka didapat nilai ketepatan klasifikasi sebesar 94,69%. Artinya bahwa model regresi logistik biner mampu memprediksi dengan tepat pengamatan sebesar 94,69%.

4.3 Analisis *Classification and Regression Tree* (CART)

Analisis CART diawali dengan langkah pembentukan pohon klasifikasi yang maksimal, kemudian dilakukan pemangkasan sampai diperoleh pohon klasifikasi yang optimal. Berikut adalah hasil analisis CART dalam memodelkan klasifikasi faktor-faktor yang mempengaruhi status kesejahteraan rumah tangga di Kota Batu.

4.3.1 Pembentukan Pohon Klasifikasi Maksimal

Langkah awal yang dilakukan untuk membentuk pohon klasifikasi adalah dengan menentukan variabel pemilah dan nilai variabel (*threshold*). Variabel pemilah dan nilai *threshold* dipilih dari beberapa kemungkinan pemilah dari masing-masing variabel prediktor. Perhitungan banyaknya kemungkinan pemilah disajikan dalam tabel berikut.

Tabel 4.21 Perhitungan Pemilah Setiap Variabel

X	Variabel	Skala	Banyak Kategori	Kemungkinan Pemilah
x_1	Usia KRT	Rasio	61	$61 - 1 = 60$
x_2	Jenis Kelamin	Nominal	2	$2^{2-1} - 1 = 1$
x_3	Status Perkawinan	Nominal	4	$2^{4-1} - 1 = 7$
x_4	Banyak Anggota RT	Rasio	12	$12 - 1 = 11$
x_5	Tingkat Pendidikan	Ordinal	4	$2^{4-1} - 1 = 7$
x_6	Status Pekerjaan	Nominal	6	$2^{6-1} - 1 = 31$
x_7	Kesehatan	Nominal	2	$2^{2-1} - 1 = 1$
x_8	Kepemilikan Tempat Tinggal	Nominal	2	$2^{2-1} - 1 = 1$
x_9	Bahan Bakar Memasak	Nominal	3	$2^{3-1} - 1 = 3$
x_{10}	Kepemilikan HP	Nominal	2	$2^{2-1} - 1 = 1$

Selanjutnya adalah menghitung nilai indeks Gini untuk mendapatkan nilai keheterogenan simpul. Indeks Gini melakukan pemilahan simpul pada masing-masing simpul kanan dan kiri. Kemudian, nilai indeks Gini digunakan sebagai penentu *goodness of split* dari masing-masing pemilah. Pemilah yang terpilih adalah variabel pemilah dengan nilai variabel yang memiliki nilai *goodness of split* paling maksimal. Berikut adalah besarnya kontribusi variabel sebagai pemilah terbaik pada pohon klasifikasi maksimal yang terbentuk yang ditunjukkan pada Tabel 4.22.

Tabel 4.22 Urutan Skor Variabel Terpenting

X	Variabel	Nilai Variabel	
x_1	Usia KRT	11,337	
x_4	Banyak ART	3,478	
x_6	Status Pekerjaan	2,579	
x_3	Status Perkawinan	1,958	

Lanjutan Tabel 4.22

x_2	Jenis Kelamin	1,910	
x_7	Kesehatan	1,86	
x_8	Kepemilikan Tempat Tinggal	0,638	
x_{10}	Kepemilikan HP	0,333	
x_5	Tingkat Pendidikan	0,122	

Tabel 4.22 menunjukkan bahwa terdapat 9 variabel prediktor yang menjadi pembangun pohon klasifikasi. Berdasarkan nilai variabel yang dihasilkan dapat diketahui bahwa variabel terpenting dan menjadi pemilah utama adalah variabel usia kepala rumah tangga (x_1) dengan nilai tertinggi yaitu 11,337. Variabel pemilah berikutnya banyaknya anggota rumah tangga (x_4) dengan nilai 3,478, status pekerjaan KRT (x_6) dengan nilai 2,579. Selanjutnya, status perkawinan (x_3), jenis kelamin KRT (x_2), kesehatan KRT selama satu bulan terakhir (x_7), status kepemilikan tempat tinggal (x_8), status kepemilikan HP (x_{10}), dan terakhir status pendidikan tertinggi dari KRT (x_5). Sedangkan, dari 10 variabel prediktor yang digunakan terdapat satu variabel yang tidak termasuk pembangun pohon klasifikasi yaitu variabel bahan bakar memasak (x_9).

Setelah mendapatkan pemilah terbaik yaitu variabel permilah utama x_1 , maka pada simpul utama sebagai simpul 0 terdapat 370 objek dari data *training* akan dipilah untuk memecah simpul menjadi dua buah simpul yaitu simpul kiri (t_L) dan simpul kanan (t_R). Simpul t_L terbentuk karena kriteria variabel usia KRT ≥ 66 tahun. Simpul t_R terbentuk karena kriteria variabel usia dari KRT yang memiliki usia < 66 tahun.

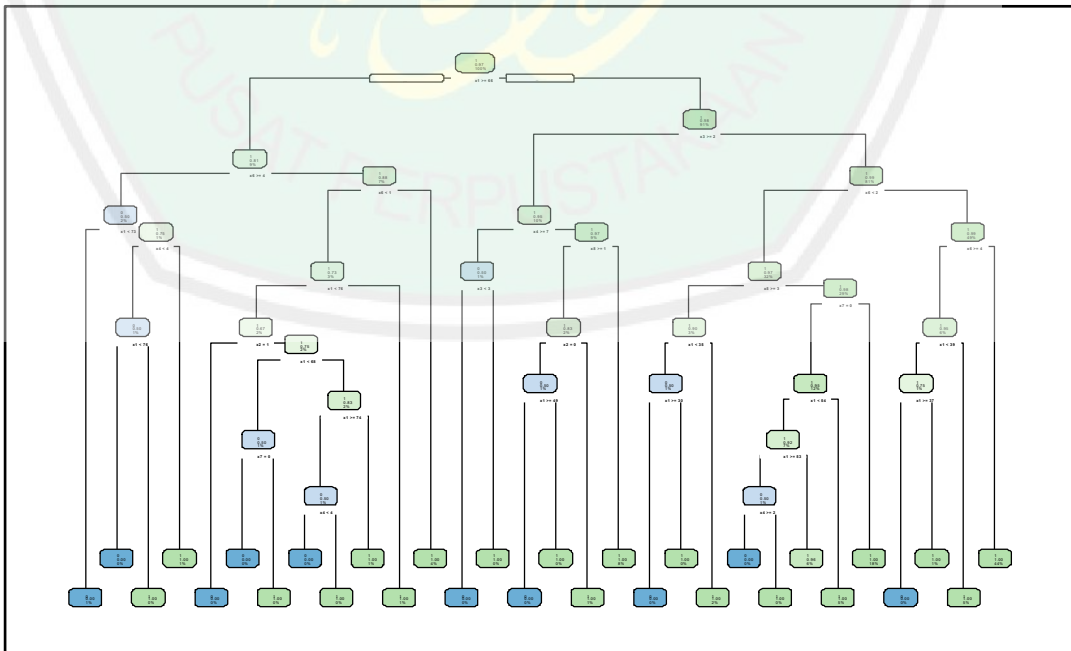
Pemberian label kelas pada simpul-simpul yang telah terbentuk berdasarkan aturan penetapan kelas jika $p(j_0|t) = \max_j p(j|t)$ dengan j adalah kelas pada simpul t .

$$p(\text{Miskin}|t) = \frac{12}{370} = 0,032$$

$$p(\text{Tidak Miskin}|t) = \frac{358}{370} = 0,968$$

Sehingga simpul diberi label kelas “Faktor Kesejahteraan Rumah Tangga”, karena peluang kelas kategori tidak miskin (rumah tangga sejahtera) lebih besar dibandingkan kelas dengan kategori miskin. Proses pelabelan kelas ini berlaku untuk semua simpul terutama pada simpul terminal. Hal tersebut karena simpul terminal sangat penting dalam memprediksi objek pada kelas tertentu apabila objek berada pada simpul terminal tersebut.

Konstruksi pohon klasifikasi maksimal berdasarkan variabel prediktor yang membangun pohon klasifikasi status kesejahteraan rumah tangga di Kota Batu ditunjukkan pada Gambar 4.10.



Gambar 4.10 Pohon Klasifikasi Maksimal

Pohon klasifikasi maksimal adalah hasil pemilahan secara biner yang dilakukan secara berulang-ulang pada data pengamatan yang menghasilkan pohon klasifikasi berukuran besar dan tingkat kedalaman yang tinggi. Gambar 4.10 menunjukkan hasil pohon klasifikasi maksimal (pohon T_{max}) yang berukuran besar dengan banyaknya simpul 59 *nodes*, terdiri dari 29 simpul induk dan 30 simpul terminal atau simpul akhir. Pohon maksimal yang diperoleh memiliki tingkat kedalaman sebesar 11.

4.3.2 Pemangkasan Pohon Klasifikasi Maksimal

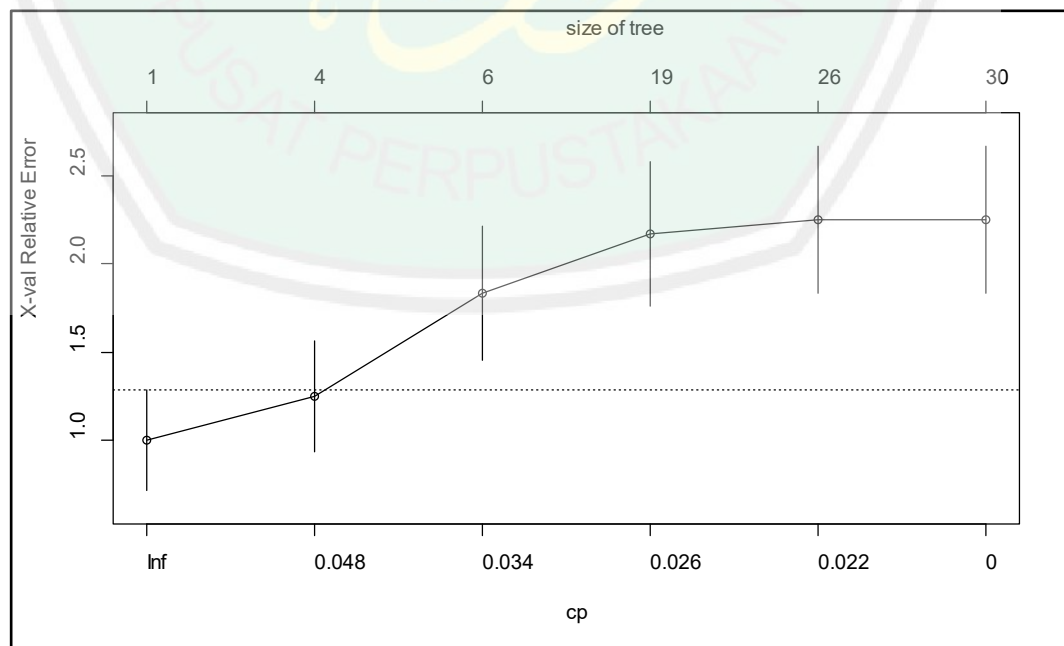
Pemangkasan atau *pruning* dilakukan untuk mengurangi kompleksitas pohon agar menjadi lebih sederhana sehingga terbentuk pohon klasifikasi optimal. Pohon yang baik dapat diperoleh dengan melakukan pemangkasan berdasarkan *cost complexity minimum*. Dilihat berdasarkan nilai *complexity parameter* yang mampu meminimalkan nilai *cross validation error* (CV Error). Nilai *complexity parameter* digunakan untuk mengontrol ukuran pohon keputusan dan untuk memilih ukuran pohon optimal. *Relative error* (Rel Error) atau kesalahan relatif $1 - R^2$ *root mean square error* adalah nilai kesalahan untuk prediksi data yang digunakan dalam memperkirakan model. Berikut adalah tabel nilai *complexity parameter* pada pohon maksimal T_{max} .

Tabel 4.23 Nilai Complexity Parameter

	CP	N Split	Rel Error	CV Error	SE
1	0,0555	0	1,000000	1,0000	0,28396
2	0,0416	3	0,833333	1,2500	0,31614
3	0,0277	5	0,750000	2,8333	0,37907
4	0,0238	18	0,333333	2,1667	0,40972
5	0,0208	25	0,166667	2,2500	0,41691

6	0,0000	29	0,083333	2,2500	0,41691
---	--------	----	----------	--------	---------

Tabel nilai *complexity parameter* menunjukkan nilai CP dari pohon yang terkecil (tanpa pemecahan) sampai pohon yang terbesar (29 pemecahan). Setiap baris dalam tabel mewakili kedalaman (*depth*) pohon yang berbeda. Semakin tinggi tingkatan pohon maka akan menghasilkan nilai kesalahan klasifikasi yang semakin rendah. *Cross validation error* biasanya meningkat ketika pohon yang terbangun berukuran besar. Sehingga tabel nilai CP memudahkan untuk membandingkan dan menentukan keputusan tentang tempat pemangkasan pohon klasifikasi maksimal menjadi pohon yang optimal. Berdasarkan Tabel 4.23, diperoleh nilai *complexity parameter* untuk setiap split berturut-turut adalah 0,0555; 0,0416; 0,0277; 0,0238; 0,0208; 0,000000. Ditunjukkan pula untuk nilai - *cross validation error* untuk setiap split berturut-turut adalah 1,0000; 1,2500; 2,8333; 2,1667; 2,2500; 2,2500. Berdasarkan nilai *cross validation error* yang terkecil adalah 1,0000 sehingga didapatkan nilai *complexity parameter* yang digunakan untuk pemangkasan pohon adalah 0,0555.

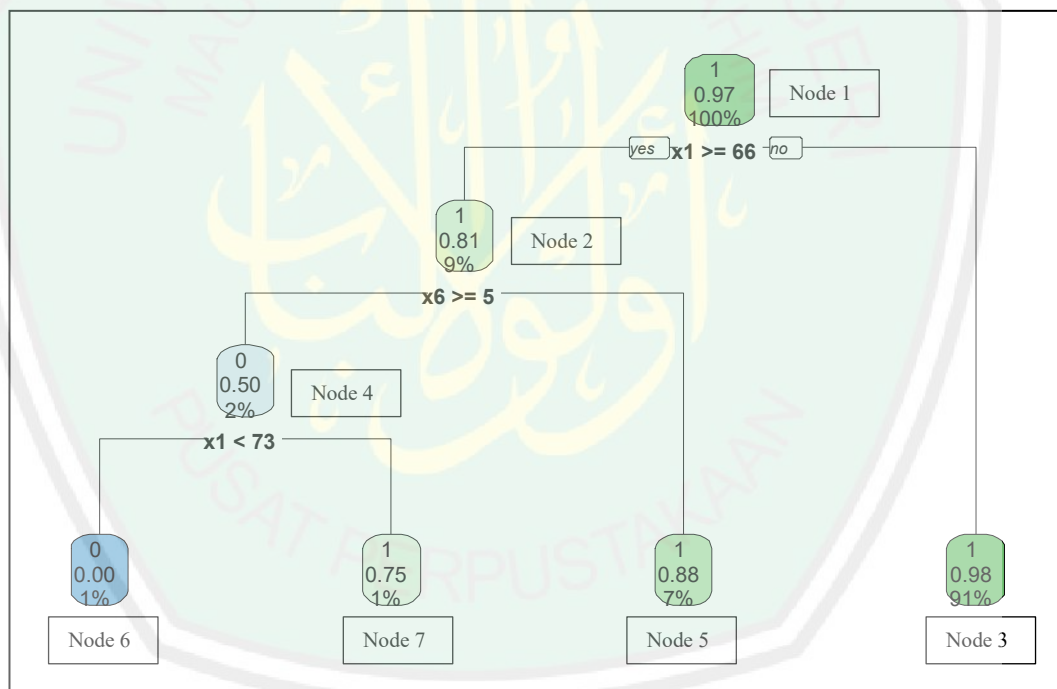


Gambar 4.11 Plot *Complexity Parameter*

Penentuan nilai *complexity parameter* untuk pemangkasan pohon dapat dilihat berdasarkan plot nilai *complexity parameter*. Gambar 4.11 menyajikan plot nilai *complexity parameter* dengan nilai kesalahan relatif (*Rel Error*), dapat dilihat bahwa nilai kesalahan klasifikasi (*Rel Error*) minimum berada pada nilai CP yang diatas 0,048 dengan ukuran pohon 4 simpul terminal. Sehingga dapat diambil nilai *complexity parameter* yaitu 0,0555 untuk pemangkasan pohon.

4.3.3 Pohon Klasifikasi Optimal

Pohon klasifikasi optimal diperoleh setelah proses pemangkasan pohon klasifikasi maksimal digambarkan sebagai berikut.



Gambar 4.12 Pohon Klasifikasi Optimal

Berdasarkan struktur pohon pada gambar 4.12 dapat diketahui bahwa simpul terminal yang dihasilkan pada pohon klasifikasi optimal sebanyak 4 simpul terminal dan terdapat 3 simpul induk. Variabel yang paling menentukan klasifikasi status kesejahteraan rumah tangga di Kota Batu adalah variabel x_1 usia

kepala rumah tangga. Selain itu, terdapat variabel lain yang juga mempengaruhi pengklasifikasian yaitu variabel x_6 status pekerjaan KRT. Urutan skor variabel dari kedua variabel tersebut disajikan pada Tabel 4.23.

Tabel 4.24 Skor Variabel Pohon Optimal

X	Variabel	Nilai Variabel	
x_1	Usia KRT	3,1846	
x_6	Status Pekerjaan	1,4949	

Tabel 4.24 menunjukkan bahwa variabel usia kepala rumah tangga (x_1) merupakan variabel pemilah utama dan paling menentukan klasifikasi faktor kesejahteraan rumah tangga di Kota Batu dengan nilai tertinggi 3,1846. Selanjutnya, variabel lain yang berkontribusi dalam pembentukan pohon klasifikasi adalah variabel status pekerjaan KRT (x_6) dengan nilai 1,4949.

Proses pemilahan diawali dengan memilah data sebanyak 370 rumah tangga berdasarkan peluang kelas miskin dan tidak miskin. Kelas miskin memperoleh peluang 0,03 ($n = 12$) sedangkan peluang untuk kelas tidak miskin adalah 0,97 ($n = 358$). Karena kelas tidak miskin bernilai lebih tinggi maka simpul dilabeli kelas tidak miskin sebagai simpul 1. Selanjutnya, simpul 1 dipilah lagi berdasarkan variabel usia kepala rumah tangga (x_1) dengan nilai *threshold* variabelnya adalah 66 tahun. Kemudian, data terbagi menjadi dua bagian yaitu simpul kiri (simpul 2) rumah tangga dengan usia kepala rumah tangga lebih dari atau sama dengan 66 tahun dengan peluang 0,81 untuk kelas tidak miskin dan peluang 0,19 untuk kelas miskin. Maka simpul dilabeli sebagai rumah tangga kelas tidak miskin. Sedangkan, simpul kanan (simpul 3) terdiri dari 338 rumah tangga yang usia kepala rumah tangganya dibawah 66 tahun yang terdiri dari kelas tidak miskin 0,98 sehingga peluang kelas miskin 0,02 sehingga dilabeli

kelas tidak miskin. Selanjutnya, setiap simpul dilakukan pemilahan kembali karena masih mempunyai tingkat keherogenan yang tinggi. Pemilahan dilakukan sampai memperoleh simpul terminal sebanyak 4 simpul. Simpul 2 dipilah kembali berdasarkan variabel pemilah x_6 dengan nilai *threshold* kategori 5 sehingga terbentuk simpul 4 (simpul kiri) dan simpul 5 (simpul kanan). Kemudian, simpul 4 berdasarkan variabel pemilah usia KRT (x_1) dengan nilai *threshold* adalah 73 tahun dipilah kembali menjadi simpul 6 (simpul kiri) dan simpul 7 (simpul kanan).

Berdasarkan Gambar 4.12 dapat diketahui bahwa terdapat 4 simpul terminal dengan nilai kedalaman 3. Karakteristik simpul terminal yang mengklasifikasikan data status kesejahteraan rumah tangga dapat dijelaskan sebagai berikut.

1. Simpul 3 diprediksi sebagai rumah tangga tidak miskin. Karakteristik rumah tangga pada simpul ini adalah variabel usia kepala rumah tangga yang berusia dibawah 66 tahun dengan peluang untuk menjadi kelas tidak miskin 0,98.
2. Simpul 5 diprediksi sebagai rumah tangga tidak miskin. Karakteristik rumah tangga pada simpul ini adalah rumah tangga dengan kepala rumah tangga yang status pekerjaannya adalah berusaha sendiri, berusaha sendiri baik dibantu oleh buruh tidak tetap, berusaha sendiri dibantu oleh buruh tetap, dan buruh/karyawan/pegawai dengan peluang 0,88.
3. Simpul 6 diprediksi sebagai rumah tangga miskin. Karakteristik pada simpul ini adalah rumah tangga dengan kepala rumah tangga yang usianya dibawah 73 tahun dengan peluang 1,00.

4. Simpul 7 diprediksi menjadi rumah tangga tidak miskin. Karakteristik rumah tangga pada simpul ini adalah rumah tangga dengan kepala rumah tangga yang usianya lebih dari atau sama dengan 73 tahun dengan peluang 0,00.

Karakteristik status kesejahteraan rumah tangga berdasarkan faktor-faktor yang mempengaruhi kesejahteraan rumah tangga menurut hasil 4 simpul terminal dari pohon optimal adalah variabel usia KRT (x_1) dengan usia kepala rumah tangga di bawah 66 dan di atas 73 tahun. Selain itu, variabel lain yang berpengaruh adalah variabel status pekerjaan (x_6), yaitu kepala rumah tangga yang mempunyai status pekerjaan berusaha sendiri, berusaha sendiri baik dibantu oleh buruh tidak tetap, berusaha sendiri dibantu oleh buruh tetap, dan buruh/karyawan/pegawai.

4.4.4 Ketepatan Klasifikasi CART

Ketepatan klasifikasi untuk pohon CART dihitung menggunakan data *testing* untuk menguji apakah pohon yang terbentuk oleh data *training* baik atau tidak dengan hasil ditunjukkan pada Tabel 4.24 berikut.

Tabel 4.25 Ketepatan Klasifikasi CART

Hasil Observasi	Taksiran Klasifikasi		Total
	Miskin	Tidak Miskin	
Miskin	0	6	6
Tidak Miskin	4	84	88
Total	4	90	94

Nilai ketepatan klasifikasi menggunakan metode *Classification and Regression Tree* (CART) dihitung dengan menggunakan APER seperti berikut:

$$\text{APER}(\%) = \left(\frac{4+6}{94} \right) \times 100\% = 0,1064 \times 100\% = 10,64\%$$

$$\text{Ketepatan Klasifikasi} = 100\% - 10,64\% = 89,36\%$$

Berdasarkan perhitungan nilai APER, perhitungan nilai APER ini adalah proporsi pengamatan yang diprediksi salah oleh fungsi klasifikasi dengan n sebagai total pengamatan. Diperoleh nilai APER atau nilai kesalahan klasifikasi penggunaan pohon CART adalah 10,64% maka didapat nilai ketepatan klasifikasi sebesar 89,36%. Artinya bahwa model CART mampu memprediksi tepat pengamatan sebesar 89,36%.

4.4 Perbandingan Nilai Ketepatan Klasifikasi

Setelah didapatkan model klasifikasi melalui dua metode yang digunakan, yaitu metode regresi logistik biner dan *Classification and Regression Tree* (CART) selanjutnya dapat diketahui hasil perbedaan dari kedua metode tersebut. Kedua metode tersebut menghasilkan perbedaan pada variabel yang signifikan terhadap model. Berdasarkan model regresi logistik biner dan CART, maka perbandingan variabel yang berpengaruh terhadap status kesejahteraan rumah tangga disajikan dalam Tabel 4.26 berikut.

Tabel 4.26 Perbandingan Variabel Signifikan

	RLB	CART
	x_1	x_1
Variabel	$x_{6(4)}$	$x_{6(1)}$
		$x_{6(2)}$
		$x_{6(3)}$
		$x_{6(4)}$

Selanjutnya dilakukan perbandingan model dari kedua metode untuk melihat model yang terbaik. Kriteria yang digunakan untuk membandingkan

adalah nilai kesalahan klasifikasi. Hasil hitung kesalahan (APER) dan ketepatan klasifikasi disajikan pada Tabel 4.27 berikut.

Tabel 4.27 Perbandingan Ketepatan Klasifikasi

Metode	APER(%)	Ketepatan Klasifikasi (%)
RLB	5,31	94,69
CART	10,64	89,36

Melalui Tabel 4.27 dapat dilihat bahwa nilai kesalahan klasifikasi untuk metode regresi logistik biner adalah 5,31% dengan nilai ketepatan klasifikasi 94,69%. Sedangkan, metode CART menghasilkan nilai kesalahan klasifikasi 10,64% dengan nilai ketepatan klasifikasi 89,36%. Dapat dikatakan bahwa nilai kesalahan klasifikasi regresi logistik biner lebih kecil dan nilai ketepatan klasifikasi lebih tinggi 5,33% dibandingkan nilai dari model CART.

Hasil perbandingan dari kedua metode memiliki perbedaan akurasi yang diamati antara dua pengklasifikasi bisa jadi tidak signifikan secara statistik. Sehingga dapat diperkuat dengan membandingkan kinerja pengklasifikasi menggunakan uji statistik selang kepercayaan atau *confidence interval*. Selang kepercayaan seringkali berguna untuk membandingkan kinerja pengklasifikasi yang berbeda untuk menentukan pengklasifikasi mana yang bekerja lebih baik pada kumpulan data tertentu. Perhitungan selang kepercayaan ini menggunakan nilai ketepatan klasifikasi yang telah diperoleh sebelumnya.

Selang kepercayaan untuk nilai ketepatan klasifikasi menggunakan persamaan (2.28) untuk metode regresi logistik biner dengan nilai akurasi 94,69%, banyaknya data *testing* 94 sampel dan nilai Z tabel = 1,96.

$$\begin{aligned}
& \frac{2 \cdot N \cdot acc_{rlb} + Z_{a/2}^2 \pm Z_{a/2} \sqrt{Z_{a/2}^2 + 4Nacc_{rlb} - 4Nacc_{rlb}^2}}{2(N + Z_{a/2}^2)} \\
&= \frac{2 \cdot 94 \cdot 0,9469 + 1,96^2 \pm 1,96 \sqrt{1,96^2 + 4 \cdot 94 \cdot 0,9469 - 4 \cdot 94 \cdot 0,9469^2}}{2(94 + 1,96^2)} \\
&= \frac{181,8588 \pm 9,347993236}{195,6832}
\end{aligned}$$

Selang kepercayaan atau akurasi kotor untuk nilai ketepatan klasifikasi regresi logistik biner berada pada selang 0,8816 – 0,9771, sehingga nilai ketepatan klasifikasi yang diperoleh berada pada selang tersebut. Kemudian selang kepercayaan untuk metode CART dengan nilai akurasi 89,36% seperti berikut:

$$\begin{aligned}
& \frac{2 \cdot N \cdot acc_{CART} + Z_{a/2}^2 \pm Z_{a/2} \sqrt{Z_{a/2}^2 + 4Nacc_{CART} - 4Nacc_{CART}^2}}{2(N + Z_{a/2}^2)} \\
&= \frac{2 \cdot 94 \cdot 0,8936 + 1,96^2 \pm 1,96 \sqrt{1,96^2 + 4 \cdot 94 \cdot 0,8936 - 4 \cdot 94 \cdot 0,8936^2}}{2(94 + 1,96^2)} \\
&= \frac{171,8384 \pm 12,33264}{195,6832}
\end{aligned}$$

Selang kepercayaan atau akurasi kotor untuk nilai ketepatan klasifikasi CART berada pada selang 0,8151 – 0,9412, sehingga nilai ketepatan klasifikasi yang diperoleh berada pada selang tersebut. Menurut hasil kedua selang tersebut, dapat dilihat bahwa selang kepercayaan bagi nilai ketepatan regresi logistik biner memiliki margin paling kecil dibandingkan selang CART. Artinya bahwa nilai interval yang semakin sempit memberikan nilai statistik yang semakin akurat. Sehingga dapat dikatakan bahwa model terbaik yang dapat digunakan untuk

memprediksi faktor dari status kesejahteraan rumah tangga di Kota Batu adalah metode regresi logistik biner.



BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan, maka dapat diperoleh kesimpulan sebagai berikut.

1. Model klasifikasi yang terbentuk dengan menggunakan 10 variabel sebagai faktor yang mempengaruhi status kesejahteraan rumah tangga menggunakan metode regresi logistik biner diperoleh dua variabel yang berpengaruh signifikan terhadap model. Model regresi logistik biner yang diperoleh dalam bentuk logit sebagai berikut:

$$g(x) = 6,540 - 0,064 x_1 + 0,915 x_{6(1)} + 15,522 x_{6(2)} + 0,976 x_{6(3)} \\ - 1,456 x_{6(4)} + 16,6486 x_{6(5)}$$

Sedangkan, metode CART (*Classification and Regression Tree*) membangun model menghasilkan pohon optimal dengan 4 simpul terminal. Variabel terpenting dalam menentukan dalam pemilahan simpul untuk klasifikasi adalah variabel x_1 yaitu usia kepala rumah tangga dan variabel x_6 yaitu status pekerjaan dari kepala rumah tangga.

2. Hasil yang diperoleh menggunakan metode regresi logistik biner, variabel-variabel yang teridentifikasi paling berpengaruh terhadap status kesejahteraan rumah tangga di Kota Batu adalah:
 - a. Usia kepala rumah tangga, variabel usia yang lebih muda cenderung menjadi rumah tangga tidak miskin 0,938 kali dibandingkan usia lebih tua.

- b. Status pekerjaan kepala rumah tangga kategori buruh/karyawan/pegawai, variabel status pekerjaan sebagai buruh/karyawan/pegawai cenderung menjadi rumah tangga tidak miskin 0,233 kali dibandingkan status pekerjaan kategori lain.

Metode CART memberikan hasil bahwa variabel yang paling berpengaruh pada klasifikasi status kesejahteraan rumah tangga di Kota Batu adalah:

- a. Usia kepala rumah tangga, variabel usia dibawah 66 tahun berpeluang menjadi kelas tidak miskin 0,98.
 - b. Status pekerjaan kepala rumah tangga, variabel status pekerjaan pada kategori berusaha sendiri, berusaha sendiri dibantu oleh buruh tidak tetap, berusaha sendiri dibantu oleh buruh tetap, dan sebagai buruh/karyawan/pegawai berpeluang menjadi kelas miskin 0,88.
3. Perbandingan hasil dari analisis kedua metode menunjukkan bahwa metode regresi logistik biner memiliki nilai ketepatan klasifikasi 94,69% sedangkan CART menghasilkan nilai ketepatan klasifikasi 89,36%. Jadi nilai ketepatan klasifikasi regresi logistik biner lebih tinggi 5,33% dibandingkan nilai dari model CART. Sehingga model terbaik yang dapat digunakan untuk memprediksi faktor kesejahteraan rumah tangga di Kota Batu adalah metode regresi logistik biner.

5.2 Saran

Adapun saran dari penelitian ini adalah sebagai berikut:

1. Dalam penelitian selanjutnyadapat menggunakan metode klasifikasi yang lain untuk membandingkan hasil klasifikasi status kesejahteraan rumah

tangga di Kota Batu dan mengembangkan faktor-faktor lain sebagai variabel yang mungkin.

2. Pengolahan menggunakan metode regresi logistik biner dan CART yang telah dilakukan didapatkan informasi bahwa usia KRT dan status pekerjaan KRT berpengaruh terhadap model status kesejahteraan rumah tangga di Kota Batu. Melalui gambaran deskriptif diketahui bahwa sebagian besar KRT bekerja sebagai buruh/karyawan/pegawai, oleh sebab itu perlu adanya usaha pemerintah dalam melakukan pemberdayaan masyarakat dan mengembangkan potensi masyarakat di berbagai sektor bidang kerja sehingga kedepannya mampu meningkatkan kesejahteraan hidup.

DAFTAR PUSTAKA

- Agresti, A. 1990. *An Introduction to Categorical Data Analysis*. New York: John Wiley & Sons Inc.
- Badan Pusat Statistik. *Survei Sosial dan Ekonomi Nasional*, (Online), (<http://www.silastik.bps.go.id>), diakses 25 Januari 2020.
- Badan Pusat Statistik. *Pengertian Kesejahteraan dan Kemiskinan*, (Online), (<http://www.bps.go.id>), diakses 1 Februari 2020.
- Badan Pusat Statistik. *Survei Sosial dan Ekonomi Nasional*, (Online), (<http://www.sirusa.bps.go.id>), diakses 5 Februari 2020.
- Badan Pusat Statistik Kota Batu. 2019. *Kota Batu dalam Angka 2019*. Batu: BPS Kota Batu.
- Badan Pusat Statistik Kota Batu. 2019. *Statistik Kesejahteraan Kota Batu 2019*. Batu: BPS Kota Batu.
- Bappenas. 2010. *Evaluasi Pelayanan Keluarga Berencana bagi Masyarakat Miskin (Keluarga Prasejahtera/KPS dan Keluarga Sejahtera-I/KS-I)*. Jakarta: Bappenas.
- Bappenas. 2018. *Analisis Wilayah dengan Kemiskinan Tinggi*. Jakarta: Bappenas.
- Breiman, L., Friedman, J. H., Olshen, R. A., dan Stone, C. J. 1998. *Classification And Regression Trees*. New York: Chapman & Hall/CRC.
- Cahyat, A., Gonner, C., Haug, M., 2007. *Mengkaji Kemiskinan dan Kesejahteraan Rumah Tangga Sebuah Panduan dengan Contoh dari Kutai Barat, Indonesia*. Bogor: Center of International Forestry Research.
- Fobia, Maria Elviana, dkk. 2020. *Perbandingan Metode Classification and Regression Tree (CART) dan Metode Regresi Logistik Biner dalam Mengklasifikasi Status Wanita Bekerja di Kota Kupang*. Jurnal Statistika Industri dan Komputasi. Vol. 5, No. 1.
- Hosmer, D.W dan Lemeshow, S. 2000. *Applied Logistic Regresion*. New York: John Wiley & Sons Inc.
- Johnson, Richard A; Wichern, Dean W. 2007. *Applied Multivariate Statistical Analysis*. Amerika Serikat: Pearson Prentice Hall.
- Lewis, R.J. 2000. *An Introduction to Classification And Regression Tree (CART) Analysis*. Annual Meeting of the Society For Academic Emergency

Medicine in San Fransisco. California: Department of Emergency Medicine.

Mahalisa, Galih. 2016. *Metode Pohon Gabungan pada CART untuk Analisa Kesejahteraan Rumah Tangga di Banjarmasin*. Jurnal Technologia. Vol. 7, No. 1.

Ramandhani, R., Sudarno, dan Safitri, D. 2017. *Metode Bootstrap Aggregating Regresi Logistik Biner untuk Ketepatan Klasifikasi Kesejahteraan Rumah Tangga di Kota Pati*. Jurnal Gaussian. Vol. 6, No. 1.

Steinberg dan Phillip, Colla. *CART: Tree-Structured Non-Parametric Data Analysis*. San Diego: Salford Systems.

Sulianto. 2011. *Ekonometrika Terapan – Teori dan Aplikasi dengan SPSS*. Yogyakarta: ANDI.

Sumartini, S. H dan Purnami, S. W. 2015. *Penggunaan Metode Classification and Regression Tree (CART) untuk Klasifikasi Rekurensi Pasien Kanker Serviks di RSUD Dr. Soetomo Surabaya*. Jurnal Sains dan Seni ITS. Vol. 4, No. 2.

Suryadarma, D., Akhamdi, Hastuti, dan Toyamah, N. 2005. *Ukuran Objektif Kesejahteraan Keluarga untuk Penargetan Kemiskinan: Hasil Uji Coba Sisrem Pemantauan Kesejahteraan oleh Masyarakat di Indonesia*. Jakarta: Lembaga Penelitian SEMERU.

Tanjung, R. H dan Kartiko. 2017. *Penerapan Metode CART (Classification and Regression Trees) untuk Menentukan Faktor-faktor yang Mempengaruhi Pembayaran Kredit oleh Nasabah (Studi Kasus Bank BRI Unit Aek Tarum-Sumatera Uatra)*. Jurnal Statistika Industri dan Komputasi. Vol. 2, No. 2.

Waluyo, Agung dkk. 2015. *Perbandingan Klasifikasi Nasabah Kredit Menggunakan Regresi Logistik Biner dan CART (Classification and Regression Trees)*. Media Satistika. Vol. 7, No. 2.

Yohannes, Yesihac; Hoddinott, John. 1999. *Classification And Regression Tree: An Introduction*. Washington, DC: International Food Policy Research Institut.

LAMPIRAN

Lampiran 1: Data Penelitian Hasil SUSENAS Kota Batu Tahun 2018

NO.	Pengeluaran Per Kapita Per Bulan	Usia KRT	Jenis Kelamin	Status Perkawinan	Banyak ART	Status Pendidikan	Status Pekerjaan	Kesehatan	Tempat Tinggal	Bahan Bakar Masak	Telepen Seluler
1	3153994,048	71	LK	Kawin	2	SD	Pkrj Klg	Tidak	Milik Sdr	Gas/Listrik	Tidak
2	1067349,206	52	LK	Kawin	6	SMP	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
3	967305,5556	46	LK	Kawin	3	SMP	Usaha Sdr	Tidak	Sewa	Gas/Listrik	Ya
4	577090,4762	30	LK	Belum	6	SMA	Brh/Kyw	Tidak	Sewa	Gas/Listrik	Ya
5	1893271,429	60	LK	Kawin	5	SMA	UDBT	Ya	Milik Sdr	Gas/Listrik	Ya
6	1249454,762	35	LK	Kawin	4	SMP	Usaha Sdr	Tidak	Milik Sdr	Gas/Listrik	Ya
7	870677,5794	39	LK	Kawin	6	PT	Usaha Sdr	Tidak	Lainnya	Gas/Listrik	Ya
8	1259815,476	63	LK	Cerai Mati	2	SD	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Tidak
9	995323,0595	54	LK	Kawin	4	SMP	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
10	572412,6984	45	LK	Kawin	6	SMA	Brh/Kyw	Tidak	Lainnya	Gas/Listrik	Ya
11	1110880,952	43	LK	Kawin	4	SMP	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
12	956275,4365	37	LK	Kawin	3	SMA	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
13	1411133,333	54	LK	Kawin	3	SMP	Usaha Sdr	Tidak	Milik Sdr	Gas/Listrik	Ya
14	1095068,452	49	LK	Kawin	4	SMP	Usaha Sdr	Tidak	Milik Sdr	Gas/Listrik	Tidak

15	1628133,333	63	LK	Kawin	5	SD	UDBTT	Tidak	Milik Sdr	Gas/Listrik	Ya
⋮											
⋮											
⋮											
85	392308,3333	70	PR	Cerai Mati	4	SD	Usaha Sdr	Tidak	Milik Sdr	Gas/Listrik	Tidak
86	1182185,119	31	LK	Kawin	4	SMA	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
87	405690,4762	50	PR	Cerai Hidup	7	SMA	Brh/Kyw	Tidak	Milik Sdr	Gas/Listrik	Ya
88	519482,1429	59	LK	Kawin	6	SD	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Ya
89	880413,6905	56	LK	Kawin	3	SMA	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Ya
90	1266352,063	53	LK	Kawin	3	SMP	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Tidak
91	935884,5238	44	LK	Kawin	1	SD	UDBTT	Tidak	Milik Sdr	Gas/Listrik	Ya
92	332676,1905	66	LK	Kawin	3	SD	Pkrj Bbs	Tidak	Milik Sdr	Gas/Listrik	Tidak
⋮											
⋮											
⋮											
460	1154216,667	52	LK	Kawin	5	PT	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Ya
461	716336,3095	36	LK	Kawin	3	SD	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Ya
462	522654,7619	56	LK	Kawin	2	SD	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Tidak
463	544373,0159	39	LK	Kawin	4	SD	Brh/Kyw	Ya	Milik Sdr	Gas/Listrik	Ya
464	1687744,048	39	PR	Cerai Mati	3	SMP	Usaha Sdr	Ya	Milik Sdr	Gas/Listrik	Ya

LIBRARY OF NAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF

[illegible]

⋮												
⋮												
85	Miskin	0	70	2	4	4	1	1	2	1	1	2
86	Tidak Miskin	1	31	1	2	4	3	4	2	1	1	1
87	Miskin	0	50	2	3	7	3	4	2	1	1	1
88	Tidak Miskin	1	59	1	2	6	1	4	1	1	1	1
89	Tidak Miskin	1	56	1	2	3	3	4	1	1	1	1
90	Tidak Miskin	1	53	1	2	3	2	4	1	1	1	2
91	Tidak Miskin	1	44	1	2	1	1	2	2	1	1	1
92	Miskin	0	66	1	2	3	1	5	2	1	1	2
⋮												
⋮												
⋮												
460	Tidak Miskin	1	52	1	2	5	4	4	1	1	1	1
461	Tidak Miskin	1	36	1	2	3	1	4	1	1	1	1
462	Tidak Miskin	1	56	1	2	2	1	4	1	1	1	2
463	Tidak Miskin	1	39	1	2	4	1	4	1	1	1	1
464	Tidak Miskin	1	39	2	4	3	2	1	1	1	1	1

Keterangan:

Y	: Status Kesejahteraan Rumah Tangga	X7	: Kesehatan KRT dalam 1 Bulan Terakhir
	1: Miskin		1: Ya
	2: Tidak Miskin		2: Tidak
X1	: Usia Kepala Rumah Tangga	X8	: Status Kepemilikan Tempat Tinggal
X2	: Jenis Kelamin Kepala Rumah Tangga		1: Milik sendiri
	1: Laki-laki		2: Kontrak/sewa
	2: Perempuan		3: Lainnya
X3	: Status Perkawinan Kepala Rumah Tangga	X9	: Bahan Bakar Memasak
	1: Belum kawin		1: Listrik/gas/elpiji
	2: Kawin		2: Minyak tanah
	3: Cerai hidup		3: Kayu
	4: Cerai mati		4: Lainnya
X4	: Banyak Anggota Rumah Tangga	X10	: Kepemilikan Telepon Seluler (HP)
X5	: Tingkat Pendidikan Tertinggi Kepala Rumah Tangga		1: Ya
	1: SD sederajat		2: Tidak
	2: SMP sederajat		
	3: SMA sederajat		
	4: Perguruan tinggi		
X6	: Status Pekerjaan Kepala Rumah Tangga		
	1: Berusaha sendiri		
	2: Berusaha dibantu buruh tidak tetap		
	3: Berusaha dibantu buruh tetap		
	4: Buruh/karyawan/ pegawai		
	5: Pekerjaan bebas		
	6: Pekerja keluarga		

Lampiran 3: Script dan Output

```
#Membuka File Data
> mydata <- read_excel("D:/data.xlsx")
> data=mydata
> head(data)
> data$y <- as.factor(data$y)
> data$x2 <- as.factor(data$x2)
> data$x3 <- as.factor(data$x3)
> data$x5 <- as.factor(data$x5)
> data$x6 <- as.factor(data$x6)
> data$x7 <- as.factor(data$x7)
> data$x8 <- as.factor(data$x8)
> data$x9 <- as.factor(data$x9)
> data$x10 <- as.factor(data$x10)
> str(data)
```

```
#Nama Miskin dan Tidak Miskin
> levels(data$y) = c("M","TM")
```

```
#Cek Kelas Bias
> nrow(data)
> table(data$y)
> prop.table(table(data$y))
```

```
#Stratifikasi Data Training 80% dan Testing 20%
> library(dplyr)
> m <- filter(data, y==0)
> tm <- filter(data, y==1)
> set.seed(100)
> acak.m <- sample(1:nrow(m), 0.8*nrow(m))
> acak.tm <- sample(1:nrow(tm), 0.8*nrow(tm))
> train80 <- rbind(m[acak.m,], tm[acak.tm,])
> test20 <- rbind(m[-acak.m,], tm[-acak.tm,])
```

```
#Regresi Logistik Biner
#Model 1
> rlb <- glm(y~. , data=train80, family=binomial)
> summary(rlb)

Call:
glm(formula = y ~ ., family = binomial, data=train80)
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.4390   0.0260   0.0914   0.2018   1.0936

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  23.45489  4629.50949   0.005  0.99596
x1          -0.10265   0.03942  -2.604  0.00922 **
x21           0.21311   1.54499   0.138  0.89029
x31          -14.13889  4629.50923  -0.003  0.99756
x32          -18.26487  4629.50934  -0.004  0.99685
x33          -15.03853  4629.50947  -0.003  0.99741
x4           -0.06797   0.23417  -0.290  0.77162
x51           1.38288   1.20168   1.151  0.24982
x52           0.31697   1.24865   0.254  0.79961
x53          -0.92210   1.30659  -0.706  0.48036
x61           0.80916   1.20391   0.672  0.50151
x62          18.27444  3621.88368   0.005  0.99597
x63           1.97536   1.13360   1.743  0.08141 .
x64          -2.24090   0.91273  -2.455  0.01408 *
x65          18.71764  9003.78796   0.002  0.99834
x71          -0.98108   0.79648  -1.232  0.21804
x81          -0.67661   1.07289  -0.631  0.52828
x82          17.18600  2789.42547   0.006  0.99508
```

```

x92          16.55631 5521.27404    0.003  0.99761
x101          1.23140    0.97639    1.261  0.20724
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 105.893  on 369  degrees of freedom
Residual deviance:  70.968  on 350  degrees of freedom
AIC: 110.97
Number of Fisher Scoring iterations: 19

```

```

#Uji serentak
> logLik(rlb)
'log Lik.' -35.48414 (df=20)

#Uji Multiko VIF
> library(car)
> vif(rlb)
      GVIF Df GVIF^(1/(2*Df))
x1  2.184737 1    1.478086
x2  3.881216 1    1.970080
x3  6.937744 3    1.381030
x4  1.280712 1    1.131686
x5  1.757749 3    1.098566
x6  2.421465 5    1.092466
x7  1.349464 1    1.161664
x8  1.425447 2    1.092667
x9  1.000000 1    1.000000
x10 2.095865 1    1.447710

#Uji Kesesuaian Model (Hosmer Lemeshow Test)
> library(ResourceSelection)
> hoslem.test(rlb$y, fitted(rlb), g=10)
    Hosmer and Lemeshow goodness of fit (GOF) test

data:  rlb$y, fitted(rlb)
X-squared = 7.9756, df = 8, p-value = 0.4359

```

```

#Model 2
> rlb2 <- glm(y~x1+x6, data=train80, family=binomial(link="logit"))
> summary(rlb2)
Call:
glm(formula = y ~ x1 + x6, family = binomial(link = "logit"),
    data = train80)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9732   0.1062   0.1763   0.2512   1.0785

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   6.54006    1.55861   4.196 2.72e-05 ***
x1            -0.06461    0.02541  -2.543  0.0110 *
x61            0.91589    1.12186   0.816  0.4143
x62           15.52239   1465.37588   0.011  0.9915
x63            0.97632    0.86965   1.123  0.2616
x64           -1.45682    0.74238  -1.962  0.0497 *
x65           16.48588   3596.45655   0.005  0.9963
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)
    Null deviance: 105.893  on 369  degrees of freedom
Residual deviance:  86.769  on 363  degrees of freedom

```

```

AIC: 100.77
Number of Fisher Scoring iterations: 17

#Uji serentak
> logLik(rlb2)
'log Lik.' -43.38465 (df=7)

#Hosmer Lemeshow Test
> library(ResourceSelection)
> hoslem.test(rlb2$y, fitted(rlb2), g=10)
    Hosmer and Lemeshow goodness of fit (GOF) test

data:  rlb2$y, fitted(rlb2)
X-squared = 9.8139, df = 8, p-value = 0.2783

```

```

#Odds Ratio Model 2
> exp(coef(rlb2))
      (Intercept)          x1          x61          x62          x63
6.923275e+02  9.374298e-01  2.498994e+00  5.511744e+06  2.654659e+00
          x64          x65
2.329758e-01  1.444530e+07

```

```

#Tabel APER
#Hasil Klasifikasi Data Testing
#Prediksi Data Testing
> prob.prediksi <- predict(rlb2, test20, type = "response")

#Hitung Score Optimal Minimizes Misclassification Error
> library(InformationValue)
> opt <- optimalCutoff(test20$y, prob.prediksi)
> opt
[1] 0.67
> library(InformationValue)
> confusionMatrix(test20$y, prob.prediksi, threshold = opt)
      M TM
0 0  1
1 4 89

```

```

#CART
#Membangun Pohon
> library(rpart)
> library(rpart.plot)
> fit=rpart(y~., method="class", data=train80, control=rpart.control(minsplit=1, minbucket=1, cp=0,000))
> printcp(fit) #display the results
Classification tree:
rpart(formula = y ~ ., data = train80, method = "class", control = rpart.control(minsplit = 1, minbucket = 1, cp = 0, 0))
Variables actually used in tree construction:
[1] x1 x2 x3 x4 x5 x6 x7 x8

Root node error: 12/370 = 0.032432
n= 370
      CP nsplit rel error xerror  xstd
1 0.055555      0 1.000000 1.0000 0.28396
2 0.041667      3 0.833333 1.2500 0.31614
3 0.027778      5 0.750000 1.8333 0.37907
4 0.023810     18 0.333333 2.1667 0.40972
5 0.020833     25 0.166667 2.2500 0.41691
6 0.000000     29 0.083333 2.2500 0.41691

> plotcp(fit) # visualize cross-validation results
> summary(fit) # detailed summary of splits

call:

```

```

rpart(formula = y ~ ., data = train80, method = "class", control
= rpart.control
(minsplit = 1, minbucket = 1, cp = 0, 0))
n= 370
      CP nsplit rel error  xerror  xstd
1 0.05555555    0 1.0000000 1.000000 0.2839553
2 0.04166667    3 0.8333333 1.250000 0.3161387
3 0.02777778    5 0.7500000 1.833333 0.3790695
4 0.02380952   18 0.3333333 2.166667 0.4097168
5 0.02083333   25 0.1666667 2.250000 0.4169143
6 0.00000000   29 0.0833333 2.250000 0.4169143

> fit$variable.importance
      x1      x4      x6      x3      x2      x7
11.3371509 3.4784913 2.5791949 1.9579956 1.9106396 1.8626520
      x8      x10     x5
0.6381463 0.3333333 0.1226041

#Plot Tree
> plot(fit, uniform=TRUE, main="Maximal Classification Tree")
> text(fit, use.n=TRUE, all=TRUE, cex=.8)
> rpart.plot(fit)

```

```

#Pemangkasan Pohon
> cp=fit$cpstable[which.min(fit$cpstable[, "xerror"]), "CP"]
> cp
[1] 0.0555555
> pfit<- prune(fit, cp = 0.05555)
> pfit
n= 370

node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 370 12 1 (0.03243243 0.96756757)
 2) x1>=65.5 32 6 1 (0.18750000 0.81250000)
   4) x6>=4.5 6 3 0 (0.50000000 0.50000000)
     6) x1< 73 2 0 0 (1.00000000 0.00000000) *
     7) x1>=73 4 1 1 (0.25000000 0.75000000) *
   5) x6< 4.5 26 3 1 (0.11538462 0.88461538) *
 3) x1< 65.5 338 6 1 (0.01775148 0.98224852) *

```

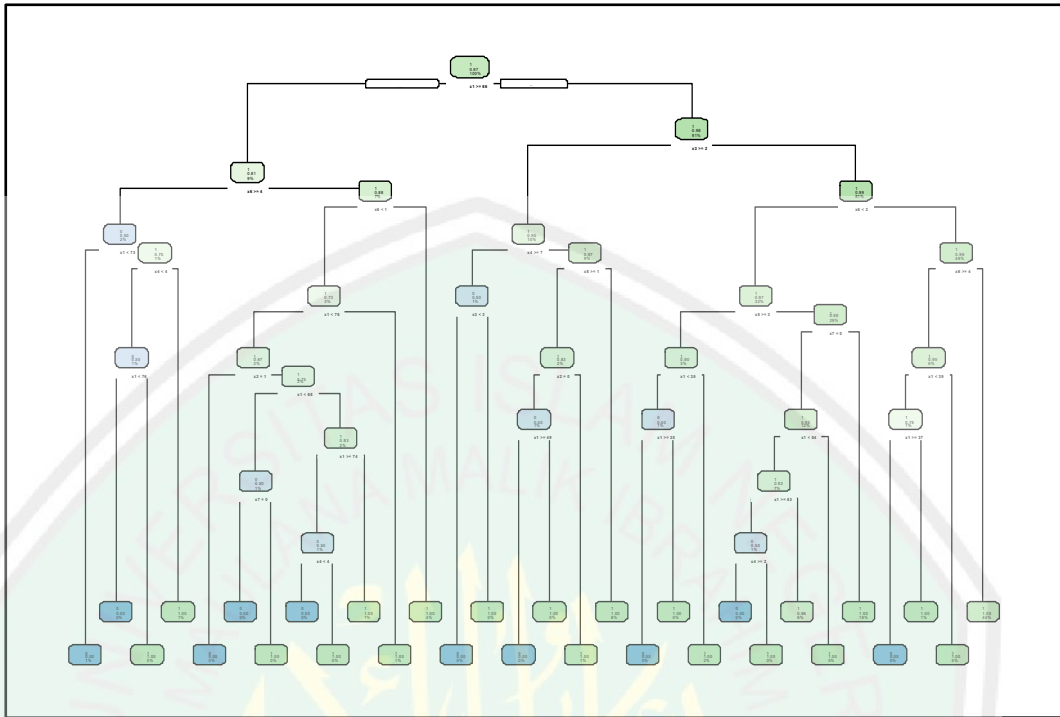
```

#ketepatan klasifikasi Data Testing
> model_pred2 = predict(fit, test20, type="class")
> table(predicted = model_pred2, actual=test20$y)
      actual
predicted 0 1
         0 0 6
         1 4 84

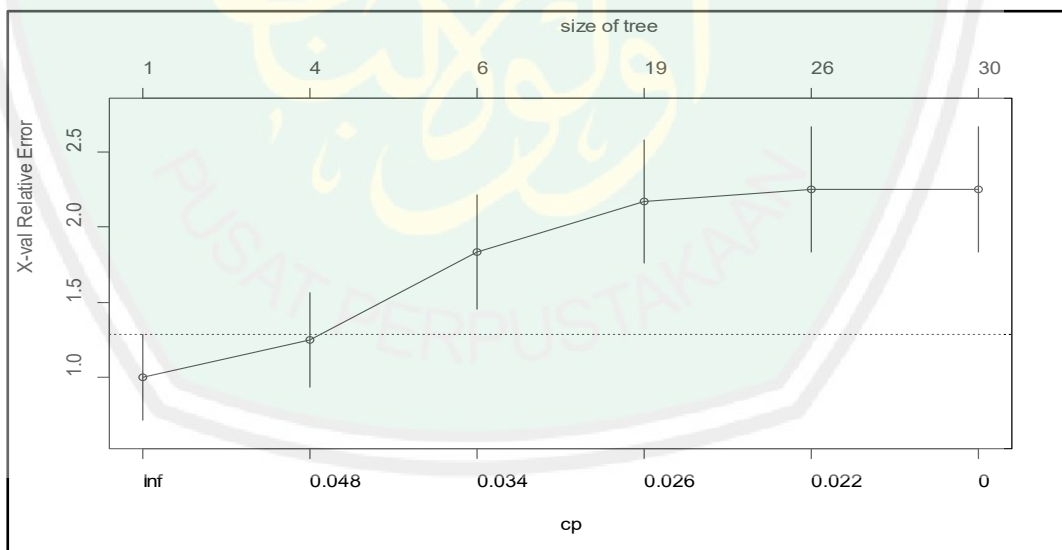
```

Lampiran 4: Output Plot

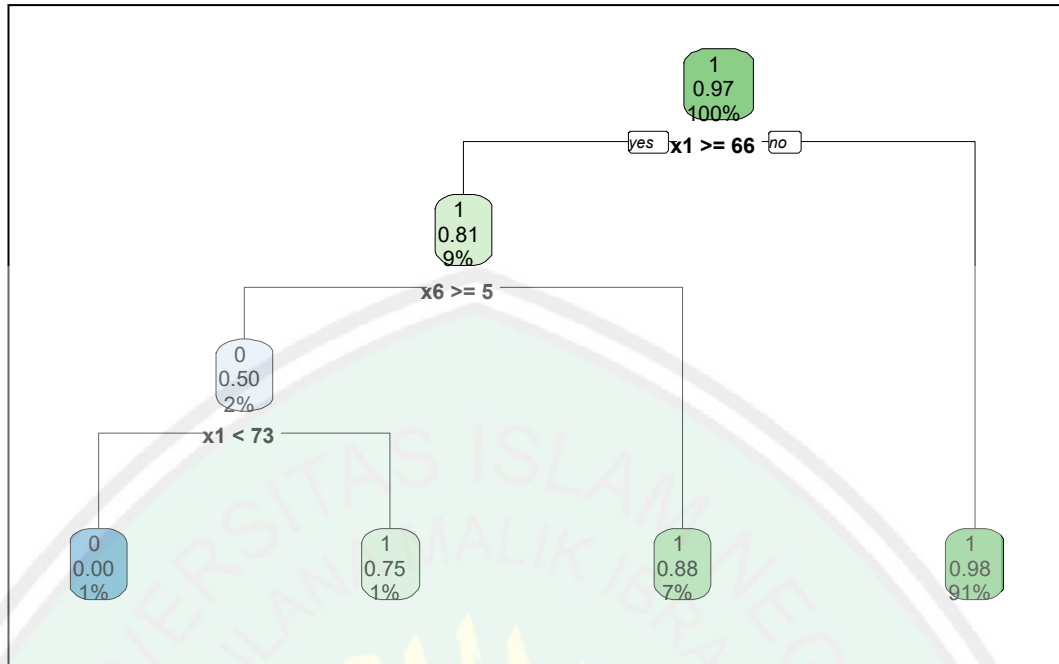
#Pohon Klasifikasi Maksimal



#Plot CV



#Pohon Klasifikasi Optimal



Lampiran 5: Tabel Chi Square dan Z

Tabel Chi Square ($\chi^2_{0,05;7} = 14,067$; $\chi^2_{0,05;8} = 15,507$; $\chi^2_{0,05;20} = 31,41$)

α		0.1	0.05	0.025	0.01	0.005
db	1	2.70554	3.84146	5.02390	6.63489	7.87940
	2	4.60518	5.99148	7.37778	9.21035	10.59653
	3	6.25139	7.81472	9.34840	11.34488	12.83807
	4	7.77943	9.48773	11.14326	13.27670	14.86017
	5	9.23635	11.07048	12.83249	15.08632	16.74965
	6	10.64464	12.59158	14.44935	16.81187	18.54751
	7	12.01703	14.06713	16.01277	18.47532	20.27774
	8	13.36156	15.50731	17.53454	20.09016	21.95486
	9	14.68366	16.91896	19.02278	21.66605	23.58927
	10	15.98717	18.30703	20.48320	23.20929	25.18805
	11	17.27501	19.67515	21.92002	24.72502	26.75686
	12	18.54934	21.02606	23.33666	26.21696	28.29966
	13	19.81193	22.36203	24.73558	27.68818	29.81932
	14	21.06414	23.68478	26.11893	29.14116	31.31943
	15	22.30712	24.99580	27.48836	30.57795	32.80149
	16	23.54182	26.29622	28.84532	31.99986	34.26705
	17	24.76903	27.58710	30.19098	33.40872	35.71838
	18	25.98942	28.86932	31.52641	34.80524	37.15639
	19	27.20356	30.14351	32.85234	36.19077	38.58212
	20	28.41197	31.41042	34.16958	37.56627	39.99686
	21	29.61509	32.67056	35.47886	38.93223	41.40094
	22	30.81329	33.92446	36.78068	40.28945	42.79566
	23	32.00689	35.17246	38.07561	41.63833	44.18139
	24	33.19624	36.41503	39.36406	42.97978	45.55836
	25	34.38158	37.65249	40.64650	44.31401	46.92797

Tabel Z ($Z_{\frac{\alpha}{2}} = Z_{\frac{0,05}{2}} = Z_{0,025} = 1,96$)

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,8	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,7	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,6	0,0002	0,0002	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,5	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681

RIWAYAT HIDUP



Aisa Khoirul Umaroh, lahir di Malang pada 18 November 1997. Akrab dengan sapaan Aisa. Anak bungsu dari tiga bersaudara dari pasangan Bapak Yanto dan Ibu Ngatiani. Kakak pertama bernama Anif Munandar dan kakak kedua bernama Anis Safa'atul Anggraeni.

Adapun riwayat pendidikannya diawali di TK Nusa Indah 01 Batu pada tahun 2002 sampai 2004. Pendidikan dasarnya ditempuh di SDN Sumberejo 03 Batu lulus pada tahun 2010. Setelah itu melanjutkan sekolah menengah di SMP Negeri 01 Batu lulus tahun 2013 dan Madrasah Aliyah Negeri Batu lulus pada tahun 2016. Pada tahun yang sama melanjutkan pendidikannya di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Jurusan Matematika melalui jalur SNMPTN.

Selama menjadi mahasiswa, berperan aktif di organisasi kampus baik intra maupun ekstra dan komunitas dalam rangka mengembangkan kompetensi akademik, potensi atau *softskill* dan membangun jaringan. Aisa pernah menjadi pengurus Divisi Publik Relasi Eksternal tahun 2017 dan Bendahara Umum tahun 2018 di HMJ “Integral” Matematika. Selain itu, menjadi anggota Departemen Pengembangan Minat dan Bakat DEMA Fakultas Sains dan Teknologi tahun 2019. Sekaligus menjadi kader Pergerakan Mahasiswa Islam Indonesia (PMII) Rayon “Pencerahan Galileo” tahun 2018. Aisa menjadi salah satu penerima Beasiswa Bank Indonesia pada tahun 2018-2019 sehingga dirinya aktif dalam komunitas mahasiswa penerima beasiswa Bank Indonesi, Generasi Baru Indonesi (GenBI) Korkom Malang. Selain organisatoris, Aisa juga termasuk mahasiswa akademis. Hal tersebut dibuktikan dengan indeks prestasi kumulatif diatas 3,5 dan pernah menjadi asisten praktikum statistika elementer tahun 2019. Penulis dapat dihubungi melalui email: aisakumaroh16ixh@gmail.com.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax.(0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Aisa Khoirul Umaroh
NIM : 16610030
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Perbandingan Metode Regresi Logistik Biner dan *Classification And Regression Tree* pada Klasifikasi Status Kesejahteraan Rumah Tangga di Kota Batu
Pembimbing I : Ria Dhea Layla Nur Karisma, M.Si
Pembimbing II : Muhammad Khudzaifah, M.Si

No.	Tanggal	Hal	Tanda Tangan	
1.	04 Februari 2020	Konsultasi Bab I	1.	
2.	13 Februari 2020	Konsultasi Bab II		2.
3.	20 Februari 2020	Konsultasi Bab III	3.	
4.	02 Maret 2020	Mencari Integrasi Ayat Al-Qur'an		4.
5	12 Maret 2020	Penambahan Integrasi Ayat di Bab II	5.	
6	30 Maret 2020	Konsultasi Bab I, II, III, IV, dan ACC Sempurna		6.
7	10 Mei 2020	Perbaikan Integrasi Ayat Al-Qur'an pada Bab II	7.	
8	13 Juni 2020	Konsultasi Persiapan Sidang I		8.
9	15 Juni 2020	Konsultasi Persiapan Sidang II	9.	
10	16 Juni 2020	ACC untuk disidangkan		10.

Malang, 4 September 2020
Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001