

**IMPLEMENTASI ALGORITMA C5.0 PADA KLASIFIKASI DATA
SOSIAL MASYARAKAT
(Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota
Singkawang)**

SARI DEVI ASRI

H1011181053

SKRIPSI



**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK**

2023

**IMPLEMENTASI ALGORITMA C5.0 PADA KLASIFIKASI DATA
SOSIAL MASYARAKAT
(Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota
Singkawang)**

SARI DEVI ASRI

H1011181053

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Matematika pada Program Studi Matematika



**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK**

2023



UNIVERSITAS TANJUNGPURA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

SURAT KETERANGAN

Nomor : 5095/UN22.8/PJ.00.01/2023

Pengelola jurnal Bimaster Jurusan Matematika/Program Studi Matematika dengan ini menerangkan bahwa:

Nama : Sari Devi Asri
NIM : H1011181053
Jurusan/Prodi : Matematika/Matematika
Tanggal Lulus : 4 Mei 2023
Pembimbing : 1. Drs. Helmi, M.Si
2. Nur'ainul Miftahul Huda, S.Si., M.Si

Telah menyerahkan Artikel Ilmiah dengan judul

Implementasi Algoritma C5.0 Pada Klasifikasi Data Sosial Masyarakat (Studi Kasus: Kelayakan Penerimaan BLT di Kelurahan Condong Kita Singkawang)

Sebagai Tugas Akhir pada Jurusan Matematika/Program Studi Matematika dan sudah diterbitkan pada jurnal Bimaster periode terbit Volume 12, No. 3 (2023), hal 259 - 268 dengan laman

<http://jurnal.untan.ac.id/index.php/jbmstr/article/view/66693>

Dengan surat keterangan ini diberikan untuk dipergunakan sebagaimana mestinya.

Pontianak, 22 Juni 2023

Mengetahui,

Wakil Dekan FMIPA UNTAN

Yudha Arman, D.Sc
NIP. 197805132003121002

Pengelola Jurnal Bimaster


Fransiskus Fran, M.Si
NIP. 198804152019031014

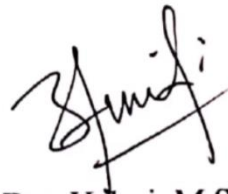
**IMPLEMENTASI ALGORITMA C5.0 PADA KLASIFIKASI DATA
SOSIAL MASYARAKAT**
**(Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota
Singkawang)**

Tanggung Jawab Yuridis Material Pada

Sari Devi Asri
NIM.H1011181053

Disetujui Oleh

Pembimbing I,



Drs. Helmi, M.Si.
NIP. 196410171998021001

Pembimbing II,



Nur'ainul Miftahul Huda, M.Si.
NIP. 199411142020122014

Disahkan Oleh,
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura Pontianak



Dr. Gusrizal, S.Si, M.Si.
NIP.197108022000031001

IMPLEMENTASI ALGORITMA C5.0 PADA KLASIFIKASI DATA SOSIAL MASYARAKAT

**(Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota
Singkawang)**

INTISARI

Algoritma C5.0 merupakan salah satu algoritma pohon keputusan yang dapat memproses data sosial masyarakat menjadi sebuah aturan yang bisa dijadikan masukan dalam pengambilan keputusan. Pada penelitian ini, peneliti menganalisis variabel yang paling menentukan kelayakan untuk Bantuan Langsung Tunai (BLT). Adapun variabel yang digunakan dalam penelitian ini adalah pendapatan, umur, pendidikan, pekerjaan, kepemilikan rumah, jumlah tanggungan dan keputusan kelayakan. Tujuan penelitian ini adalah mengklasifikasi data sosial masyarakat untuk kelayakan penerimaan bantuan program BLT menggunakan metode Algoritma C5.0 serta mengetahui tingkat akurasi dan laju *error* hasil klasifikasi metode Algoritma C5.0. Ada beberapa langkah yang dilakukan yaitu menginput data yang diteliti. Selanjutnya pemilihan *node* akar diawali dengan menghitung nilai *entropy*. Kemudian proses dilanjutkan dengan mencari nilai *gain*. Setelah itu mencari nilai *gain ratio*. Penentuan cabang untuk masing-masing *node* dengan menghitung nilai *gain ratio* tertinggi dari variabel bebas yang ada. Penelitian ini menghasilkan *decision tree* dari kasus yang diangkat menunjukkan bahwa ada beberapa variabel yang mempengaruhi dalam penentuan masyarakat yang mendapatkan BLT. Setelah dilakukan analisis didapatkan hasil tingkat akurasi sebesar 81,429%. Dengan tingkat akurasi yang tinggi bisa dikembangkan menjadi sebuah aturan yang dapat memberikan prediksi atau masukan dalam membuat keputusan kelayakan penerima bantuan langsung tunai (BLT).

Kata kunci : C5.0, bantuan langsung tunai, gain ratio

IMPLEMENTATION OF C5.0 ALGORITHM ON SOCIAL COMMUNITY DATA CLASSIFICATION

**(Case Study : Feasibility of Accepting BLT in Condong Village Singkawang
City)**

ABSTRACT

The C5.0 algorithm is a decision tree algorithm that can process community social data into rules that can be used as input in decision making. In this study, researchers analyzed the variables that most determine eligibility for Direct Cash Assistance (BLT). The variables used in this study are income, age, education, employment, home ownership, number of dependents and eligibility decisions. The purpose of this study was to classify community social data for the eligibility of receiving BLT program assistance using the C5.0 Algorithm method and to determine the accuracy and error rate of the classification results of the C5.0 Algorithm method. There are several steps taken, namely inputting the data studied. Next, the selection of the root node begins with calculating the entropy value. Then the process is continued by finding the gain value. After that look for the value of the gain ratio. Determination of the branch for each node by calculating the value of the highest gain ratio of the existing independent variables. This study produced a decision tree from the cases raised showing that there were several variables that influenced the determination of the people who received BLT. After the analysis was carried out, the results obtained an accuracy rate of 81.429%. With a high level of accuracy it can be developed into a rule that can provide predictions or input in making decisions regarding the eligibility of recipients of direct cash assistance (BLT).

Keywords: C5.0, direct cash assistance, gain ratio

KATA PENGANTAR

Puji syukur penulis panjatkan atas kehadiran Allah Yang Maha Esa yang telah melimpahkan rahmat dan hidayah-nya, sehingga penyusunan skripsi ini dapat diselesaikan. Adapun judul dari skripsi ini adalah **“Implementasi Algoritma C5.0 Pada Klasifikasi Data Sosial Masyarakat (Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota Singkawang)”**. Skripsi ini dibuat sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika pada Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, universitas Tanjungpura Pontianak.

Dalam penyusunan dan penulisan skripsi ini dapat terselesaikan berkat bimbingan, arahan, dukungan serta motivasi dari berbagai pihak sehingga dalam penyusunan skripsi ini dapat berjalan dengan lancar. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih kepada semua pihak yang telah membantu penulisan skripsi ini, terutama kepada:

1. Kedua orang tua, kakak dan abang ipar serta adik penulis yang selalu memberi kasih sayang dan dukungan, serta menjadi motivasi penulis hingga skripsi ini dapat terselesaikan.
2. Bapak Drs. Helmi, M.Si., selaku Dosen Pembimbing pertama dan ibu Nur'ainul Miftahul Huda, M. Si., selaku Dosen Pembimbing kedua yang telah membimbing dan memberikan motivasi, arahan, masukan serta nasihat kepada penulis selama penulisan skripsi ini.
3. Ibu Dr. Yundari, M. Sc., selaku Dosen Pembimbing Akademik yang sudah membimbing dan memberikan motivasi serta masukan dari awal perkuliahan sampai saat ini.
4. Ibu Dr. Yundari, M.Sc., selaku Dosen Penguji pertama dan ibu Dr. Nilamsari Kusumastuti, M.Sc., selaku Dosen Penguji kedua yang telah memberikan masukan serta saran dalam penulisan skripsi ini.
5. Seluruh dosen di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam yang sudah membimbing dan memberikan ilmu yang bermanfaat bagi penulis selama menjalankan studi.

6. Comdev dan *Outraching* Universitas Tanjungpura yang sudah membantu baik secara materi maupun motivasi dari awal perkuliahan sampai sekarang.
7. Semua pihak Kelurahan Condong Kota Singkawang yang telah menerima penulis dengan baik dan mengizinkan penulis untuk ikutserta dalam mewawancarai masyarakat, sehingga diperoleh data untuk penulisan skripsi ini.
8. Rina Megasari, Ririn Febriyanti, Anggi, Nurul Fadhillah, Yusi Sania, Dimas Catur Ramadhani, Arif Febriyanto dan Rudi Iswahyudi yang telah banyak membantu, memberikan dukungan dan menemani penulis dalam menyelesaikan skripsi ini.
9. Seluruh teman-teman Matematika angkatan 2018 (Fermat'18) yang tidak dapat disebutkan satu persatu yang telah memberikan motivasi dan dukungan kepada penulis.
10. Author dan penerbit dari aplikasi Webtoon, KakaoWebtoon, Wattpad, NovelToon, serta situs web lainnya dan teman online yang selama ini menemani penulis melewati masa perkuliahan sampai dengan sekarang dalam menyusun skripsi yang membuat penulis kembali semangat ketika tidak menemukan ide dalam penulisan skripsi ini.
11. Penulis yang telah bersedia menyelesaikan skripsi ini dengan ikhlas, semangat dan sabar.
12. Semua pihak yang tidak dapat penulis sebutkan satu per satu.

Dengan demikian, penulis berharap skripsi ini dapat memberikan kontribusi bagi kemajuan pendidikan khususnya bidang Matematika serta bermanfaat bagi pembaca dan dipergunakan sebagai mestinya. Apabila ada kesalahan maupun kekurangan dalam penulisan skripsi ini, kritik dan saran yang sifatnya membangun dari pembaca sangat penulis harapkan demi memaksimalkan skripsi ini.

Pontianak, April 2023

Penulis

DAFTAR ISI

INTISARI.....	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
DAFTAR SIMBOL	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Tinjauan Pustaka	4
1.6 Metodologi Penelitian	4
BAB II LANDASAN TEORI.....	8
2.1 Analisis Deskriptif	8
2.2 Implementasi.....	9
2.3 <i>Data Mining</i>	9
2.4 <i>Decision Tree</i>	10
2.5 <i>Confusion Matrix</i>	12
2.6 <i>Non-Probability Sampling</i>	14
BAB III HASIL DAN PEMBAHASAN.....	16
3.1 Algoritma C5.0	16
3.2 Analisis Statistika Deskriptif.....	18
3.3 Perhitungan Algoritma C5.0	20
3.4 Mengukur Tingkat Akurasi dan Laju <i>Error</i>	28
BAB V KESIMPULAN.....	31
DAFTAR PUSTAKA	32
LAMPIRAN	34

DAFTAR GAMBAR

Gambar 1.1 <i>Flowchart</i> Algoritma C5.0.....	6
Gambar 2.1 Struktur Pohon Klasifikasi <i>Decision Tree</i>	12
Gambar 3.1 Diagram Pie Keputusan Kelayakan	18
Gambar 3.2 Diagram Batang Keputusan Kelayakan.....	19
Gambar 3.3 Hasil Pembentukan Cabang di <i>Node</i> Akar	24
Gambar 3.4 Proses Pembentukan Pohon Keputusan	25
Gambar 3.5 Pohon Klasifikasi Algoritma C5.0	27

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	13
Tabel 3.1 Variabel-Variabel yang Mempengaruhi Kelayakan Penerimaan BLT.	20
Tabel 3.2 Hasil Perhitungan <i>Entropy</i> , <i>Gain</i> dan <i>Gain Ratio</i>	22
Tabel 3.3 Tingkat Akurasi Klasifikasi Algoritma C5.0	29

DAFTAR SIMBOL

S	: Kasus
m	: Banyaknya kelas pada variabel A_i
p_i	: Proporsi dari S dan S_i
$A_{i,j}$	: Variabel ke-i pada kategori ke - j
A_i	: Variabel ke - i
$ S_i $	: Banyaknya kasus pada kategori ke - i
$ S $	: Banyaknya kasus dalam S
$Gain(A_i)$	: Nilai gain dari suatu variabel
$\sum_{i,j=1}^m Entropy(A_{i,j})$	: Jumlah nilai entropy pada variabel ke-i pada kategori ke -j
TP	: Jumlah nilai benar dinyatakan benar
FP	: Jumlah nilai salah dinyatakan benar
FN	: Jumlah nilai benar dinyatakan salah
TN	: Jumlah nilai salah dinyatakan salah

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data adalah segala fakta, angka, atau teks yang dapat diproses oleh komputer. Salah satu bentuk pengolahan suatu data yaitu *data mining*. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam mengambil sebuah keputusan. Munculnya *data mining* didasarkan pada kenyataan bahwa jumlah data yang tersimpan dalam basis data semakin besar (Prasetyo, 2012). Banyak fungsi yang dapat dilakukan menggunakan *data mining*, diantaranya adalah deskripsi, estimasi, prediksi, klasifikasi, pengklusteran, dan asosiasi. Klasifikasi adalah suatu metode yang digunakan untuk memberikan label dari suatu data/objek baru. Ada beberapa macam pengklasifikasian dalam *data mining* yaitu *decision tree*, *naive Bayes*, *Support Vector Machine (SVM)* dan lain-lain (Larose, 2005).

Decision tree atau pohon keputusan adalah salah satu teknik klasifikasi. *Decision tree* adalah *top-down* pohon rekursif dari algoritma induksi, yang menggunakan ukuran seleksi variabel untuk memilih variabel yang diuji. Pohon rekursif adalah salah satu cara untuk menyelesaikan persamaan recurrence secara langsung dengan menggambar, sedangkan algoritma induksi adalah pembuktian pernyataan yang berkaitan dengan objek diskrit yang menunjukkan hasil yang kompleks (Han, Jiawei dan Kamber, Micheline, 2011). Pohon yang dibentuk tidak selalu berupa pohon biner. Jika semua fitur dalam data set menggunakan dua macam nilai kategorikal maka bentuk pohon yang didapatkan berupa pohon biner. Sebaliknya jika terdapat lebih dari dua macam nilai kategorikal atau menggunakan tipe numerik maka pohon yang dihasilkan tidak berupa pohon biner. Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan, salah satunya yaitu algoritma C5.0 (Prasetyo, 2014).

Algoritma C5.0 merupakan klasifikasi algoritma yang cocok untuk kumpulan data besar dan penyempurnaan dari algoritma C4.5 dimana prosesnya hampir sama, hanya saja algoritma C5.0 punya kelebihan dari algoritma sebelumnya.

Dibandingkan dengan algoritma klasifikasi lainnya algoritma C5.0 memiliki kinerja lebih baik, selain itu hasil kinerja yang didapatkan dari proses klasifikasi algoritma C5.0 juga memiliki tingkat akurasi yang tinggi. Dalam algoritma C5.0 pemilihan atribut yang akan diproses menggunakan ukuran *gain ratio* untuk memilih atribut uji pada setiap *node* di dalam *tree*. Ukuran ini digunakan untuk memilih atau membentuk *node* pada pohon. Atribut dengan nilai *gain ratio* tertinggi akan terpilih sebagai parent bagi node selanjutnya (Kusrini dan Luthfi, 2009). Salah satu contoh dari pengklasifikasian Algoritma C5.0 yaitu tentang data sosial masyarakat.

Data sosial adalah data tentang bagaimana orang berinteraksi dalam konteks sosial, seringkali dengan data yang diperoleh dari survei, wawancara atau layanan jejaring sosial (Martin, 2005). Tujuannya mungkin hanya untuk memahami perilaku manusia atau bahkan untuk memperoleh data yang digunakan untuk penelitian atau dipublikasikan. Aplikasi metode klasifikasi dalam bidang sosial salah satunya adalah untuk memberi keputusan atas kelayakan penerimaan program bantuan masyarakat miskin yaitu Bantuan Langsung Tunai (BLT). Salah satu kelurahan di Kota Singkawang yaitu Kelurahan Condong mendapatkan program bantuan tersebut, namun program yang diselenggarakan oleh pemerintah masih belum tergolong sukses, karena masih banyak masyarakat miskin yang belum mendapatkan program bantuan tersebut dan sebaliknya, masyarakat yang tergolong mampu yang mendapatkan bantuan. Untuk menentukan kelayakan tersebut dibutuhkan beberapa variabel seperti pendapatan masyarakat (kelas menengah dan rendah), pekerjaan, jumlah tanggungan, kepemilikan rumah, pendidikan dan umur. Setelah menghasilkan pohon keputusan dari algoritma C5.0 dilakukan uji tingkat akurasi dan laju error menggunakan *confusion matrix*.

Di bidang pembelajaran mesin dan khususnya masalah klasifikasi statistik, *confusion matrix*, juga dikenal sebagai matriks *error* adalah tata letak tabel khusus yang memungkinkan visualisasi kinerja suatu algoritma, biasanya pembelajaran yang diawasi; dalam pembelajaran tanpa pengawasan biasanya disebut matriks pencocokan (Stehman dan Stephen, 1997). *Confusion matrix* merupakan sebuah metode untuk mengevaluasi yang menggunakan tabel *matrix*. Hasil evaluasi dengan

menggunakan *confusion matrix* menghasilkan nilai akurasi, serta laju error (Pramana dkk, 2018).

Pada tugas akhir ini penulis mencoba menerapkan algoritma C5.0 dan *confusion matrix* yang diaplikasikan dalam memutuskan kelayakan penerimaan program BLT di Kelurahan Condong Kota Singkawang serta melihat tingkat akurasi dan laju error dari algoritma C5.0. Berdasarkan latar belakang tersebut, maka penulis tertarik untuk membuat penelitian ilmiah dengan judul “Implementasi Algoritma C5.0 Pada Klasifikasi Data Sosial Masyarakat (Studi Kasus : Kelayakan Penerimaan BLT di Kelurahan Condong Kota Singkawang)”.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka permasalahan yang dapat dirumuskan dalam penelitian ini sebagai berikut :

1. Bagaimana klasifikasi data sosial masyarakat untuk kelayakan penerimaan bantuan program BLT menggunakan algoritma C5.0?
2. Bagaimana hasil tingkat akurasi dan laju error hasil klasifikasi metode algoritma C5.0 menggunakan *confusion matrix*?

1.3 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai dalam penelitian ini adalah :

1. Mengklasifikasi data sosial masyarakat untuk kelayakan penerimaan bantuan program BLT menggunakan metode algoritma C5.0.
2. Mengetahui tingkat akurasi dan laju error hasil klasifikasi metode algoritma C5.0 menggunakan *confusion matrix*.

1.4 Batasan Masalah

Berdasarkan uraian dari latar belakang, maka batasan masalah pada penelitian ini adalah menggunakan teknik *Non-Probability Sampling* yaitu *Systematic Sampling* dalam mengolah data, mengetahui tingkat akurasi dan laju error hasil klasifikasi metode algoritma C5.0 menggunakan *confusion matrix*. Penelitian berlokasi di Kelurahan Condong Kota Singkawang mengenai data umur,

pendidikan terakhir, kepemilikan rumah, jumlah tanggungan, pekerjaan, dan pendapatan masyarakat. Data diambil pada tahun 2021.

1.5 Tinjauan Pustaka

Pada penelitian ini menggunakan referensi dari Wijaya, Hasibuan dan Ramadhani, (2018) tentang implementasi algoritma C5.0 dalam klasifikasi pendapatan masyarakat dengan studi kasus : Kelurahan Mesjid Kecamatan Medan Kota. Penelitian ini membahas tentang kelayakan penerimaan bantuan BLT menggunakan aplikasi WEKA dan diperoleh pohon keputusan dengan hasil gain tertinggi yaitu jumlah anggota keluarga.

Penelitian yang berkaitan dengan metode Algoritma C5.0 dilakukan oleh Kastawan, Wiharta dan Sudarman, (2018) tentang implementasi algoritma C5.0 pada penilaian kinerja pegawai negeri sipil. Penelitian ini membahas tentang menggunakan algoritma C5.0 untuk memproses data kinerja pegawai menjadi sebuah aturan yang bisa dijadikan masukan dalam pengambilan keputusan. Berdasarkan hasil evaluasi terhadap 184 data kinerja pegawai didapat tingkat akurasi sebesar 96.08%.

Referensi penelitian yang membahas tentang perbandingan Algoritma C5.0, CART dan CHAID dilakukan oleh Yusuf (2007) tentang perbandingan algoritma decision tree C5.0, CART dan CHAID dengan kasus prediksi status resiko kredit di Bank X. Penelitian ini membahas perbandingan tiga algoritma yang mana dapat mengetahui algoritma manakah yang tepat untuk memprediksi status resiko kredit di Bank X melalui cara *credit scoring* menggunakan teknik *decision tree*. Berdasarkan hasil penelitian tersebut ketiga algoritma menghasilkan model tree yang berbeda untuk data set yang sama dengan tingkat keakuratan dari Algoritma C5.0 sebesar 87,27%, CART sebesar 87,27% dan CHAID sebesar 87,15%. Dari analisis statistik diperoleh bahwa tidak ada perbedaan performansi yang signifikan di antara ketiga algoritma tersebut.

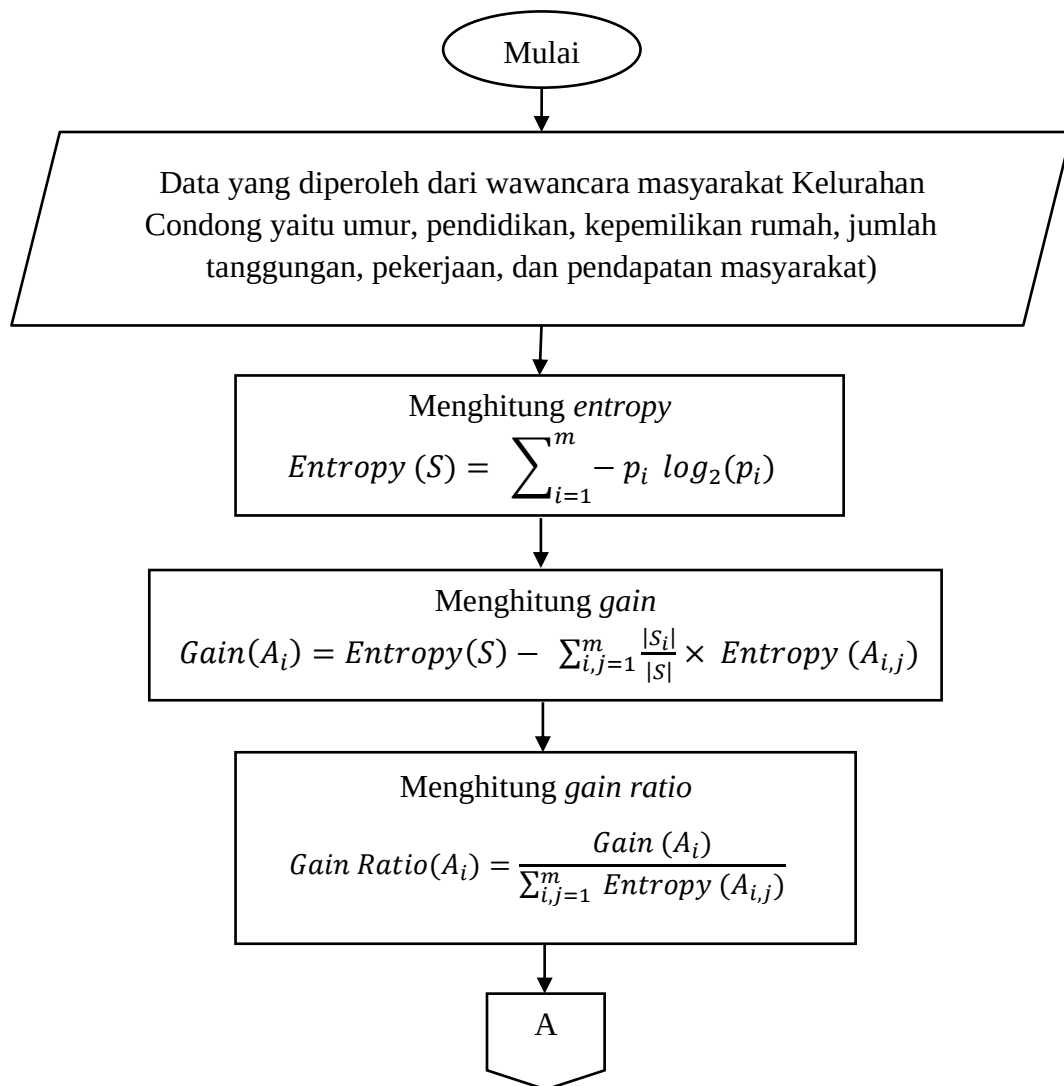
1.6 Metodologi Penelitian

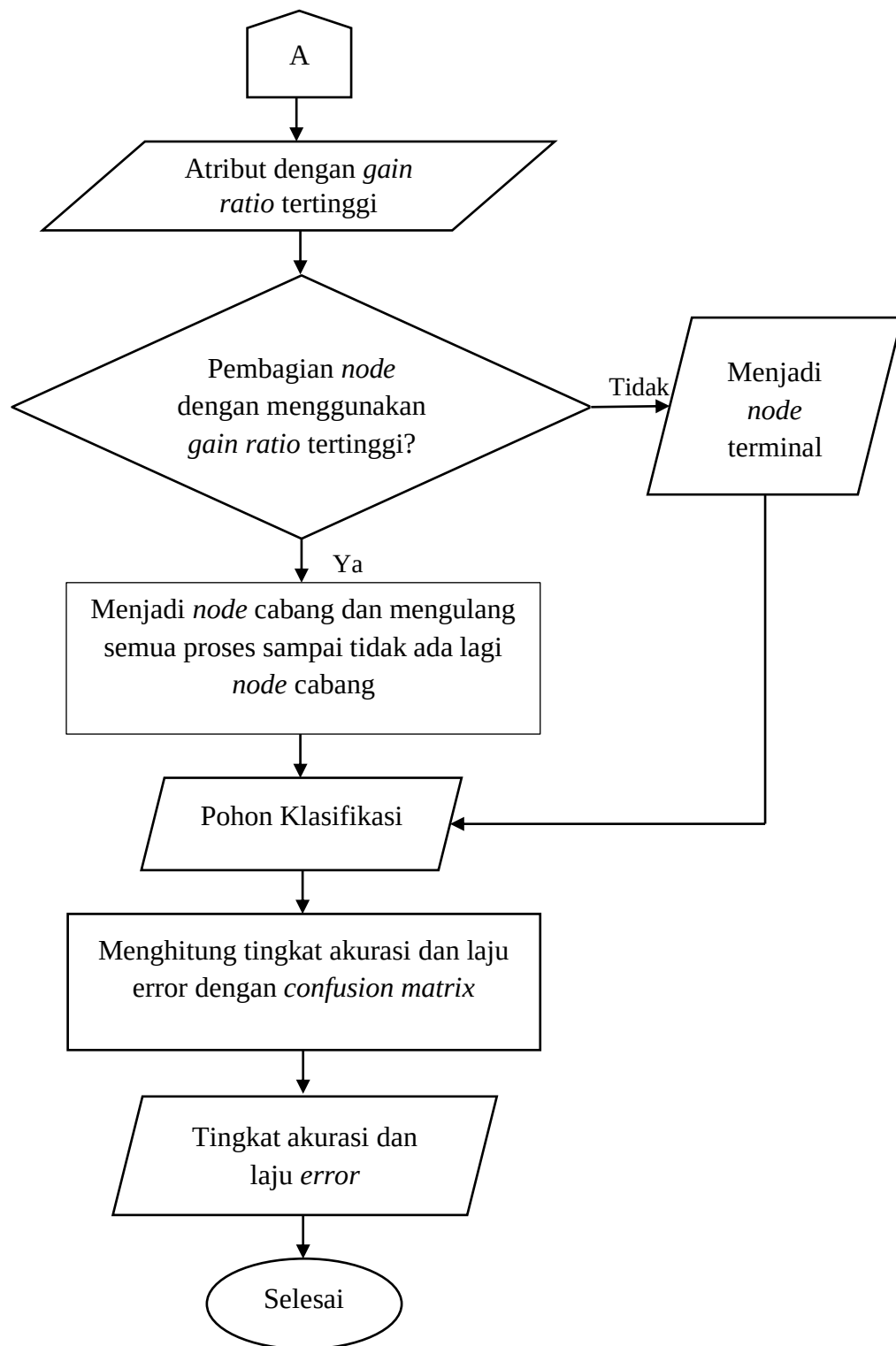
Pada penelitian ini, data yang digunakan adalah data primer. Pengambilan data dilakukan dengan menemui responden secara langsung dan mencatat data,

penelitian berlokasi di Kelurahan Condong Kota Singkawang mengenai data umur, pendidikan terakhir, kepemilikan rumah, jumlah tanggungan, pekerjaan, dan pendapatan masyarakat. Populasi dalam penelitian ini adalah data sosial seluruh Kepala Keluarga masyarakat Kelurahan Condong Kota Singkawang sebanyak 492 KK, sedangkan yang menjadi sampel sebanyak 70 KK setelah dilakukan teknik *Non-Probability Sampling* yaitu *Systematic Sampling*. Selanjutnya, mempelajari konsep dasar data *mining*, mengklasifikasi data menggunakan metode Algoritma C5.0 untuk menentukan pohon keputusan dengan menghitung *entropy* serta *information gain*, setelah itu menghitung tingkat akurasi/ketepatan dari model klasifikasi dengan menggunakan *Confusion Matrix*.

Dalam analisis menggunakan Algoritma C5.0 ada beberapa langkah yaitu, menginput data yang diteliti. Selanjutnya pemilihan *node* akar diawali dengan menghitung nilai *entropy*. Kemudian proses dilanjutkan dengan mencari nilai *gain*. Setelah itu mencari nilai *gain ratio*. Penentuan cabang untuk masing-masing *node* dengan menghitung nilai *gain ratio* tertinggi dari variabel bebas yang ada. Perhitungan untuk menentukan cabang pada metode ini dilakukan secara manual dan peneliti menggunakan bantuan software *Microsoft Excel 2016*. Kemudian kelas dibagi dalam cabang yang telah ditentukan. Ulangi penentuan cabang sampai pembagian kelas hingga semua kelas pada cabang memiliki kelasnya masing-masing.

Langkah analisis Algoritma C5.0 untuk lebih jelasnya dapat dilihat pada Gambar 1.1





Gambar 1.1 Flowchart Algoritma C5.0