

**ALGORITMA C4.5 DALAM PENGAJUAN KREDIT UNTUK
PEMBELIAN RUMAH DI CAHAYA DARUSSALAM 2 BEKASI**



SKRIPSI

Diajukan untuk memenuhi salah satu syarat kelulusan Program Strata Satu (S1)

DICKI RIZKI AMARULLAH

2016320010

Program Studi Sistem Informasi

Fakultas Informatika

Universitas Bina Insani

2020

PERSEMBAHAN

Teruslah Berbuat Baik Meskipun Ada Orang Yang Tidak Menyukai Perbuatan Baikmu.

Dengan mengucap puji syukur kepada Allah S.W.T. Penulisan Skripsi ini kupersembahkan untuk:

1. Bapak Duriat dan Ibu Yuyun Agustianingsih tercinta yang telah membesarkan aku dan selalu membimbing, mendukung, memotivasi, memberi apa yang terbaik bagiku serta selalu mendoakan aku untuk meraih kesuksesanku.
2. Bapak Uus Rusmawan selaku Dosen Pembimbing Skripsi yang telah membimbing, mendukung, memotivasi, serta memberi ilmu dan nasehat kepada aku untuk meraih kesuksesan.
3. Adik Adikku Tsaniya Desita Anjani dan Faqih Muhammad Syauqi yang telah menjadi curahan hatiku, dan menemaniku disaat aku mengerjakan skripsi dirumah ketika PSBB yang telah memberiku semangat, aku selalu sayang mereka.
4. Teman Teman SI 16A dan SI 16B yang selalu setia dan memberikan semangat.
5. Mufti Abdillah, dan Marketing Perumahan Cahaya Darussalam 2 Bekasi lainnya yang sudah membantu dan memberikan semangat.

Tanpa mereka,

aku dan karya ini tak akan pernah ada



LEMBAR PERNYATAAN DIRI

Dengan ini saya:

Nama : Dicki Rizki Amarullah
NPM : 2016320010
Program Studi : Sistem Informasi
Judul Skripsi : **ALGORITMA C4.5 DALAM PENGAJUAN KREDIT UNTUK PEMBELIAN RUMAH DI CAHAYA DARUSSALAM 2 BEKASI**

Menyatakan dengan sebenarnya

Bahwa dalam penyusunan Skripsi didasarkan pada data factual dan dapat dipertanggungjawabkan serta merupakan karya asli penulis **BUKAN** karya pihak lain. Apabila dikemudian hari diketahui bahwa skripsi penulis disusun berdasarkan data fiktif dan atau merupakan karya tiruan dan atau karya orang lain. Penulis bersedia menerima Sanksi Akademis dalam bentuk apapun.

Pernyataan ini adalah persyaratan dalam penyusunan skripsi.

Bekasi, 19 Mei 2020

Dicki Rizki Amarullah

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Dicki Rizki Amarullah
NPM : 2016320010
Perguruan Tinggi : Bina Insani University
Program Studi : Sistem Infomasi

Dengan ini menyetujui untuk memberikan ijin kepada pihak **Sekolah Tinggi Manajemen Informatika & Komputer Bina Insani University**, Hak Bebas Royalti Non- Eksklusif (Non-exclusive Royalty-Free Right) atas skripsi kami yang berjudul "**“ALGORITMA C4.5 DALAM PENGAJUAN KREDIT UNTUK PEMBELIAN RUMAH DI CAHAYA DARUSSALAM 2 BEKASI.”**"

Dengan Hak Bebas Royalti Non-Eksklusif ini pihak Sekolah Tinggi Manajemen Informatika & Komputer Bina Insani University berhak menyimpan, mengalih-media atau format-kan, mengelolaannya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta skripsi tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Sekolah Tinggi Manajemen Informatika & Komputer Bina Insani University, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam skripsi saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 19 Mei 2020
Yang menyatakan,

Dicki Rizki Amarullah

LEMBAR PERSETUJUAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Dicki Rizki Amarullah
NPM : 2016320010
Program Studi : Sistem Infomasi
Jenjang : Strata Satu (S1)
Judul Skripsi : Algoritma C4.5 Dalam Pengajuan Kredit Untuk Pembelian Rumah Di Cahaya Darussalam 2 Bekasi.

Telah dipertahankan pada periode I-2018 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Ilmu Komputer (S.Kom) pada Program Strata Satu (S1) Program Studi Sistem Informasi di Sekolah Tinggi Manajemen Informatika dan Komputer Bina Insani.

Bekasi, 19 Mei 2020

Dosen Pembimbing

(Uus Rusmawan, S.Pd, M.Kom)

Ketua STMIK BINA INSANI

Ketua Program Studi Sistem Informasi

(Solikin, S.Si, MT)

(Endang Retnoningdih, M.Kom)

LEMBAR PENGUJIAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Dicki Rizki Amarullah
NPM : 2016320010
Program Studi : Sistem Informasi
Judul Skripsi : Algoritma C4.5 Dalam Pengajuan Kredit Untuk Pembelian Rumah Di Cahaya Darussalam 2 Bekasi.

Telah dipertahankan pada periode I-2018 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Ilmu Komputer (S.Kom) pada Program Strata Satu (S1) Program Studi Sistem Informasi di Sekolah Tinggi Manajemen Informatika dan Komputer Bina Insani.

Bekasi, 19 Mei 2020

DEWAN PENGUJI

Ketua Penguji : (.....)

Penguji I : (.....)

Penguji II : (.....)

Ketua Program Studi Sistem Informasi

(Endang Retnoningdih, M.Kom)

 UNIVERSITAS BINA INSANI	LEMBAR KONSULTASI BIMBINGAN SKRIPSI
FAKULTAS INFORMATIKA UNIVERSITAS BINA INSANI	

Nama Mahasiswa : Dicki Rizki Amarullah
 Nomor Pokok Mahasiswa : 2016320010
 Program studi / Kelas : Sistem Informasi / SI-16B
 Jenjang Pendidikan : Satrata Satu S-1
 Dospem : Uus Rusmawan, S.Pd, M.Kom
 Judul Skripsi : Algoritma C4.5 Dalam Pengajuan Kredit Untuk Pembelian Rumah Di Cahaya Darussalam 2 Bekasi.

No	Hari / Tanggal	Masalah Yang dikonsultasikan	Paraf Dosen Pembimbing
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

- Dimulai pada tanggal :
- Diakhiri pada tanggal :
- Jumlah Bimbingan :

Ketua Program Studi Sistem Informasi

Disetujui oleh,
Dosen Pembimbing

(Endang Retnoningdih, M.Kom)

(Uus Rusmawan, S.Pd, M.Kom)

KATA PENGANTAR

Puji syukur alhamdullillah, penulis panjatkan kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan skripsi ini dengan baik. Dimana skripsi ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul skripsi, yang penulis ambil sebagai berikut, “**ALGORITMA C4.5 DALAM PENGAJUAN KREDIT UNTUK PEMBELIAN RUMAH DI CAHAYA DARUSSALAM 2 BEKASI.**”

Tujuan penulisan skripsi ini dibuat sebagai salah satu syarat kelulusan Progrram Strata 1 (S1). Sebagai bahan penulisan diambil berdasarkan hasil penelitian (eksperimen), observasi dan beberapa sumber literatur yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan proposal skripsi ini tidak akan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Bina Insani
2. Dekan Fakultas Informatika Universitas Bina Insani
3. Ketua Jurusan Sistem Informasi Universitas Bina Insani.
4. Bapak/ibu dosen Program Studi Sistem Informasi Fakultas Informatika
5. Universitas Bina Insani yang telah memberikan penulis panduan dengan semua bahan yang diperlukan.
6. Pembimbing Skripsi.
7. Staff / karyawan / dosen di lingkungan Universitas Bina Insani.
8. Orang tua tercinta yang telah memberikan dukungan moral maupun spritual.

9. Rekan-rekan mahasiswa kelas SI 16 B, SI 16 A.
10. Staff/ Karyawan Perumahan Cahaya Darussalam 2 Bekasi yang telah membantu dan memberikan dukungan dalam menyelesaikan penelitian.

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan skripsi ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan dimasa yang akan datang.

Akhir kata semoga skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Bekasi, 16 Januari 2019

Penulis

Dicki Rizki Amarullah

ABSTRAK

Dicki Rizki Amarullah (2016320010)

Penerapan Algoritma C4.5 Untuk Menentukan Faktor Persetujuan Kredit Rumah Subsidi Pada Perumahan Cahaya Darussalam 2 Bekasi.

Peminat rumah subsidi melalui KPR Bank semakin meningkat, data konsumen rumah semakin hari semakin bertambah jumlahnya, baik konsumen yang kreditnya di setujui maupun tidak, tetapi data konsumen belum dimanfaatkan secara maksimal dalam menemukan informasi penting bagi perusahaan, seperti mencari kriteria apa saja yang menentukan persetujuan kredit rumah, sehingga bisa mempercepat proses, tidak perlu menunggu pihak bank melakukan pengecekan yang membutuhkan waktu cukup lama, sehingga konsumen bisa mengetahui hasilnya lebih cepat. Dengan pemanfaatan data konsumen rumah yang maksimal, data-data tersebut bisa memberikan informasi yang belum diketahui sebelumnya, sehingga perusahaan membutuhkan alat bantu analisis untuk menemukan informasi itu. *Data mining* adalah teknik untuk menggali infomasi yang tersembunyi dan selama ini hanya tersimpan sebagai arsip saja. Algoritma C4.5 adalah satu metode data mining yang bertujuan mengasilkan aturan klasifikasi melalui pohon keputusan (*classification rules*) dan prediksi (*prediction*) yang berkaitan dengan artibut sebagai parameter yang dibuat sebagai kriteria dalam pembentukan pohon, yang memperkuat aturan klasifikasi yang ditemukan. Hasil dari penelitian ini, menemukan kriteria - kriteria apa saja yang menentukan persetujuan kredit rumah subsidi.

Kata Kunci: Algoritma C4.5, Aturan Klasifikasi, *Data Mining*, Rumah

ABSTRACT

Dicki Rizki Amarullah (2016320010)

Application of C4.5 Algorithm to Determine Factors for Approval of Subsidized Home Loans in Cahaya Darussalam 2 Bekasi.

Enthusiasts of subsidized houses through Bank KPR are increasing, home consumer data is increasing in number, both consumers whose credit is approved or not, but consumer data has not been maximally utilized in finding important information for the company, such as finding what factors determine credit approval home, so that it can speed up the process, no need to wait for the bank to check that takes a long time, so consumers can find out the results faster. With the maximum utilization of home consumer data, these data can provide information that has not been known before, so companies need analytical tools to find that information. Data mining is a technique to explore hidden information and has only been kept as an archive. C4.5 algorithm is a data mining method that aims to produce classification rules through classification rules and prediction related to attributes as parameters created as criteria in tree formation, which strengthens the classification rules found. The results of this study, find out what my factors determine the approval of subsidized home loan.

Keywords: *C4.5 Algorithm, Classification Rules, Data Mining*

DAFTAR ISI

PERSEMBAHAN.....	i
LEMBAR PERNYATAAN DIRI	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	iii
LEMBAR PERSETUJUAN SKRIPSI	iv
LEMBAR PENGUJIAN SKRIPSI.....	v
KATA PENGANTAR.....	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xvii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah	1
1.2. Identifikasi Masalah	4
1.3. Batasan Masalah.....	5
1.5. Tujuan dan Manfaat Penelitian.....	6
1.6. Sistematika Penulisan.....	7
BAB II LANDASAN TEORI	9
2.1. Tinjauan Pustaka	9
1. Algoritma C4.5	9
2. Pohon Keputusan (Decission Tree)	16
3. Pengertian Klasifikasi	17
4. Pengertian Information Gain	18
5. Pengertian Entropy	18
6. Pengertian Kredit Pemilikan Rumah Subsidi	18
7. Pengertian RapidMiner Studio 8.2.....	18
8. Pengertian Confusion Matrix.....	19
9. Pengertian Data Mining.....	21
2.2. Penelitian Terkait.....	26
BAB III METODOLOGI PENELITIAN	28

3.1. Teknik Pengumpulan Data	28
3.2. Tahapan Data Mining	29
3.3. Metode Algoritma C4.5.....	30
3.4. Pengujian Hasil.....	32
3.5. Kerangka Pemikiran	32
3.5.1. Deskripsi Kerangka Pemikiran	32
BAB IV HASIL DAN PEMBAHASAN.....	35
4.1 Tinjauan Perusahaan.....	35
4.1.1. Sejarah Perusahaan	35
4.1.2. Struktur Organisasi dan Fungsi.....	38
4.1.2.1. Struktur Organisasi	38
4.1.2.2. Fungsi.....	40
4.2. Pembahasan	46
4.2.1 Proses Bisnis Sistem	46
4.2.2 Atribut Keputusan.....	49
4.2.3 Pengelompokan Nilai Atribut Keputusan	49
4.2.4 Data Pengajuan Kredit Pada Perumahan Cahaya Darussalam 2	54
4.2.5 Klasifikasi Jumlah Nilai Atribut	62
4.2.6 Perhitungan Manual Decision Tree Algoritma C4.5	65
4.2.7 Hasil Akhir Pohon Keputusan (Decision Tree)	101
4.2.8 Mengubah Pohon Menjadi Aturan (Tree to Rules)	102
4.3. Hasil.....	104
4.3.1 Analisa Hasil.....	104
4.3.2 Pengujian Algoritma C4.5	107
4.4. Implementasi Algoritma C4.5 dengan <i>RapidMiner Studio 8.2</i>	110
1. Add Data Sample	110
2. Pembuatan Pohon Keputusan Menggunakan Algoritma C4.5	115
3. Konversi pohon keputusan menjadi aturan klasifikasi	118
4. Pengujian Algoritma C4.5	121
5. Perbandingan dengan Algoritma Naïve Bayes	124
BAB V SIMPULAN DAN SARAN.....	127
5.1 Simpulan.....	127
5.2 Saran – Saran.....	128
DAFTAR PUSTAKA	130

DAFTAR RIWAT HIDUP	132
---------------------------------	------------

DAFTAR GAMBAR

Gambar II. 1 Pohon Keputusan Hasil Perhitungan Node 1.	12
Gambar II. 2 Pohon keputusan hasil perhitungan node 1.1.	13
Gambar II. 3 Pohon Keputusan Dari Node 1.1.2	15
Gambar II. 4 Confusion Matrix.....	20
Gambar II. 5 Tahapan Knowledge Discovery in Database (KDD)	23
Gambar III. 1 Tahap Penerapan Algoritma C4.5	31
Gambar III. 2 Kerangka Pemikiran.....	34
Gambar IV. 1 Struktur Organisasi Perumahan Cahaya Darussalam 2 Bekasi.....	40
Gambar IV. 2 Activity Diagram Proses Pengajuan Kredit Rumah Subsidi.....	48
Gambar IV. 3 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node akar 1.	77
Gambar IV. 4 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.1.	80
Gambar IV. 5 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.1.1.	83
Gambar IV. 6 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.	86
Gambar IV. 7 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.	89
Gambar IV. 8 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.1.	92
Gambar IV. 9 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.1.	95
Gambar IV. 10 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.2.....	98
Gambar IV. 11 Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.2.1.....	101
Gambar IV. 12 Pohon Keputusan / Decision Tree akhir.	102
Gambar IV. 13 Hasil Performa Pengujian Algoritma C4.5.	108

Gambar IV. 14 Hasil pengujian Akurasi Algoritma C4.5.....	108
Gambar IV. 15 Hasil pengujian Presisi Algoritma C4.5.....	108
Gambar IV. 16 Hasil Performa Pengujian Algoritma Naïve Bayes.....	109
Gambar IV. 17 Hasil Pengujian Akurasi Algoritma Naïve Bayes.....	109
Gambar IV. 18 Hasil Pengujian Presisi Algoritma Naïve Bayes.....	110
Gambar IV. 19 Splash Screen RapidMiner Studio 8.2	111
Gambar IV. 20 Landing Page RapidMiner Studio 8.2.....	112
Gambar IV. 21 Home Page RapidMiner Studio 8.2	112
Gambar IV. 22 Import Data Sample dari My Computer.	113
Gambar IV. 23 Mencari File Data Sample.....	113
Gambar IV. 24 Pilih Atribut yang akan Digunakan untuk sampling.....	114
Gambar IV. 25 Change Role Atribute keputusan menjadi label.....	114
Gambar IV. 26 Simpan Data Sample.....	115
Gambar IV. 27 Pembuatan Pohon Keputusan (Decision Tree).	116
Gambar IV.28 Parameter Operator Pohon Keputusan (Decision Tree).....	116
Gambar IV. 29 Pohon Keputusan (Decision Tree) Pengajuan Kredit Rumah Subsidi.....	117
Gambar IV. 30 Deskripsi dari Pohon Keputusan (Decision Tree) Pengajuan Kredit Rumah Subsidi.....	117
Gambar IV. 31 Hasil Pohon Keputusan Pengajuan Kredit Rumah Subsidi.	118
Gambar IV. 32 Operator Konversi Tree To Rules.....	119
Gambar IV. 33 Operator Decision Tree didalam Operator Tree To Rules.....	119
Gambar IV. 34 Parameter Operator Decision Tree.....	120
Gambar IV. 35 Hasil Konversi Pohon Keputusan Menjadi Aturan Klasifikasi. 120	
Gambar IV. 36 Pengujian Algoritma C4.5 Menggunakan Confusion Matrix. ...	121
Gambar IV. 37 Parameter Operator Decision Tree.....	122
Gambar IV. 38 Hasil Pengujian Akurasi Algoritma C4.5 Menggunakan Confusion Matrix.	122
Gambar IV. 39 Hasil Pengujian Presisi Algoritma C4.5 Menggunakan Confusion Matrix.	123

Gambar IV. 40 Hasil Pengujian Performa Algoritma C4.5 Menggunakan Confusion Matrix.....	123
Gambar IV. 41 Pengujian Perbandingan Performa dengan algoritma Naïve Bayes Menggunakan Confusion Matrix.....	124
Gambar IV. 42 Hasil Pengujian Akurasi Algoritma Naïve Bayes Menggunakan Confusion Matrix.....	125
Gambar IV. 43 Hasil Pengujian Presisi Algoritma Naïve Bayes Menggunakan Confusion Matrix.....	125
Gambar IV. 44 Hasil Pengujian Performa Algoritma Naïve Bayes Menggunakan Confusion Matrix.....	126

DAFTAR TABEL

Tabel II. 1 Database Keputusan Bermain Golf	9
Tabel II. 2 Mencari Akar Node 1 Menggunakan Information Gain dan Entropy ..	11
Tabel II. 3 Menghitung Akar Node 1.1	12
Tabel II. 4 Menghitung Node 1.1.2 Menggunakan Information Gain dan Entropy ..	14
Tabel IV. 1 Atribut sample decision tree algoritma C4.5	49
Tabel IV. 2 Pengelompokan nilai atribut jenis kelamin.....	50
Tabel IV. 3 Pengelompokan nilai atribut usia.....	50
Tabel IV. 4 Pengelompokan nilai atribut pendidikan.	50
Tabel IV. 5 Pengelompokan nilai atribut Profesi	51
Tabel IV. 6 Pengelompokan nilai atribut Total Masa Kerja.	51
Tabel IV. 7 Pengelompokan nilai atribut Status Kerja.	51
Tabel IV. 8 Pengelompokan nilai atribut Status Pernikahan.	52
Tabel IV. 9 Pengelompokan nilai atribut Jumlah Tanggungan.....	52
Tabel IV. 10 Pengelompokan nilai atribut Status Tempat Tinggal.....	52
Tabel IV. 11 Pengelompokan nilai atribut Total Penghasilan	53
Tabel IV. 12 Pengelompokan nilai atribut Kredit Lainnya	53
Tabel IV. 13 Pengelompokan nilai atribut Tunggakan.	54
Tabel IV. 14 Pengelompokan nilai atribut Keterangan.....	54
Tabel IV. 15 Data Pengajuan Kredit Rumah Subsidi 2019.	55
Tabel IV. 16 Klasifikasi jumlah nilai atribut berdasarkan jenis kelamin.....	62
Tabel IV. 17 Klasifikasi jumlah nilai atribut berdasarkan Usia.....	62
Tabel IV. 18 Klasifikasi jumlah nilai atribut berdasarkan Pendidikan.	62
Tabel IV. 19 Klasifikasi jumlah nilai atribut berdasarkan profesi.	63
Tabel IV. 20 Klasifikasi jumlah nilai atribut berdasarkan Total Masa Kerja.	63
Tabel IV. 21 Klasifikasi jumlah nilai atribut berdasarkan Status Kerja.	63
Tabel IV. 22 Klasifikasi jumlah nilai atribut berdasarkan Status Pernikahan. ...	64
Tabel IV. 23 Klasifikasi jumlah nilai atribut berdasarkan Jumlah Tanggungan..	64

Tabel IV. 24 Klasifikasi jumlah nilai atribut berdasarkan Status Tempat Tinggal.	64

Tabel IV. 25 Klasifikasi jumlah nilai atribut berdasarkan Total Pendapatan.	65
Tabel IV. 26 Klasifikasi jumlah nilai atribut berdasarkan Kredit Lainnya.	65
Tabel IV. 27 Klasifikasi jumlah nilai atribut berdasarkan Tunggakan.	65
Tabel IV. 28 Seluruh atribut dan nilai atribut untuk mencari Information Gain dan entropy	66

Tabel IV. 29 Hasil perhitungan Information Gain dan Entropy node akar 1.....	74
Tabel IV. 30 Hasil perhitungan Information Gain dan Entropy node 1.1.....	78
Tabel IV. 31 Hasil perhitungan Information Gain dan Entropy node 1.1.1.....	81
Tabel IV. 32 Hasil perhitungan Information Gain dan Entropy node 1.2.....	84
Tabel IV. 33 Hasil perhitungan Information Gain dan Entropy node 1.2.1.....	87
Tabel IV. 34 Hasil perhitungan Information Gain dan Entropy node 1.2.1.1.....	90
Tabel IV. 35 Hasil perhitungan Information Gain dan Entropy node 1.2.1.1.1...	93
Tabel IV. 36 Hasil perhitungan Information Gain dan Entropy node 1.2.2.....	96
Tabel IV. 37 Hasil perhitungan Information Gain dan Entropy node 1.2.2.1.....	99
Tabel IV. 38 Mengubah Pohon Keputusan/decision tree menjadi klasifikasi aturan.	102

Tabel IV. 39 Hasil Analisis Menggunakan Algoritma C4.5	104

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Rumah adalah salah satu kebutuhan primer bagi manusia, di era sekarang sulit kiranya untuk membangun rumah secara langsung terlebih di kota besar. Harga tanah dan harga bahan bangunan yang setip saat mengalami kenaikan dan tidak sesuai dengan rata-rata gaji yang diperoleh membuat warga sulit untuk memndirikan sebuah rumah. Pertambahan penduduk yang dari tahun ke tahun semakin padat, hal ini mendorong para developer berlomba-lomba untuk mengembangkan usahanya di bidang perumahan. Karena setiap orang pasti membutuhkan rumah tempat tinggal. Dengan kondisi perekonomian yang tidak stabil, para developer memutar otak untuk bisa menarik para konsumen dengan berbagai cara. Salah satu cara yaitu dengan membangun rumah dengan harga terjangkau. KPR atau Kredit Kepemilikan Rumah adalah merupakan salah satu jenis pelayanan kredit yang diberikan oleh bank kepada nasabah yang mengajukan kredit khusus untuk terpenuhinya kebutuhan dalam mendirikan rumah atau memperbaiki rumah [Zefriyenni and Yuliana IU.,2014:73].

Pihak developer bekerja sama dengan pihak bank untuk memudahkan masyarakat dalam mendapatkan rumah yaitu dengan kredit kepemilikan rumah (KPR). Bank memiliki peranan yang sangat penting untuk menjunjung tinggi

perekonomian rakyat, karena melalui bank unit-unit yang memiliki kelebihan modal dapat di salurkan kepada masyarakat melalui pinjam kredit.

Menurut mentri Pekerjaan Umum dan Perumahan Rakyat Republik Indonesia, kebutuhan perumahan di Indonesia masih tinggi. Untuk memenuhinya tidak dapat hanya melalui dana APBN, namun juga kerja sama dengan perbankan, pengembang/Developer, dan pemerintah daerah. Data Kementerian PUPR memperlihatkan, *trend* pencapaian Program Sejuta Rumah dari tahun ke tahun terus meningkat. Pada awal tahun dicanangkannya, 2015, program ini menorehkan realisasi sebanyak 669.770 unit. Lalu, pada tahun 2016 sebanyak 805.169 unit, dan tahun 2017 sebanyak 904.758 unit. Hingga akhir November 2018 lalu, rumah yang dibangun tercatat mencapai 1.041.323 unit. Usaha perumahan ini sangat menjanjika, karena setiap tahun peminatnya terus meningkat

Salah satunya adalah PT Cahaya Indorahmat Pratamajaya membangun Perumahan Cahaya Darussalam 2 yang lokainya strategis, seperti dekat dengan pintu tol, Sekolah, statsiun, pasar dll. Disamping meningkatnya jumlah konsumen, ditemukan juga banyak kekurangan yang terjadi, diantaranya permasalahan waktu untuk mendapatkan hasil persetujuan kredit dari pihak BANK.

Berdasarkan hasil observasi yang dilakukan peneliti, waktu yang dibutuhkan konsumen untuk mendapatkan persetujuan kredit dari BANK dimulai setelah pengisian formulir pengajuan KPR adalah minimal paling cepat 2 bulan dan selambat lambatnya 3 bulan. Pengetahuan tentang konsumen yang

mendapat persetujuan ataupun tidak dari pihak BANK dapat dimanfaatkan untuk menentukan kriteria - kriteria apa saja yang menentukan persetujuan kredit diterima oleh bank, sehingga konsumen yang ingin membeli rumah subsidi tidak perlu menunggu waktu lama untuk mendapatkan hasil persetujuan kredit. Hasil persetujuan tersebut bisa di dapatkan ketika mengisi formulir di kantor pemasaran perumahan.

Tetapi untuk menentukan kebijakan tersebut pihak developer harus mengetahui dan mencermati data persetujuan kredit konsumen dari BANK. Untuk menemukan informasi kriteria - kriteria yang menentukan persetujuan kredit dari database konsumen, diperlukan suatu metode. Metode digunakan dalam mencari kriteria - kriteria yang menentukan persetujuan kredit rumah subsidi adalah metode kaidah klasifikasi dengan menggunakan algoritma C4.5

Meminjam dengan cara kredit sudah merupakan hal biasa di masyarakat. Sebelum mendapatkan kredit, seseorang harus melalui survey yang akan dilakukan oleh seorang analisis kredit untuk mengetahui apakah pemohon kredit layak atau tidak layak untuk mendapat kredit. Seorang analisis kredit harus benar-benar teliti dalam memprediksi pemohon kredit tersebut dalam pemberian kredit agar tidak terjadi kredit macet. Perlu adanya suatu penunjang keputusan untuk membantu seorang analisis kredit dalam memprediksi pemohon kredit. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Pohon keputusan merupakan salah satu metode klasifikasi yang paling populer karena mudah untuk dipahami. C4.5 merupakan algoritma pohon keputusan yang sering digunakan untuk membuat suatu pohon keputusan karena memiliki tingkat akurasi yang tinggi dalam menentukan keputusan. Algoritma C4.5 adalah suksesor dari ID3 dimana pemilihan root dan parent bukan hanya berdasar information Information Gain saja tetapi juga split information untuk mendapatkan Information Gain Ratio. Dataset yang digunakan dalam penelitian ini yaitu sebanyak 1000 data dengan proporsi 70% disetujui dan 30% data debitur yang ditolak. Dalam laporan ini dibahas kinerja algoritma pohon keputusan C4.5 pada identifikasi kelayakan kredit oleh debitur. Dari penelitian yang dilakukan, diketahui nilai precision terbesar dicapai oleh algoritma C4.5 dengan partisi data 90%:10% dengan nilai sebesar 78,08 %. Nilai recall terbesar partisi data 80%:20% dengan nilai sebesar 96,4 %. Dari hasil data latih

yang sama, ID3 menghasilkan precision sebesar 71,51% dan recall sebesar 92,09%. Hasil akhir dari penelitian ini membuktikan bahwa pada kasus ini algoritma C4.5 memiliki tingkat akurasi yang tinggi dan lebih baik dari ID3. [Rafik et al, 2015:1768]

Hal ini tentu akan mempengaruhi volume penjualan rumah. Data konsumen rumah akan terus bertambah setiap harinya dan menyebabkan penyimpanan data sangat besar. Selama ini data konsumen hanya dijadikan arsip saja, belum dimanfaatkan untuk menemukan pola menentukan kriteria persetujuan kredit rumah subsidi, jika kumpulan data tersebut dan diolah maka arsip data konsumen akan memberikan informasi yang berguna untuk meningkatkan volume penjualan rumah subsidi tersebut. Sehingga pada penelitian ini, peneliti melakukan pemanfaatan data konsumen rumah cahaya Darussalam 2 Bekasi untuk menemukan pola kriteria persetujuan kredit rumah subsidi.

1.2. Identifikasi Masalah

Berdasarkan latar belakang masalah yang telah dikemukakan adapun masalah yang dapat diidentifikasi adalah sebagai berikut:

1. Lambatnya approval kredit oleh pihak bank / Pemberi kredit.
2. Rendahnya volume penjualan rumah yang dilakukan oleh marketing penjualan.
3. Tidak adanya proses validasi yang dilakukan marketing pemasaran terhadap pengajuan kredit konsumen.

1.3. Batasan Masalah

Dalam penelitian perlu adanya pembatasan masalah agar penelitian lebih terarah dan memudahkan dalam pembahasan sehingga tujuan penelitian dapat tercapai. Adapun beberapa batasan yang digunakan dalam penelitian adalah sebagai berikut:

1. Penelitian menganalisa kriteria – kriteria apa saja yang menentukan kredit disetujui oleh pihak bank dan memberikan prediksi kepada pemohon kredit apakah layak untuk diteruskan kepada pemberi kredit/pihak bank.
2. Data transaksi penjualan rumah subsidi yang digunakan pada penelitian ini mulai dari tanggal 01 Januari 2019 sampai dengan 31 Desember 2019 pada Perumahan Cahaya Darussalam 2 Bekasi.
3. Data Sample yang digunakan sebanyak 314 Data, dengan 11 kriteria yang digunakan yaitu Nama, Usia, Pendidikan, Profesi, Total masa kerja, Status Kerja, Status Pernikahan, Jumlah Tanggungan, Status Tempat Tinggal, Total Penghasilan, Kredit Lainnya, Tunggakan, dan Keterangan.
4. Perangkat lunak untuk mengolah data konsumen menggunakan RapidMiner Studio Verison 8.2 dengan menggunakan algoritma C4.5.
5. Objek penelitian ini dilakukan pada perumahan cahaya darussalam 2 Bekasi sehingga hasil *output* hanya dapat digunakan pada perumahan cahaya darussalam 2 Bekasi.

1.4. Perumusan Masalah

Merupakan rumusan pertanyaan dari identifikasi masalah berdasarkan latar belakang dan identifikasi masalah tersebut dapat dirumuskan suatu permasalahan, yaitu:

1. Bagaimana cara agar approval kredit lebih cepat?
2. Bagaimana cara agar volume penjualan meningkat?
3. Bagaimana proses validasi yang dapat dilakukan marketing pemasaran?

1.5. Tujuan dan Manfaat Penelitian

Dari hasil penelitian tujuan dan manfaat yang diperoleh, yaitu:

Tujuan yang diperoleh antaranya, yaitu:

1. Marketing pemasaran melakukan validasi formulir pengajuan kredit agar dapat memberikan hasil prediksi permohonan kredit dapat/tidak diteruskan kepada pihak bank.
2. Menambah volume penjualan rumah dengan menggunakan prediksi kriteria persetujuan kredit berdasarkan hasil penelitian yang dilakukan.

Manfaat yang diperoleh, yaitu:

1. Dapat menambah pelayanan kepada konsumen rumah dan kelangsungan kegiatan penjualan rumah pada perumahan Cahaya Darussalam 2 Bekasi.

2. Membantu marketing dalam menambah volume penjualan rumah.
3. Membantu dalam melakukan strategi promosi pada perumahan Cahaya Darussalam 2 Bekasi.
4. Menjadi acuan mahasiswa Bina Insani atau kampus lain, sebagai bahan referensi melakukan penelitian selanjutnya yang berhubungan dengan metode klasifikasi dan prediksi tersebut.

1.6. Sistematika Penulisan

Secara garis besar skripsi ini terdiri dari 5 (lima) bab dan beberapa lampiran.

Adapun setiap bab dan sub-sub bab yang terdiri dari:

BAB I PENDAHULUAN

Bab ini akan membahas Latar Belakang Masalah pada Perumahan Cahaya Darussalam 2 Bekasi dalam menentukan kriteria – kriteria yang dapat menentukan persetujuan kredit rumah subsidi, batasan masalah dalam penelitian, tujuan dan manfaat penelitian, dan Sistematika Penulisan.

BAB II LANDASAN TEORI

Pada bab ini berisikan tinjauan pustaka yang berisi tentang sumber acuan penulis. Penelitian terkait dengan

permasalahan yang diangkat berasal dari jurnal tentang penerapan data mining menggunakan algoritma C4.5 lima tahun terakhir dari tahun penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tentang metode pengumpulan data yang dilakukan, dan model pengembangan yang digunakan oleh penulis serta kerangka pemikiran untuk memecahkan masalah yang sedang diteliti.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang metodologi pencarian aturan klasifikasi dari data konsumen rumah dengan metode algoritma C4.5 dan diuraikan dengan jelas mengenai langkah-langkah metode algoritma C4.5. Sehingga mendapatkan hasil yang dibutuhkan dalam penelitian ini.

BAB V SIMPULAN DAN SARAN

Bab ini merupakan penutup dari penulisan laporan skripsi di mana simpulan menjawab identifikasi masalah yang dibahas dan saran ditujukan untuk pihak yang mendapat manfaat dari hasil penelitian ini.

BAB II

LANDASAN TEORI

2.1. Tinjauan Pustaka

Dari berbagai aspek dan sudut pandang yang berbeda-beda mencakup sesuai dengan hal-hal yang berkaitan dengan penelitian sebagai berikut:

1. Algoritma C4.5

“Algoritma C4.5 Merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah di pahami dengan bahasa alami. Dan mereka juga dapat di ekspresikan dalam bentuk bahasa basis data seperti *Structure Query Language* untuk mencari *record* pada kategori tertentu.” [Kusrini dan Luthfi, 2009: 13]. Tahapan algoritma C4.5 untuk membangun pohon keputusan, yaitu:

1. Pilih atribut sebagai akar.

Berikut data keputusan bermain golf:

Tabel II. 1
Database Keputusan Bermain Golf

No	Outlook	Temperature	Humidity	Windy	Play
1	Sunny	Hot	High	False	No
2	Sunny	Hot	High	True	No
3	Cloudy	Hot	High	False	Yes
4	Rainy	Mild	High	False	Yes
5	Rainy	Cool	Normal	False	Yes
6	Rainy	Cool	Normal	True	Yes
7	Cloudy	Cool	Normal	True	Yes

No	Outlook	Temperature	Humidity	Windy	Play
8	Sunny	Mild	High	False	No
9	Sunny	Cool	Normal	False	Yes
10	Rainy	Mild	Normal	False	Yes
11	Sunny	Mild	Normal	True	Yes
12	Cloudy	Mild	High	True	Yes
13	Cloudy	Hot	Normal	False	Yes
14	Rainy	Mild	High	True	No

Sumber: Kusrini (2009)

Untuk memilih atribut sebagai akar, didasarkan pada nilai *Information Gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *Information Gain* menggunakan rumus persamaan 1 berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Sumber: Kusrini (2009: 16)

Dimana:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut a

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Untuk menghitung nilai entropi menggunakan persamaan 2 berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

Sumber: Kusrini (2009)

Dimana:

S : himpunan kasus

A : fitur

n : jumlah partisi S

P_i : proporsi dari S_i terhadap S

Hitung jumlah kasus, jumlah kasus untuk keputusan **yes**, jumlah keputusan **no**, dan entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut **OUTLOOK**, **TEMPERATURE**, **HUMIDITY**, dan **WINDY**. dan lakukan perhitungan *Information Gain* untuk setiap atribut.

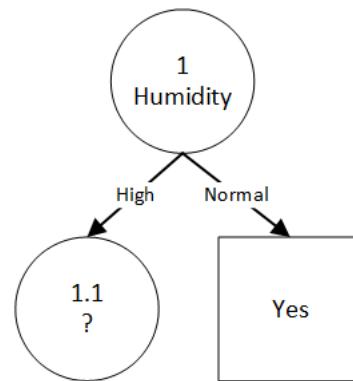
Tabel II. 2
Mencari Akar Node 1 Menggunakan *Information Gain* dan *Entropy*

Node			Jumlah Kasus (S)	Tidak (S ₁)	Ya (S ₂)	Entropy	Information Gain
1	TOTAL		14	4	10	0.863120569	0.258532037
	OUTLOOK						
		CLOUDY	4	0	4		
		RAINY	5	1	4	0.721928095	
		SUNNY	5	3	1	0.970950594	
	TEMPERATURE						0.183850925
		COOL	4	0	4	0	
		HOT	4	2	2	1	
		MILD	6	2	4	0.918295834	
	HUMIDITY						0.370506501
		HIGH	7	4	3	0.985228136	
		NORMAL	7	0	7	0	
	WINDY						0.005977711
		FALSE	8	2	6	0.811278124	
		TRUE	6	4	2	0.918295834	

Sumber: Kusrini (2009)

Dari hasil tabel diatas atribut dengan Information Gain tertinggi adalah **HUMIDITY**, yaitu sebesar 0.37. Dengan demikian **HUMIDITY** dapat menjadi node akar. Ada dua nilai atribut tersebut, nilai atribut **NORMAL** sudah mengklasifikasikan kasus menjadi 1, yaitu keputusan **Yes**, sehingga tidak perlu dilakukan perhitungan lebih lanjut, tetapi untuk nilai atribut **HIGH** masih perlu dilakukan perhitungan lagi.

Dari hasil tersebut dapat digambarkan pohon keputusan sementara sebagai berikut:



Sumber: Kusrini (2009)

Gambar II. 1
Pohon Keputusan Hasil Perhitungan Node 1.

2. Buat cabang untuk tiap nilai-nilai.

Menghitung jumlah kasus, jumlah kasus untuk keputusan **Yes**, jumlah kasus untuk keputusan **No**, dan entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut **OUTLOOK**, **TEMPERATURE**, **HUMIDITY**, dan **WINDY** yang dapat menjadi node akar dari nilai atribut **HIGH**. Dan lakukan perhitungan *Information Gain* untuk setiap atribut.

Tabel II. 3
Menghitung Akar Node 1.1

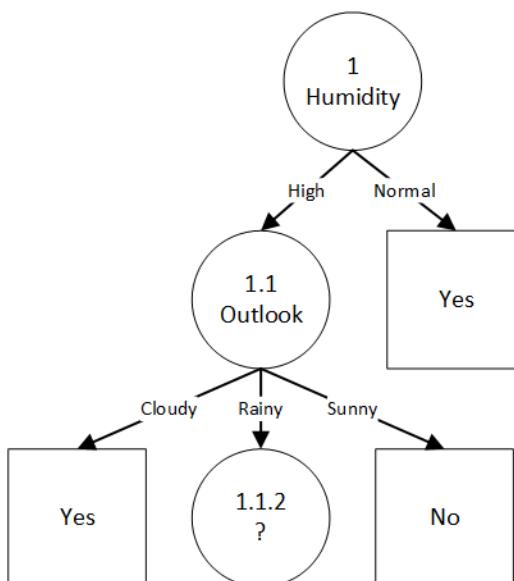
Node			Jumlah Kasus (S)	Tidak (S ₁)	Ya (S ₂)	Entropy	Information Gain
1.1	HUMIDITY-HIGH		7	4	3	0.985228136	
	OUTLOOK						0.69951385
		CLOUDY	2	0	2	0	
		RAINY	2	1	1	1	
		SUNNY	3	3	0	0	
	TEMPERATURE						0.020244207
		COOL	4	0	4	0	

Node		Jumlah Kasus (S)	Tidak (S ₁)	Ya (S ₂)	Entropy	Information Gain
	HOT	4	2	2	1	
	MILD	6	2	4	0.918295834	
	WINDY					0.005977711
	FALSE	8	2	6	0.811278124	
	TRUE	6	4	2	0.918295834	

Sumber: Kusrini (2009)

Dari hasil diatas diketahui bahwa atribut dengan Information Gain tertinggi adalah **OUTLOOK**, yaitu sebesar 0.67. dengan demikian **OUTLOOK** dapat menjadi node cabang dari nilai atribut **HIGH**. Ada 3 nilai atribut dari **OUTLOOK**, yaitu **CLOUDY**, **RAINY**, dan **SUNNY** sudah mengklasifikasikan kasus menjadi satu dengan keputusan **No**, sehingga tidak perlu dilakukan perhitungan lebih lanjut, tetapi untuk nilai atribut **RAINY** masih perlu dilakukan perhitungan lagi

Pohon keputusan yang terbentuk sampai tahap ini adalah sebagai berikut:



Sumber: Kusrini (2009)

Gambar II. 2

Pohon keputusan hasil perhitungan node 1.1.

3. Bagi kasus dalam cabang.

Menghitung jumlah kasus jumlah kasus untuk keputusan **Yes**, jumlah kasus untuk keputusan **No**, dan entropy dari semua kasus yang dibagi berdasarkan atribut **TEMPERATURE** dan **WINDY** yang dapat menjadi cabang dari nilai atribut **RAINY**. Lakukan perhitungan Information Gain untuk tiap atribut

Tabel II. 4

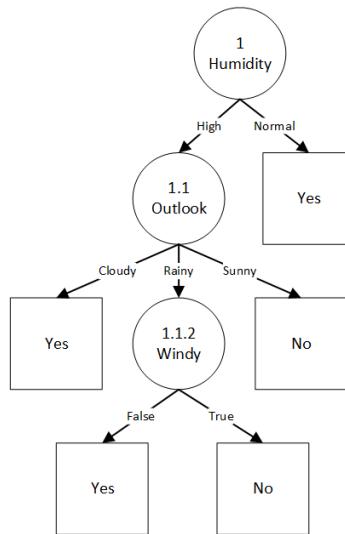
Menghitung Node 1.1.2 Menggunakan Information Gain dan Entropy

Node		Jumlah Kasus (S)	Tidak (S ₁)	Ya (S ₂)	Entropy	Information Gain
1.1.2	HUMIDITY - HIGH dan OUTLOOK - RAINY	2	1	1	1	
	TEMPERATURE					0
	COOL	0	0	0	0	
	HOT	0	0	0	0	
	MILD	2	1	1	1	
	WINDY					1
	FALSE	1	0	1	0	
	TRUE	1	1	0	0	

Sumber: Kusrini (2009)

Dari hasil diatas dapat diketahui bahwa atribut dengan Information Gain tertinggi adalah **WINDY**, yaitu sebesar 1. Dengan demikian **WINDY** dapat menjadi node cabang dari nilai atribut **RAINY**. Ada 2 nilai dari atribut **WINDY** yaitu **FALSE** and **TRUE** sudah mengklasifikasikan kasus menjadi satu dengan keputusan **No**, sehingga tidak perlu dilakukan perhitungan lebih lanjut untuk nilai atribut ini.

Pohon keputusan yang terbentuk sampai tahap ini adalah sebagai berikut:



Sumber: Kusrini (2009)

Gambar II. 3
Pohon Keputusan Dari Node 1.1.2

Berdasarkan gambar II.3 semua kasus sudah masuk dalam kelas, dengan demikian, pohon keputusan pada gambar II.3 merupakan pohon keputusan terakhir yang terbentuk.

Berikut adalah Kondisi berhenti pada pembuatan *decision tree*:

1. Semua data latih memiliki kelas yang sama. Pada kasus ini bentuklah sebuah leaf yang berisi kelas tersebut berdasarkan pengetahuan pakar atau memilih pada hasil acak
2. Data latih kosong (tidak memiliki isi). Pada kondisi ini pembentukan tree dihentikan.
3. Suatu variable tidak memiliki nilai (kosong). Jika kondisi ini terjadi. Buatlah sebuah *leaf* yang berisi kelas yang paling banyak muncul serta cek kondisi *leaf size* yang ditentukan.

2. Pohon Keputusan (Decission Tree)

“Pohon keputusan merupakan representasi sederhana dari teknik klasifikasi untuk sejumlah kelas berhingga, di mana node internal maupun nodul akar ditandai dengan nama kriteria, rusuk-rusuknya diberi label nilai kriteria yang mungkin dan mode daun ditandai dengan kelas kelas yang berbeda”. [Fajar, 2013]

Pohon keputusan biasanya digunakan untuk mendapatkan informasi untuk tujuan pengambilan sebuah keputusan. Pohon keputusan dimulai dengan sebuah node (titik awal) yang dipakai oleh user untuk mengambil tindakan. Dari root ini, user memecahkan sesuai algoritma Tree. Hasil akhirnya adalah sebuah pohon keputusan dengan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil serta hasilnya.

Pada metode pohon keputusan terdapat tiga algoritma yaitu ID3, C4.5, dan C5.0 masing-masing algoritma terdapat perbaikan dari setiap versinya. Algoritma C4.5 merupakan versi perbaikan dari ID3 sedangkan C5.0 merupakan versi perbaikan dari C4.5.

Dalam pembuatan decision Tree dibutuhkan data, data umumnya dibagi menjadi training set dan testing. Training set digunakan oleh algoritma klasifikasi untuk membentuk sebuah model classifier. Testing set digunakan untuk mengukur sejauh mana classifier berhasil melakukan klasifikasi dengan benar titik karena itu, data yang ada pada testing set

seharusnya tidak boleh ada pada training set sehingga dapat diketahui Apakah model classifier sudah benar dalam melakukan klasifikasi.

Manfaat dari pohon keputusan adalah kemampuannya untuk melakukan break down untuk proses pengambilan keputusan yang kompleks menjadi lebih mudah sehingga pengambil keputusan akan lebih meninterpretasikan solusi dari permasalahan.

Decision Tree menggunakan struktur data *tree* sebagai model dalam proses penentuan kelas dari suatu data. Terdapat tiga jenis *node* pada *Decision Tree*:

1. *Root Node*, merupakan node yang tidak memiliki edge masukan dan memiliki nol atau lebih edge keluaran.
2. *Internal Node*, memiliki tepat satu edge masukan dan tidak mempunyai edge keluaran.
3. *Leaf* atau *Terminal Node*. Mempunyai tepat satu edge masukan dan tidak mempunyai edge keluaran.

Pada sebuah *Decision Tree*, setiap *leaf* memiliki sebuah nama kelas. *Root node* dan *internal node* berisi aturan kondisional yang digunakan untuk memisahkan data yang memiliki karakteristik berbeda. [Adinugroho and Sari, 2018: 56]

3. Pengertian Klasifikasi

“Suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi

dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan”
[Nofriansyah, 2014: 10]

4. Pengertian Information Gain

“Information Gain (S,A) merupakan Perolehan informasi dari atribut A relative terhadap output data S. Perolehan informasi didapat dari output data atau variabel dependent S yang dikelompokkan berdasarkan atributA, dinotasikan dengan Information Gain (S,A) “ [Nofriansyah, 2014: 22]

5. Pengertian Entropy

“Entropy(S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S.” [Nofriansyah, 2014: 21]

6. Pengertian Kredit Pemilikan Rumah Subsidi

Kredit Pemilikan Rumah merupakan kredit yang diperuntukan kepada masyarakat berpenghasilan rendah dalam rangka memenuhi kebutuhan perumahan atau perbaikan rumah yang telah dimiliki. Bentuk subsidi yang diberikan berupa subsidi meringankan kredit dan subsidi menambah dana pembangunan atau perbaikan rumah. Kredit subsidi ini diatur tersendiri oleh Pemerintah, sehingga tidak setiap masyarakat yang mengajukan kredit dapat diberikan fasilitas ini. [Hudiyanto, 2017:8]

7. Pengertian RapidMiner Studio 8.2

RapidMiner Studio 8.2 is a *visual workflow designer* that makes data scientists more productive, from the rapid prototyping of ideas to designing mission-critical predictive models. RapidMiner brings artificial intelligence to the enterprise through an open and extensible data science platform. Built for analytics teams, RapidMiner unifies the entire data science lifecycle from data prep to machine learning to predictive model deployment. More

than 625,000 analytics professionals use RapidMiner products to drive revenue, reduce costs, and avoid risks. [RapidMiner: 2001]

Berdasarkan pengertian diatas dapat disimpulkan bahwa RapidMiner Studio 8.2 adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambangan teks, dan analisis prediktif.

8. Pengertian Confusion Matrix

Confusion matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining atau Sistem Pendukung Keputusan. Pada pengukuran kinerja menggunakan confusion matrix, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Nilai True Negative (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan False Positive (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, True Positive (TP) merupakan data positif yang terdeteksi benar. False Negative (FN) merupakan kebalikan dari True Positive, sehingga data positif, namun terdeteksi sebagai data negatif.

		Actual Values	
		1 (Positive)	0 (Negative)
Predictive Values	1 (Positive)	TP (True Positive)	FP (False Positive) Type I Error
	0 (Negative)	FN (False Negative) Type II Error	TN (True Negative)

Sumber: Medium (2019).

Gambar II. 4
Confusion Matrix

Presisi adalah data yang diambil berdasarkan informasi yang kurang dalam klasifikasi biner, presisi dapat dibuat sama dengan nilai prediksi positif. Berikut ini adalah aturan presisi:

$$\text{Presisi} = (\text{TP} / (\text{TP} + \text{FP})) * 100\%$$

Ket:

TP = True Positive yaitu jumlah data positif yang terkласifikasi dengan benar oleh sistem.

FP = False Positive yaitu jumlah data positif namun terkласifikasi salah oleh sistem.

Recall adalah data penghapusan yang berhasil diambil dari data yang relevan dengan kueri. Dalam klasifikasi biner. Recall dikenal sebagai sensitivitas. Munculnya data relevan yang diambil adalah menyetujui dengan query dapat dilihat dengan recall. Berikut ini adalah peran recall:

$$\text{Recall} = (\text{TP} / (\text{TP} + \text{FN})) * 100\%$$

Ket:

TP = True Positive yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem.

FN = False Negative yaitu jumlah data negatif namun terklasifikasi salah oleh sistem.

Akurasi adalah persentase dari total data yang diidentifikasi dan dinilai. Berikut ini adalah aturan akurasi

$$\text{Akurasi} = ((\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})) * 100\%$$

Ket:

TP = True Positive yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem.

FP = False Positive yaitu jumlah data positif namun terklasifikasi salah oleh sistem.

FN = False Negative yaitu jumlah data negatif namun terklasifikasi salah oleh sistem.

9. Pengertian Data Mining

“Data mining is the process of discovering useful patterns and trends in large data sets. “[Larose, 2015: 4]

Berdasarkan pengertian diatas dapat disimpulkan Data Mining adalah proses menemukan pola dan tren yang bermanfaat dalam data yang luas.

“Data Mining adalah rangkaian proses ekstraksi atau penggalian data dan informasi yang besar yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting.” [Indrajani, 2011: 289]

“Istilah data mining dan knowledge discovery in databases (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain.” [Kusrini dan Luthfi, 2009: 7]

Berdasarkan pengertian-pengertian di atas dapat disimpulkan bahwa Data Mining adalah proses menggali dan mengekstrak informasi yang selama ini tidak diketahui atau tersembunyi dari suatu basis data untuk menemukan pengetahuan, yang penting untuk kepentingan perusahaan atau organisasi tersebut.

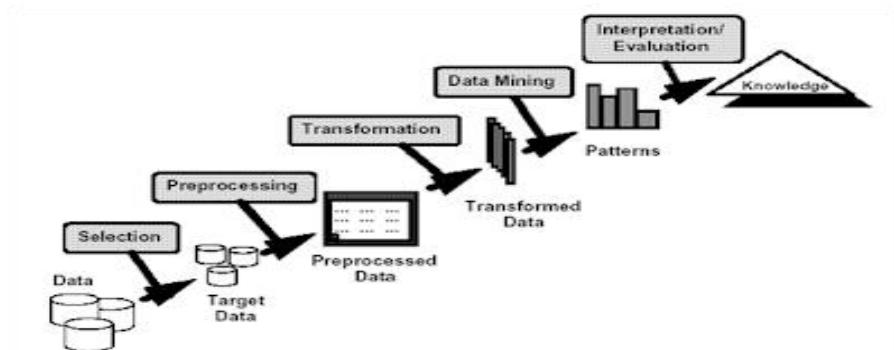
1. Proses Tahapan *Data Mining*

“*Data Mining* merupakan salah satu dari rangkaian *Knowledge Discovery in Database (KDD)*. KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah data. Serangkaian proses tersebut memiliki tahapan sebagai berikut.” [Retno, 2017: 2]

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*)
2. Integrasi data (penggabungan data dari beberapa sumber)

3. Transformasi Data (data diubah menjadi bentuk yang sesuai untuk *mining*)
4. Aplikasi teknik *Data Mining*, proses ekstrasi pola dari data yang ada.
5. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan)
6. Presentasi pengetahuan (dengan teknik visualisasi)

Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna.



Sumber: Retno (2017)

Gambar II. 5
Tahapan Knowledge Discovery in Database (KDD)

2. Pengelompokan *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu: [Kusrini dan Luthfi, 2009:10-12]

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari data untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menentukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelesan untuk suatu pola atau kecenderungan. [Kusrini dan Luthfi, 2009:10-12]

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik dari pada kearah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya. [Kusrini dan Luthfi, 2009:10-12]

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang. Contoh prediksi bisnis dan penelitian adalah:

- a. Prediksi harga beras dalam tiga bulan yang akan datang.
- b. Prediksi persentasi kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi. [Kusrini dan Luthfi, 2009:10-12]

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori , yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah:

- a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau tidak.
- b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- c. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa. [Kusrini dan Luthfi, 2009:10-12]

2.2. Penelitian Terkait

Beberapa penelitian yang terkait dalam penulisan skripsi ini adalah sebagai berikut:

1. Kecelakaan merupakan suatu kejadian yang tidak terencana begitupun pada sebuah proyek konstruksi dimana kecelakaan sering terjadi hal ini disebabkan oleh berbagai faktor. Kita lihat pada Industri jasa konstruksi yang merupakan salah satu sektor industri yang memiliki risiko kecelakaan kerja yang cukup tinggi. Banyaknya kecelakaan kerja yang terjadi tidak terlepas dari faktor Human Error, tentunya berdampak pada kinerja dan pekerjaan yang dilaksanakan, Metode yang digunakan dalam analisis ini adalah Algoritma C4.5 yang merupakan salah satu algoritma modern untuk melakukan Data Mining, Algoritma C4.5 disebut juga dengan pohon keputusan (decision tree) yaitu merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon, dan pada setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas, Konsep dari pohon keputusan ini adalah dengan mengumpulkan data selanjutnya dibuatkan decision tree yang kemudian akan dihasilkan rule-rule solusi permasalahan. Dari hasil penelitian faktor-faktor yang menjadi penyebab terjadinya kecelakaan kerja konstruksi yang sering terjadi adalah Lingkungan Tempat Kerja, Rambu-Rambu Keselamatan dan Pekerja dan Cara kerja. [Elisa, 2017:36].
2. Lemahnya pengawasan dalam proses pemberian kredit kepada karyawan PT. X Group menyebabkan tingginya kredit macet. Dalam menyalurkan kreditnya, PT. X Group haruslah pintar dalam menilai para nasabah dimasa yang akan datang apakah akan menguntungkan atau tidak. Faktor ini sangatlah penting bagi pihak perusahaan karena hal ini akan menunjukkan bahwa layak atau tidaknya suatu usaha atau individu yang akan diberikan pinjaman atau kredit, pada penelitian ini digunakan teknik data mining klasifikasi dengan metode C4.5 untuk mengetahui apakah nasabah tergolong nasabah lancar ataupun tidak. Dengan penelitian ini diharapkan dapat membantu pihak perusahaan dalam membaca pola pembayaran dari nasabahnya sehingga dapat menentukan apakah nasabah tersebut layak mendapatkan kredit atau tidak dan menghasilkan rule dari pohon keputusan yang diterapkan pada implementasi sistem klasifikasi data nasabah kredit di PT. X Group. [Pratama dkk, 2018: 121].
3. Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga, pada koperasi permasalahan kredit merupakan permasalahan manajemen, dimana jika banyak nasabah yang menunggak dalam pembayaran maka akan

mengganggu system keuangan yang ada, untuk itulah penelitian ini menerapkan proses analisa kredit nasabah terlebih dahulu sebelum diambil sebuah keputusan pemberian kredit, analisa keputusan memberikan kredit menggunakan algoritma klasifikasi C4.5 dan Naïve Bayes dimana kedua algoritma tersebut dilakukan penilaian, mana algoritma yang paling akurat dalam menganalisa kemampuan nasabah dalam membayar kredit, analisa berdasarkan data history. Hasil yang didapatkan dari perbandingan kedua algoritma tersebut, bahwa tingkat akurasi yang lebih baik adalah menganalisa menggunakan algoritma klasifikasi C4.5 yaitu 88.90 % sedangkan untuk tingkat akurasi menggunakan algoritma klasifikasi Naïve Bayes yaitu 80.00%. [Marsipah, 2016: 187]

4. Penelitian ini bertujuan untuk mengetahui bagaimana Sistem dan Prosedur Kepemilikan Rumah Subsidi Kredit (KPR) Cabang Manado PT. Bank Tabungan Negara. Jenis penelitian ini adalah deskriptif dengan pendekatan kualitatif. Teknik pengumpulan data yang dilakukan adalah wawancara, observasi dan dokumentasi. Langkah-langkah yang diambil untuk menganalisis data yang diperoleh dilakukan di dua fase, mereka (1) menganalisis bagaimana Sistem dan Prosedur Rumah Subsidi Kredit Kepemilikan (KPR) di Cabang Mandao PT. Bank Tabungan Negara (BTN) sudah baik, dan (2) menganalisis apakah faktor-faktor yang menyebabkan kredit macet di Kepemilikan Rumah Kredit (KPR) Cabang Manado PT. Bank Tabungan Negara dan bagaimana penyelesaiannya kredit macet dilakukan oleh Cabang Manado PT. Bank Tabungan Negara. Hasil dari Penelitian menunjukkan bahwa Sistem dan Prosedur Kredit Kepemilikan Rumah Subsidi (KPR) baik dan cocok dengan standar yang diterapkan oleh pemerintah. Faktor itu menyebabkan kredit macet dalam Sistem dan Prosedur Kredit Pemilikan Rumah (KPR) di Manado Cabang PT. Bank Tabungan Negara sakit debitur, debitur tidak punya pekerjaan, rumah itu bukan standar, dan karakter debitur. Upaya penyelesaian adalah untuk mengumpulkan dan jika itu tidak dapat ditagih maka keputusan diambil dengan lelang. Kata kunci: Sistem dan Prosedur, Kredit Kepemilikan Rumah Subsidi, dan Kredit Buruk. [Takalamingan Dkk, 2018:830]

Berdasarkan penelitian-penelitian diatas dapat disimpulkan bahwa data transaksi perusahaan maupun organisasi masih belum dimanfaatkan secara maksimal sehingga menghasilkan tumpukan data yang memakan memori penyimpanan. Jika dianalisa dengan baik kumpulan data tersebut dapat menghasilkan informasi yang bermanfaat. Salah satunya analisa menggunakan algoritma C4.5 yang dapat menemukan informasi bermanfaat.

BAB III

METODOLOGI PENELITIAN

3.1. Teknik Pengumpulan Data

Dalam Metodologi Penelitian adapun teknik pengumpulan data yang akan dibahas dalam penulisan ini adalah sebagai berikut.

1. Observasi

Merupakan metode yang dilakukan penulis dengan cara mendatangi langsung tempat riset yang ingin di teliti oleh penulis. Penulis melakukan pengamatan langsung ke Perumahan Cahaya Darussalam 2 Bekasi dengan beberapa acuan yaitu menganalisa dan mengamati proses kegiatan pengajuan kredit rumah.

2. Wawancara

Untuk mendapatkan data – data yang benar dan akurat, maka dilakukan tanya jawab secara langsung kepada Marketing dan Developer Cahaya Darussalam 2 terkait dengan masalah yang sudah dibahas pada bab 1.2. Sehingga masalah tersebut dapat diselesaikan yang ada di Perumahan Cahaya Darussalam 2. Salah satu hal yang di wawancara misalnya bertanya kepada marketing pemasaran tentang bagaimana proses pengajuan pembelian rumah sampai mendapatkan hasil kredit..

3. Studi Pustaka

Merupakan metode yang digunakan penulis sebagai pendukung dan referensi. Media untuk studi pustaka yang digunakan peneliti yaitu: Buku

dan jurnal yang berhubungan dengan penelitian dan penulisan. Hal ini dilakukan untuk membantu peneliti dalam menentukan landasan berpikir dan sebagai pijakan yang kuat dalam membangun kerangka berpikir.

3.2. Tahapan Data Mining

Data mining adalah proses untuk mengumpulkan data, pemakaian data, historis data untuk menemukan pola atau hubungan dalam kumpulan data yang besar. Tahapan yang dilakukan secara umum adalah pembersihan data, integrasi data, transformasi data, ekstrasi pola, evaluasi pola dan presentasi pola , berdasarkan tahapan yang disebutkan dapat diuraikan sebagai berikut :

- a. Pembersihan data (untuk membuang data yang tidak konsisten dan noise).

Sebelum menggunakan data untuk dianalisa dengan algoritma C4.5 terdapat tahap pembersihan data. Tahapan yang dilakukan adalah pemilihan data mana saja yang akan digunakan, dalam penelitian ini data yang digunakan adalah data transaksi penjualan rumah baik yang lolos tahap kredit maupun tidak.

- b. Integrasi data (penggabungan data dari beberapa sumber) Pada tahap integrasi data dilakukan penggabungan data transaksi dari beberapa tanggal pada bulan oktober yang ada pada laporan data transaksi detail yang ada.

- c. Transformasi Data (data diubah menjadi bentuk yang sesuai untuk di-mining) Pada tahap transformasi data adalah proses mengubah data transaksi penjualan rumah pada Perumahan Cahaya Darussalam 2 dengan

RapidMiner Studio 8.2, yaitu diubah menjadi bentuk data tabular yang berisikan angka biner yaitu 1 dan 0 dengan ekstensi ***.XLS (Ms.Excel).

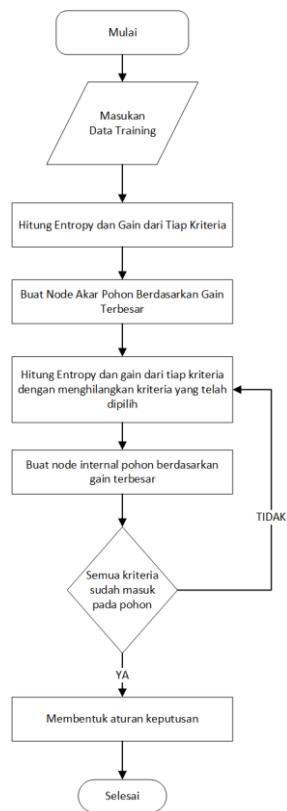
- d. Aplikasi teknik data mining, proses ekstraksi pola dari data yang ada.Pada tahap ini setelah data ditransformasikan maka akan dilakukan proses data mining untuk menemukan kaidah asosiasi yang dibutuhkan dengan menggunakan algoritma C4.5. Pada proses ini peneliti menghitung menggunakan software RapidMiner Studio 8.2.
- e. Evaluasi pola yang ditemukan / intepretasi. Pada tahap terakhir dalam proses data mining ini peneliti mengidentifikasi pola-pola menarik yang bisa diterjemahkan kedalam bahasa yang mudah dipahami.
- f. Presentasi pengetahuan (dengan teknik visualisasi). Pada tahap ini peneliti melakukan penyajian pengetahuan mengenai teknik yang digunakan untuk memperoleh pengetahuan yang mudah dipahami oleh user.

3.3. Metode Algoritma C4.5

Pada penerapan menggunakan algortima C4.5 dilakukan dalam beberapa tahap yaitu :

- a. Data training dimasukan
- b. Menghitung *Information Gain* dan *entropy* dari masing-masing kriteria data training yang ada unutk menentukan kriteria mana yang akan menjadi akar pada pohon keputusan/
- c. Membuat node akar dari pemilihan kriteria dengan memilih Information Gain terbesar.

- d. Menghitung *Information Gain* dan *entropy* dari masing masing kriteria dengan menghilangkan kriteria yang sudah dipilih sebelumnya.
- e. Melakukan pengecekan pada semua kriteria apakah kategori pada kriteria yang memiliki nilai *Information Gain* tertinggi sudah masuk pada kelas yang sama. Jika belum, maka ulangi proses (d) dan (e), jika sudah maka lanjutkan pada proses berikutnya.
- f. Kemudian aturan keputusan dibuat mengikuti pohon yang telah dibentuk sebelumnya.



Sumber: Hasil Penelitian (2020)

Gambar III. 1
Tahap Penerapan Algoritma C4.5

3.4. Pengujian Hasil

Pada tahap ini setelah melakukan metode *data mining* dan metode algoritma C4.5 melakukan pengujian dan hasil perancangan aturan klasifikasi menggunakan *software data mining RapidMiner Studio 8.2*. Sistem diuji dengan prosedur-prosedur untuk melakukan eksplorasi dan permodelan dari data-data yang ada sehingga mendapatkan informasi yang dibutuhkan.

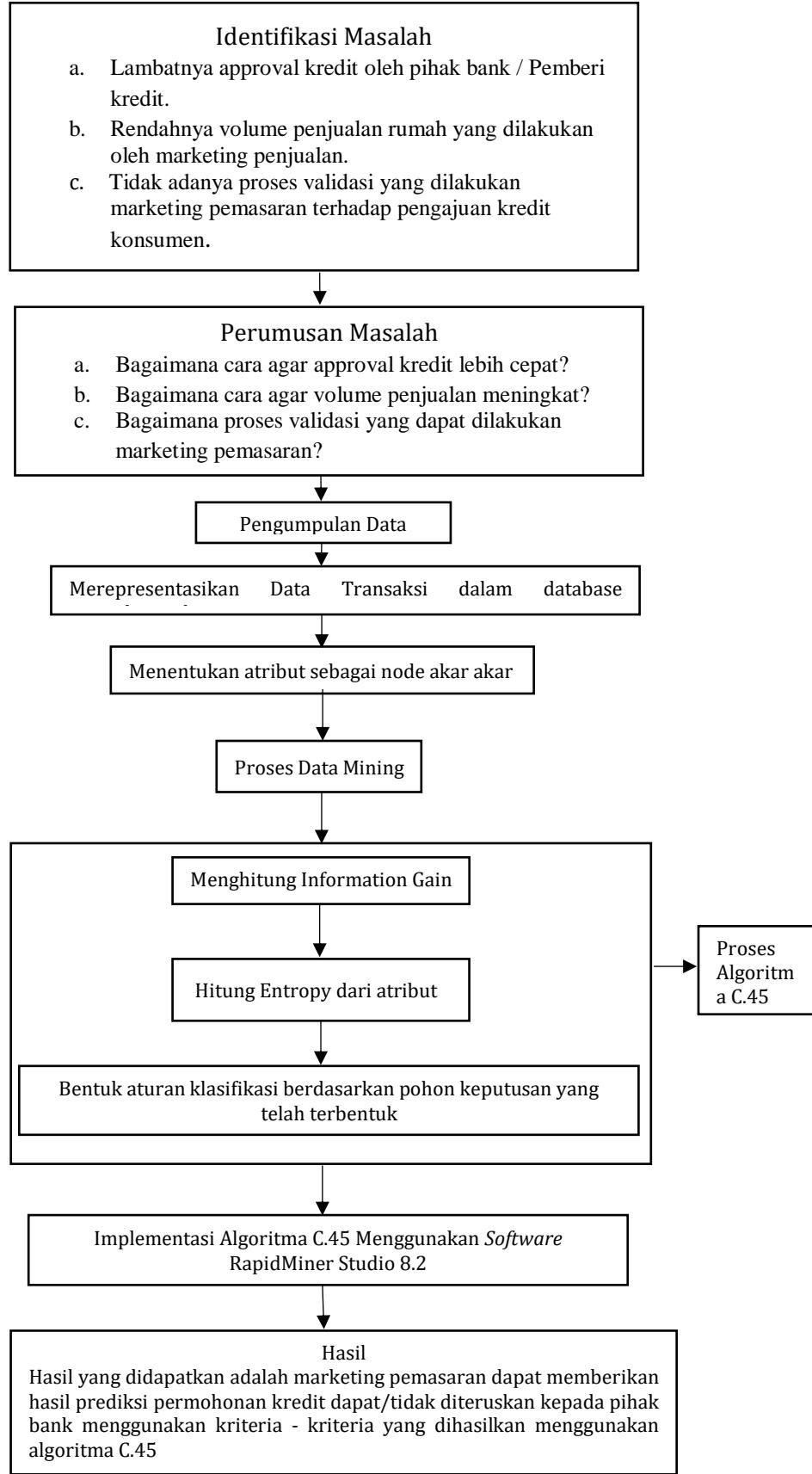
3.5. Kerangka Pemikiran

Pada tahap ini penulis membuat kerangka pemikiran yang berutujuan untuk memecahkan suatu permasalahan yang ada pada Perumahan Cahaya Darussalam 2 Bekasi, Bentuk kerangka pemikiran ditujukan pada gambar III.1.

3.5.1. Deskripsi Kerangka Pemikiran

1. Permasalahan
2. Masalah yang ditemukan pada perumahan Cahaya Darussalam 2 Bekasi adalah lambatnya approval kredit oleh pihak bank / Pemberi kredit, Rendahnya volume penjualan rumah yang dilakukan oleh marketing penjualan. maka dari itu peneliti membantu mencari factor apa saja yang dapat menentukan persetujuan kredit rumah dan memberikan hasil nya kepada konsumen.
3. Analisa Masalah
4. Analisa Masalah yang dapat disimpulkan adalah proses pencarian kriteria yang menentukan persetujuan kredit rumah subsidi berdasarkan data penjualan rumah subsidi. Adapun langkah-langkah yang digunakan yaitu :

5. Proses Algoritma C4.5
6. Pengumpulan Data penjualan rumah yang diolah menjadi aturan
7. Klasifikasi
8. Memilih atribut sebagai akar
9. Membuat cabang untuk tiap tiap nilai
10. Mengelompokan kasus kedalam cabang
11. Implementasi Algoritma C4.5
12. Penerapan penghitungan algoritma C4.5 menggunakan *software RapidMiner Studio 8.2* mempermudah dan mempercepat penemuan hasil aturan asosiasi yang diharapkan menghasilkan kriteria yang dapat menentukan persetujuan kredit rumah.
13. Hasil
14. Hasil yang didapat dari pencarian kriteria persetujuan kredit rumah adalah mempercepat konsumen mendapatkan hasil persetujuan kredit rumah.



Sumber: Hasil Penelitian (2020)

Gambar III. 2
Kerangka Pemikiran

BAB IV

HASIL DAN PEMBAHASAN

4.1 Tinjauan Perusahaan

Dalam tinjauan perusahaan ini adapun sejarah, visi – misi dan tujuan PT Cahaya Indorahmat Pratamajaya yang merupakan developer dari Perumahan Cahaya Darussalam 2 Bekasi dan akan dibahas dalam penulisan ini adalah sebagai berikut:

4.1.1. Sejarah Perusahaan

PT CAHAYA INDORAHMAT PRATAMAJAYA adalah perusahaan yang bergerak di bidang jasa pengembang perumahan, percetakan print digital, travel haji & umroh serta kegiatan usaha terkait lainnya, seperti Lembaga Pendidikan.

PT CAHAYA INDORAHMAT PRATAMAJAYA adalah perusahaan legal yang sudah mendapatkan pengesahan badan hukum perseroan berdasarkan keputusan menteri hukum dan hak asasi manusia Republik Indonesia dan Daftar Perseroan.

Dengan semua legalitas dan pendukung yang ada, kami mempersembahkan sebuah karya anak bangsa yaitu PERUMAHAN CAHAYA DARUSSALAM 2 BEKASI. Sebuah hunian yang luar biasa Asri, yang akan memudahkan semua orang untuk memiliki rumah tinggal yang terjangkau dan berkualitas. Sebagai perusahaan yang sedang tumbuh dan

berkembang dengan semangat entrepreneurship, komitmen pada kualitas, pelayanan, kepuasan dan kesuksesan pelanggan adalah tujuan utama kami

Adapun Visi, Misi dan Tujuan “PT CAHAYA INDORAHMAT PRATAMA JAYA” adalah sebagai berikut:

1. Visi

Menjadi perusahaan property terpercaya dalam menyediakan kawasan permukiman yang lebih baik dan berkesinambungan di Indonesia.

2. Misi

- a. Mengembangkan dan menata permukiman beserta lingkungannya secara berkesinambungan.
- b. Menjadi pengembang property yang inovatif dan terpercaya dalam meningkatkan keuntungan bagi semua pihak

3. Tujuan

Memudahkan semua orang untuk memiliki rumah tinggal yang terjangkau dan berkualitas. Sebagai perusahaan yang sedang tumbuh dan berkembang dengan semangat entrepreneurship, komitmen pada kualitas, pelayanan, kepuasan dan kesuksesan pelanggan.

Dan salah satu produk dari PT Cahaya Indorahmat Pratamajaya adalah Perumahan Cahaya Darussalam 2 Bekasi, Perumahan Cahaya Darussalam 2 merupakan hunian bersubsidi yang lokasinya berada di Jl. Raya Jejalen-Jabir, Sumber Jaya, Tambun Selatan, Bekasi, Jawa Barat ini merupakan bagian dari program Sejuta Rumah yang digagas pemerintah

pusat. Unit yang dikembangkan sebanyak 700 unit yang pembangunannya dimulai sejak 2015. Lokasi yang cukup strategis menjadi daya tarik yang membuat hunian ini jadi incaran para pencari rumah. Tambun Selatan termasuk kecamatan di Kabupaten Bekasi yang prospektif. Sejumlah infrastruktur yang dibangun di Bekasi, seperti pembangunan jalan tol layang Jakarta-Cikampek dan jalan tol Cibitung-Cilincing yang pintu masuknya dekat perumahan Darussalam 2. Saat ini akses menuju ke Cahaya Darussalam 2 bisa masuk dari Jl. Tol Jakarta-Cikampek dan keluar di Bekasi Timur yang jaraknya sekitar 10 km dari hunian dengan waktu tempuh sekitar 30-40 menit. Jika hendak menggunakan sarana transportasi umum, pilihan terbaik adalah menumpang KRL komuter lalu keluar di Stasiun Tambun. Jarak tempuh dari stasiun ke perumahan sekitar 7 km atau sekitar 10-20 menit berkendara.

Cahaya Darussalam 2 mudah ditemukan karena lokasinya persis di samping perumahan De Green De Jaleen. Gerbang Cahaya Darussalam 2 yang mengusung desain modern terlihat menarik untuk kelas perumahan subsidi. Jalan depan perumahan bisa dilalui kendaraan roda empat. Tersedia fasilitas masjid, area komersial berupa ruko dan taman bermain anak. Cahaya Darussalam 2 memiliki tipe 29/60 dengan 2 kamar tidur dan 1 kamar mandi. Harga saat launching sekitar Rp 130 jutaan. Harga jual rumah seken di Cahaya Darussalam 2 kemungkinan berada dalam kisaran harga Rp 250 juta hingga 300 jutaan.

4.1.2. Struktur Organisasi dan Fungsi

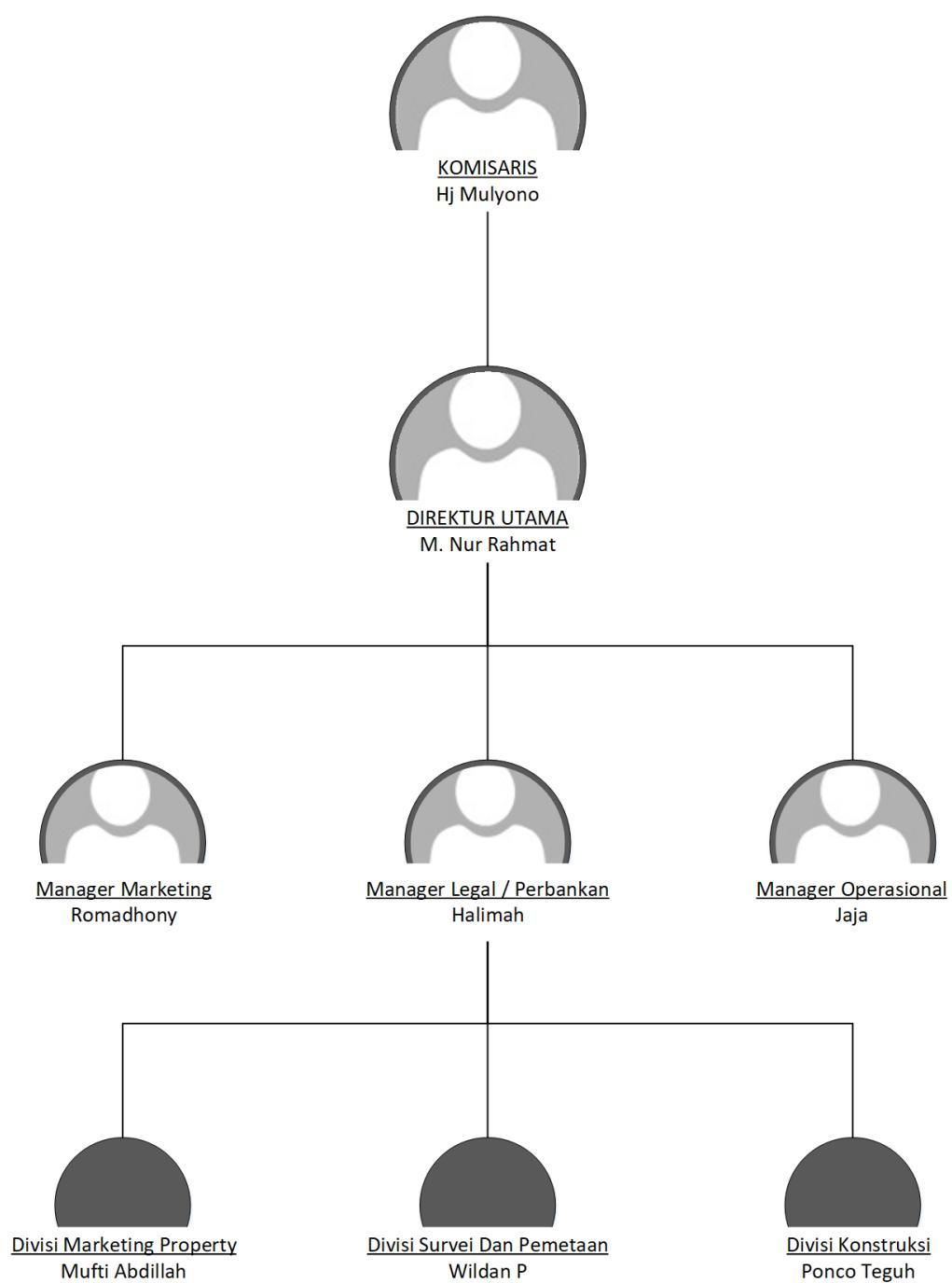
Setiap perusahaan memiliki struktur organisasi dan fungsinya masing-masing. Berikut adalah struktur organisasi beserta tugas dan fungsinya pada Perumahan Cahaya Darussalam 2 Bekasi yang ditunjukan pada Gambar xx. Struktur Organisasi Perumahan Cahaya Darussalam 2 Bekasi.

4.1.2.1. Struktur Organisasi

Struktur organisasi juga berarti susunan dari berbagai macam komponen atau unit kerja dalam sebuah organisasi. Dalam struktur organisasi terdapat pembagian kerja dan bagaimana fungsi atau kegiatan-kegiatan berbeda yang telak dikoordinasikan dan juga terdapat adanya berbagai spesialisasi dari sebuah pekerjaan, saluran perintah ataupun penyampaian laporan. Dan ketika akan mengajukan izin organisasi, para pengurus harus melampirkan struktur organisasi berikut nama-nama pengurusnya.

Struktur ini mengandung unsur-unsur spesialisasi kerja, standarisasi kerja, koordinasi, sentralisasi atau desentralisasi dalam pembuatan keputusan dan besar (ukuran) satuan kerja. Agar tujuan organisasi untuk membangun ekonomi yang kuat “Perumahan Cahaya Darussalam 2 Bekasi” Tercapai dengan baik maka perlu dibuatkan bagan organisasi.

Jadi struktur organisasi mempunyai peranan penting dalam suatu kelompok, seperti kejelasan tanggung jawab, kejelasan kedudukan, kejelasan jalur hubungan, kejelasan uraian tugas.



Sumber: Hasil Penelitian (2020)

Gambar IV. 1

Struktur Organisasi Perumahan Cahaya Darussalam 2 Bekasi.

4.1.2.2. Fungsi

Berikut adalah penjelasan yang dapat diuraikan disini adalah fungsi struktur organisasi yang ada pada Perumahan Cahaya Darussalam 2 Bekasi, yaitu:

a. Komisaris

Tugas dan Fungsinya adalah:

1. Dewan Komisaris melakukan pengawasan atas kebijakan pengurusan, jalannya pengurusan pada umumnya, baik mengenai Perseroan maupun usaha Perseroan, serta memberi nasihat kepada Direksi. Adapun, pengawasan dan pemberian nasihat dilakukan untuk kepentingan Perseroan sesuai dengan maksud dan tujuan Perseroan.
2. Dalam menjalankan tugas pengawasan dan pemberian nasihat kepada Direksi, Dewan Komisaris wajib melakukannya dengan itikad baik, kehati-hatian, dan bertanggung jawab demi kepentingan Perseroan.

3. Dewan Komisaris turut bertanggung jawab secara pribadi atas kerugian Perseroan, apabila yang bersangkutan bersalah atau lalai dalam menjalankan tugasnya sebagaimana mestinya.

b. Direktur Utama

Tugas dan Fungsinya adalah:

1. Direksi bertanggung jawab atas pengurusan Perseroan untuk kepentingan perseroan dan sesuai dengan maksud dan tujuan perseroan, sesuai dengan kebijakan yang dipandang tepat dalam batas yang telah ditentukan dalam Undang-Undang atau Anggaran Dasar.
2. Direksi wajib beritikad baik dan bertanggung jawab dalam melakukan pengurusan dalam Perseroan.
3. Direksi wajib mewakili perseroan baik di luar maupun di dalam pengadilan.
4. Direksi juga wajib membuat dan memelihara daftar pemegang saham, risalah RUPS, dan risalah rapat direksi, menyelenggarakan pembukuan perseroan, melaporkan kepemilikan sahamnya.
5. Jika mengalami kelalaian atau kerugian, setiap anggota Direksi bertanggungjawab penuh secara pribadi atas kerugian Perseroan apabila yang bersangkutan bersalah atau lalai dalam menjalankan tugasnya. Jika Direksi terdiri atas 2

(dua) anggota Direksi atau lebih, maka tanggung jawab tersebut berlaku secara tanggung renteng bagi setiap anggota Direksi.

c. Manager Marketing

Tugas dan Fungsinya adalah:

1. Manajer pemasaran bertanggung-jawab terhadap manajemen bagian pemasaran
2. Manajer pemasaran bertanggung-jawab terhadap perolehan hasil penjualan dan penggunaan dana promosi
3. Manajer pemasaran sebagai koordinator manajer produk dan manajer penjualan
4. Manajer pemasaran membina bagian pemasaran dan membimbing seluruh karyawan dibagian pemasaran
5. Manajer pemasaran membuat laporan pemasaran kepada direksi

d. Manager Legal /Perbankan

Tugas dan Fungsinya adalah:

1. Merancang & Mereview Kontrak/ Perjanjian Kerja
2. Sistematika Kontrak/ Perjanjian Kerja
3. Komparisi, Premis, Isi & Klausul Kontrak/ Perjanjian Kerja
4. Jenis-jenis Akte & Teknik Pembuatan Akte
5. Proses Pendirian Badan Usaha (PT, CV, Fa)

6. Jenis-jenis Badan Usaha & Dokumen Pelengkapnya
7. Proses Pengurusan Izin-izin Perusahaan (NPWP, SIUP, TDP, AMDAL/ UKL-UPL)
8. Keterangan Domisili Perusahaan
9. Proses Pembubaran/ Likuidasi Perusahaan
10. Proses Kepailitan
11. Pembuatan Opini Yuridis
12. Pengurusan Hak-hak Atas Tanah
13. Tata Cara Pendaftaran & Persyaratan Hak-hak Atas Tanah dan Proses Pengurusan Pengikatan Kredit

e. Manager Operasional

Tugas dan Fungsinya adalah:

1. Mengelola dan meningkatkan efektifitas dan efisiensi operasi perusahaan
2. Memangkas habis biaya-biaya operasi yang sama sekali tidak menguntungkan perusahaan
3. Meneliti teknologi baru dan metode alternatif efisiensi
4. Mengawasi produksi barang atau penyediaan jasa
5. Mengawasi tata letak operasional , persediaan dan distribusi barang
6. Membuat atau merencanakan pengembangan operasi dalam jangka pendek maupun panjang

7. Meningkatkan sistem operasional, proses dan kebijakan dalam mendukung visi dan misi perusahaan
8. Melakukan pertemuan rutin dengan direktur eksekutif secara berkala
9. Melakukan pencairan cek untuk biaya agen
10. Mengatur anggaran dan mengelola biaya
11. Mengelola program jaminan kualitas

f. Divisi Marketing Property

Tugas dan Fungsinya adalah:

1. Membuat data/listing properti yang dijualnya
2. Membuat iklan melalui media promosi (brosur, iklan koran, spanduk, banner dll)
3. Melakukan proses prospektng dan presentasi di depan calon konsumen
4. Mengantar calon konsumen ke lokasi yang diinginkan
5. Memastikan kerjasama dengan pihak notaris
6. Memastikan kerjasama dengan pihak perbankan (bila berkaitan dengan KPR)
7. Memastikan terkumpulnya persyaratan konsumen (KTP suami istri, surat nikah) dan persyaratan KPR lainnya.
8. Mengecek bangunan, sertifikat, IMB, bukti lunas PBB dll.
9. Mengingatkan akan kewajiban pajak masing-masing pihak.

10. Membantu kelancaran proses transaksi jual beli

g. Divisi Survei Dan Pemetaan

Tugas dan Fungsinya adalah:

Menentukan pemetaan yang akurat dalam bentuk perencanaan tahap awal yang hasil akhir nya adalah pemetaan topografi adapun hal hal yang perlu di rencakan seperti

1. Persiapan berupa kantor, administrasi, pengadaan peta dasar dan peta kerja, peralatan dan personil,
2. Lapangan berupa mobilisasi, orientasi,
3. Pelaksanaan berupa pematokan dan pemasangan tugu/bench mark, pengukuran kerangka dasar dan pengukuran detail situasi.
4. Pekerjaan studio berupa pengolahan data, editing data dan penggambaran, plotting peta hasil penggambaran, serta pelaporan.

h. Divisi Konstruksi

Tugas dan Fungsinya adalah:

1. Meminta laporan dan penjelasan tentang pelaksanaan pekerjaan kepada pelaksana proyek baik secara lisan maupun tulisan.
2. Menghentikan atau menolak hasil pekerjaan apabila dalam pelaksanaan menyimpang dari spek yang telah ditentukan.
3. Mengesahkan adanya perubahan baik didalam desain maupun pekerjaan.

4. Memberikan keputusan terhadap perubahan waktu pelaksanaan dengan mempertimbangkan segala resiko yang akan dihadapi.
5. Mengarahkan, mengelola, serta mengkoordinasikan pelaksanaan kontraktor dalam aspek mutu, biaya, waktu, dan keselamatan dalam pekerjaan.
6. Mengadakan rapat koordinasi yang dihadiri oleh konsultan perencana dan kontraktor. Rapat diadakan seminggu sekali.
7. Memeriksa gambar detail pelaksanaan (shop drawing).
8. Membuat laporan kemajuan pekerjaan di lapangan

4.2. Pembahasan

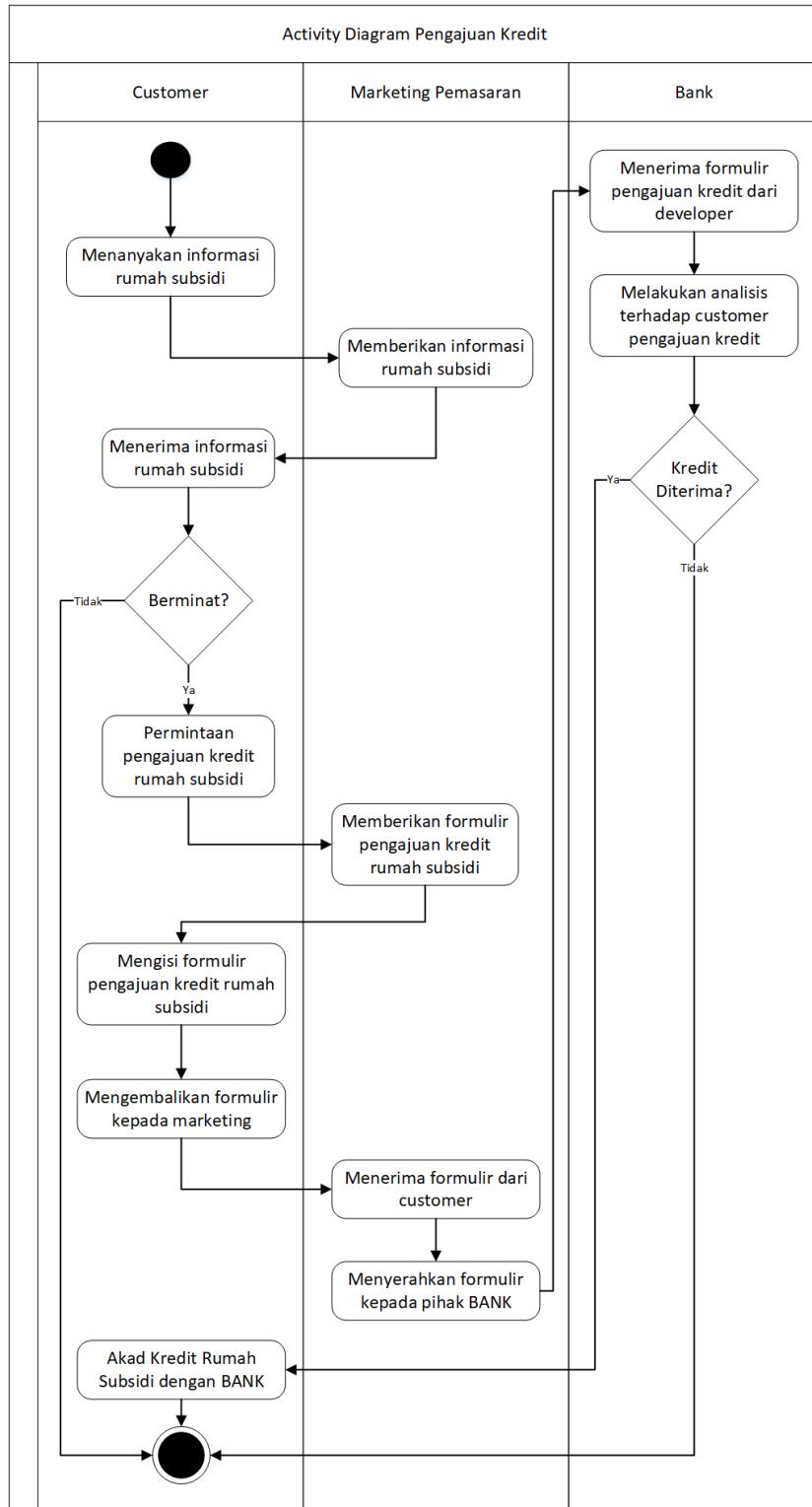
Pada sub bab ini akan dijelaskan bagaimana proses bisnis dan langkah – langkah dalam menentukan algoritma C4.5, berikut adalah langkah langkahnya:

4.2.1 Proses Bisnis Sistem

Pada proses pengajuan kredit rumah subsidi, customer datang ke pemasaran untuk menanyakan seputar rumah subsidi yang di pasarkan di Perumahan Cahaya Darussalam 2 Bekasi, setelah itu markering pemasaran memberikan formulir pengajuan kredit kepada konsumen untuk di isi, kemudian formulir pengajuan kredit dikembalikan kepada marketing pemasaran, dan marketing pemasaran langsung menyerahkan formulir pengajuan kredit dari customer kepada pihak bank untuk di analisis apakah customer tersebut dapat melakukan pinjaman KPR

subsidi atau tidak. Hasil analisis untuk perstujuan dari pihak BANK memerlukan kurun waktu 2-3 bulan, dan jika pengajuan kredit rumah subsidi tersebut diterima oleh bank maka customer melakukan akad kredit dengan pihak BANK dan akan mulai melakukan cicilan kredit sesuai dengan waktu yang sudah di tentukan.

Proses bisnis pengajuan kredit rumah subsidi oleh customer Perumahan Cahaya Darussalam 2 Bekasi, dapat dilihat pada gambar 2:



Sumber: Hasil Penelitian (2020)

Gambar IV. 2

Activity Diagram Proses Pengajuan Kredit Rumah Subsidi.

4.2.2 Atribut Keputusan

Berdasarkan data transaksi pengajuan kredit rumah subsidi yang telah dikumpulkan, selanjutnya dilakukan pemilihan atribut keputusan, adapun tabel atribut keputusan ditunjukkan pada tabel IV.1

Tabel IV. 1
Atribut sample decision tree algoritma C4.5

No	Nama Atribut	Kode Atribut
1	Nama Customer	A1
2	Jenis Kelamin	A2
3	Usia	A3
4	Pendidikan	A4
5	Profesi	A5
6	Total Masa Kerja	A6
7	Status Kerja	A7
8	Status Pernikahan	A8
9	Jumlah Tanggungan	A9
10	Status Tempat Tinggal	A10
11	Total Penghasilan	A11
12	Kredit Lainnya	A12
13	Tunggakan	A13
14	Keterangan	A14

Sumber: Hasil Penelitian (2020)

4.2.3 Pengelompokan Nilai Atribut Keputusan

Atribut sudah ditemukan, selanjutnya adalah pengelompokan nilai dari atribut tersebut, Berikut adalah pengelompokan nilai atribut jenis kelamin ditunjukkan pada Tabel IV.2:

Tabel IV. 2
Pengelompokan nilai atribut jenis kelamin.

No	Jenis Kelamin	Kode Nilai Atribut
1	Pria	JK1
2	Wanita	JK2

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Usia ditunjukkan pada Tabel IV.3:

Tabel IV. 3
Pengelompokan nilai atribut usia.

No	Usia	Kode Nilai Atribut
1	<25 Tahun	UA1
2	25-35 Tahun	UA2
3	36-45 Tahun	UA3
4	>45 Tahun	UA4

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Pendidikan ditunjukkan pada Tabel IV.4:

Tabel IV. 4
Pengelompokan nilai atribut pendidikan.

No	Pendidikan	Kode Nilai Atribut
1	SMA/SLTA	PD1
2	Diploma	PD2
3	S1	PD3

Sumber: Hasil Penelitian (2020).

Pengelompokan nilai atribut Profesi ditunjukan pada Tabel IV.5:

Tabel IV. 5
Pengelompokan nilai atribut Profesi

No	Profesi	Kode Nilai Atribut
1	PNS / Instansi / Departemen / Pemda	PR1
2	Swasta Asing / PMA	PR2
3	Swasta Besar / Menengah	PR3
4	Wiraswasta Besar / Menengah	PR4

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Total Masa Kerja ditunjukan pada Tabel IV.6:

Tabel IV. 6
Pengelompokan nilai atribut Total Masa Kerja.

No	Total Masa Kerja	Kode Nilai Atribut
1	<1 Tahun	MK1
2	1-2 Tahun	MK2
3	>2 Tahun	MK3

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Status Kerja ditunjukan pada Tabel IV.7:

Tabel IV. 7
Pengelompokan nilai atribut Status Kerja.

No	Status Kerja	Kode Nilai Atribut
1	Kontrak	SK1
2	Tetap	SK2
3	Pengusaha	SK3

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Status Pernikahan ditunjukan pada

Tabel IV.8:

Tabel IV. 8
Pengelompokan nilai atribut Status Pernikahan.

No	Status Pernikahan	Kode Nilai Atribut
1	Menikah	SP1
2	Belum Menikah	SP2

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Jumlah Tanggungan ditunjukan pada Tabel IV.9:

Tabel IV. 9
Pengelompokan nilai atribut Jumlah Tanggungan.

No	Jumlah Tanggungan	Kode Nilai Atribut
1	Tidak Ada	JT1
2	1-3 Orang	JT2
3	>3 Orang	JT3

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Status Tempat Tinggal ditunjukan pada Tabel IV.10:

Tabel IV. 10
Pengelompokan nilai atribut Status Tempat Tinggal.

No	Status Tempat Tinggal	Kode Nilai Atribut
1	Dinas	ST1
2	Keluarga	ST2
3	Milik Sendiri	ST3
4	Sewa/Kontrak	ST4

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Total Penghasilan ditunjukan pada

Tabel IV.11:

Tabel IV. 11
Pengelompokan nilai atribut Total Penghasilan

No	Total Penghasilan	Kode Nilai Atribut
1	< Rp.2.000.000	TP1
2	Rp.2.000.000 – Rp.3.500.000	TP2
3	Rp.3.600.000 – Rp.4.000.000	TP3
4	> Rp.4.000.000	TP4

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Kredit Lainnya ditunjukan pada

Tabel IV.12:

Tabel IV. 12
Pengelompokan nilai atribut Kredit Lainnya

No	Kredit Lainnya	Kode Nilai Atribut
1	Tidak Ada	KL1
2	Kredit Barang	KL2
3	Kredit Kendaraan	KL3

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut Tunggakan ditunjukan pada Tabel IV.13:

Tabel IV. 13
Pengelompokan nilai atribut Tunggakan.

No	Tunggakan	Kode Nilai Atribut
1	Tidak Ada	TN1
2	Tidak Menunggak	TN2
3	Menunggak	TN3

Sumber: Hasil Penelitian (2020)

Pengelompokan nilai atribut keterangan ditunjukan pada Tabel IV.14:

Tabel IV. 14
Pengelompokan nilai atribut Keterangan.

No	Keterangan	Kode Nilai Atribut
1	Lolos	KN1
2	Tidak Loloas	KN2

Sumber: Hasil Penelitian (2020)

4.2.4 Data Pengajuan Kredit Pada Perumahan Cahaya Darussalam 2

Atribut untuk sample penelitian sudah ditemukan, selanjutnya menentukan data yang akan digunakan dalam penelitian ini, data yang digunakan bisa dilihat pada Tabel IV.15:

Tabel IV. 15
Data Pengajuan Kredit Rumah Subsidi 2019.

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
1	AR. REZA FAHMI MUHAMMAD	JK1	UA2	PD1	PR2	MK2	SK2	SP2	JT1	ST2	TP1	KL2	TN2	KN1
2	MASNITA PANJAITAN	JK2	UA2	PD1	PR4	MK3	SK3	SP1	JT2	ST4	TP1	KL1	TN1	KN1
3	RIKO SETIAWAN	JK1	UA3	PD3	PR4	MK2	SK3	SP1	JT2	ST4	TP1	KL1	TN1	KN1
4	DINI DWI ISMYATI	JK2	UA2	PD2	PR4	MK2	SK3	SP1	JT2	ST4	TP1	KL1	TN1	KN1
5	NURYANA	JK1	UA1	PD2	PR4	MK3	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
6	SANTIH SUSILAWATI	JK2	UA1	PD2	PR3	MK2	SK1	SP2	JT1	ST4	TP2	KL2	TN2	KN2
7	RICHARD NAI NGGOLAN	JK1	UA2	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
8	EDI MULYONO	JK1	UA2	PD1	PR4	MK3	SK3	SP1	JT2	ST4	TP4	KL1	TN1	KN1
9	MUH ANGGA PUTRA	JK1	UA2	PD2	PR4	MK3	SK3	SP1	JT2	ST4	TP4	KL1	TN1	KN1
10	NURIMAN	JK1	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST4	TP3	KL1	TN1	KN1
11	AHGY MARTIN PRAMANA	JK1	UA1	PD2	PR4	MK2	SK3	SP2	JT1	ST3	TP1	KL3	TN2	KN2
12	IVAN PRASETIO	JK1	UA2	PD1	PR1	MK3	SK2	SP1	JT1	ST1	TP1	KL1	TN1	KN1
13	JUNAEDI	JK1	UA2	PD2	PR4	MK2	SK3	SP2	JT1	ST4	TP4	KL1	TN1	KN1
14	ARIS MUNANDAR	JK1	UA2	PD1	PR1	MK3	SK2	SP2	JT1	ST2	TP1	KL1	TN1	KN1
15	DETA NURAPRILIANIE RIZKИ	JK1	UA2	PD1	PR4	MK3	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
16	EKA SUSANTI	JK2	UA2	PD1	PR3	MK2	SK1	SP1	JT2	ST1	TP2	KL1	TN1	KN1

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
17	DESI AGUSTINI	JK2	UA2	PD2	PR2	MK3	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
18	RACHMAT HIDAYAT	JK1	UA3	PD1	PR3	MK2	SK2	SP1	JT2	ST4	TP3	KL3	TN2	KN2
19	NURUL FIKRI	JK1	UA2	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
20	ISNURYANTI	JK2	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
21	SUWARNIATUN	JK1	UA1	PD3	PR3	MK3	SK1	SP2	JT1	ST4	TP2	KL3	TN2	KN2
22	RIKORDIAS DOMINIUS	JK1	UA2	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
23	MEGA ASTUTI WIJAYANTI	JK2	UA2	PD1	PR1	MK2	SK2	SP2	JT1	ST4	TP4	KL1	TN1	KN1
24	RUKMAN HIDAYAT	JK1	UA2	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP1	KL3	TN2	KN2
25	ERBI MAWADAH	JK2	UA1	PD2	PR1	MK2	SK2	SP2	JT1	ST1	TP1	KL1	TN1	KN1
26	SUHERI B RASJID	JK1	UA2	PD2	PR4	MK3	SK3	SP2	JT1	ST4	TP1	KL1	TN1	KN1
27	ARISKA AMELIA	JK2	UA2	PD2	PR4	MK2	SK3	SP1	JT2	ST2	TP1	KL1	TN1	KN1
28	RALIM	JK1	UA2	PD3	PR1	MK2	SK2	SP1	JT2	ST4	TP1	KL1	TN1	KN1
29	NANA JANNATUL MA`NA	JK2	UA2	PD1	PR1	MK2	SK2	SP2	JT1	ST4	TP1	KL3	TN2	KN2
30	ASRI SEPTERIANA	JK2	UA1	PD1	PR4	MK3	SK3	SP2	JT1	ST2	TP1	KL1	TN1	KN1
31	UBAI HAQI	JK1	UA2	PD1	PR3	MK3	SK2	SP1	JT2	ST4	TP2	KL1	TN1	KN1
32	LAKSMITA PUTRI	JK2	UA2	PD1	PR1	MK2	SK2	SP2	JT1	ST4	TP1	KL3	TN2	KN2
33	JOKO PRIYONO	JK1	UA2	PD2	PR3	MK2	SK1	SP2	JT1	ST4	TP3	KL1	TN1	KN1
34	WARDIMAN ARMANSYAH	JK1	UA2	PD3	PR3	MK3	SK2	SP1	JT2	ST4	TP2	KL3	TN2	KN2
35	KARNAH	JK2	UA2	PD1	PR1	MK2	SK2	SP1	JT2	ST4	TP1	KL3	TN2	KN2

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
36	MARSIYAH	JK2	UA3	PD1	PR3	MK2	SK1	SP1	JT2	ST4	TP2	KL1	TN1	KN1
37	SUDIARTO	JK1	UA1	PD2	PR1	MK2	SK2	SP1	JT2	ST4	TP1	KL1	TN1	KN1
38	NARISTOMO PURNAMA HADI	JK1	UA1	PD2	PR1	MK2	SK2	SP2	JT1	ST4	TP1	KL3	TN2	KN2
39	MUHAMMAD FIQRI HASBALLAH	JK1	UA1	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP3	KL3	TN2	KN2
40	FERDIANTONO	JK1	UA1	PD1	PR3	MK2	SK1	SP2	JT1	ST1	TP3	KL1	TN1	KN1
41	AGUNG SURYO SUMARTONO	JK1	UA3	PD2	PR1	MK3	SK2	SP1	JT2	ST3	TP3	KL1	TN1	KN1
42	LIDIA ROSANNA	JK2	UA3	PD3	PR4	MK3	SK3	SP1	JT2	ST4	TP3	KL3	TN2	KN2
43	ELISA NOVARIANI BORU.P	JK2	UA3	PD1	PR4	MK3	SK3	SP1	JT3	ST4	TP3	KL3	TN2	KN2
44	CITRA FRANSISKA PURBA	JK2	UA2	PD2	PR2	MK2	SK2	SP2	JT1	ST2	TP3	KL1	TN1	KN1
45	AGUNG YUWONO	JK1	UA2	PD2	PR2	MK2	SK2	SP2	JT1	ST3	TP1	KL3	TN2	KN2
46	TEGUH PRIYONO	JK1	UA3	PD1	PR3	MK2	SK2	SP1	JT2	ST4	TP3	KL1	TN1	KN1
47	DENNY	JK1	UA1	PD2	PR2	MK3	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
48	MUHAMMAD SUWITO	JK1	UA1	PD2	PR3	MK2	SK1	SP2	JT1	ST4	TP3	KL3	TN2	KN2
49	SAMSON EKENEDI	JK1	UA3	PD3	PR4	MK2	SK3	SP1	JT2	ST4	TP3	KL3	TN2	KN2
50	ANTORI	JK1	UA1	PD1	PR4	MK2	SK3	SP2	JT1	ST2	TP1	KL1	TN1	KN1
51	SANWANI	JK1	UA3	PD3	PR3	MK2	SK1	SP1	JT2	ST4	TP3	KL3	TN2	KN2
52	SHARIF HIDAYATULLOH	JK1	UA3	PD3	PR3	MK2	SK2	SP1	JT2	ST4	TP2	KL3	TN2	KN2
53	SHATYA BUANA	JK1	UA3	PD3	PR3	MK2	SK1	SP1	JT2	ST4	TP3	KL3	TN2	KN2
54	ABDUL HAKIM	JK1	UA2	PD3	PR4	MK2	SK3	SP2	JT1	ST2	TP3	KL1	TN1	KN1

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
55	RIMBUNAN SILABAN	JK1	UA2	PD3	PR4	MK2	SK3	SP2	JT1	ST4	TP3	KL1	TN1	KN1
56	SAPRIYANSAH	JK1	UA3	PD3	PR4	MK3	SK3	SP1	JT2	ST4	TP2	KL3	TN3	KN2
57	AHMAD SUKENDI	JK1	UA2	PD2	PR2	MK2	SK1	SP2	JT1	ST2	TP1	KL3	TN3	KN2
58	YOSRIAN DANIL	JK1	UA4	PD1	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL3	TN3	KN2
59	MOCH.KASIM	JK1	UA2	PD3	PR3	MK2	SK1	SP1	JT2	ST4	TP3	KL1	TN1	KN1
60	MELANI OKTAVIANI	JK2	UA2	PD2	PR3	MK3	SK1	SP2	JT1	ST4	TP2	KL1	TN1	KN1
61	EDI MULYANTO	JK1	UA2	PD1	PR4	MK3	SK3	SP1	JT2	ST4	TP2	KL1	TN1	KN1
62	HERU HARDIYANTO	JK1	UA2	PD1	PR1	MK2	SK2	SP2	JT1	ST1	TP2	KL1	TN1	KN1
63	DONA KRISDIANA	JK2	UA2	PD3	PR3	MK3	SK1	SP2	JT1	ST1	TP3	KL1	TN1	KN1
64	WASPADA DAELI, SH	JK2	UA3	PD2	PR4	MK3	SK3	SP1	JT3	ST4	TP3	KL3	TN3	KN2
65	VEBRINA SARI S.	JK2	UA3	PD1	PR4	MK2	SK3	SP1	JT2	ST4	TP3	KL1	TN1	KN2
66	NURJANAH	JK1	UA2	PD2	PR3	MK2	SK1	SP1	JT2	ST4	TP3	KL1	TN1	KN1
67	RASNAWATI	JK1	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST4	TP4	KL1	TN1	KN1
68	DEDI	JK1	UA1	PD2	PR4	MK3	SK3	SP2	JT1	ST4	TP4	KL1	TN1	KN1
69	FERA FARISTIYA SUMARLIN	JK2	UA2	PD2	PR3	MK2	SK1	SP1	JT2	ST1	TP2	KL3	TN3	KN1
70	RUBES SETIANI	JK2	UA2	PD1	PR3	MK1	SK1	SP1	JT2	ST4	TP2	KL1	TN1	KN2
71	NORMAL TUMANGGOR	JK1	UA2	PD2	PR3	MK2	SK1	SP2	JT1	ST4	TP2	KL1	TN1	KN1
72	AMELIA ROSIDAH	JK2	UA1	PD1	PR4	MK2	SK3	SP2	JT1	ST2	TP2	KL1	TN1	KN1
73	TONI ARDIANSYAH	JK1	UA2	PD1	PR1	MK2	SK2	SP2	JT1	ST4	TP2	KL1	TN1	KN1

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
74	WINDA	JK2	UA2	PD1	PR3	MK1	SK2	SP2	JT1	ST4	TP2	KL1	TN1	KN2
75	BARTOLOMEUS SETIAWAN F	JK1	UA1	PD2	PR4	MK2	SK3	SP2	JT1	ST2	TP3	KL1	TN1	KN1
76	EKA SUSILOWATI	JK2	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
77	SITI AISAH	JK2	UA1	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP3	KL1	TN1	KN2
78	HARDIANTO LIM	JK1	UA2	PD1	PR3	MK2	SK1	SP1	JT3	ST1	TP3	KL1	TN1	KN1
79	NURHAYATI NISSA	JK2	UA3	PD1	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL3	TN3	KN2
80	DEVIT FITRIYANI	JK2	UA2	PD3	PR2	MK2	SK1	SP2	JT1	ST1	TP3	KL1	TN1	KN1
81	IMRAN HANAFI	JK1	UA2	PD2	PR3	MK2	SK1	SP2	JT1	ST1	TP4	KL1	TN1	KN1
82	DEVI RANI	JK2	UA1	PD3	PR4	MK2	SK3	SP2	JT1	ST4	TP3	KL1	TN1	KN1
83	ELSA NURHAYANI	JK2	UA1	PD2	PR3	MK2	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
84	DEDI PURWANTO	JK1	UA2	PD2	PR2	MK2	SK1	SP2	JT1	ST1	TP1	KL1	TN1	KN1
85	AHMAD BUSAHERI	JK1	UA1	PD1	PR2	MK1	SK1	SP2	JT1	ST3	TP1	KL1	TN1	KN2
86	IMAS MASLIAH	JK1	UA3	PD2	PR4	MK2	SK3	SP1	JT2	ST4	TP4	KL1	TN1	KN2
87	AYUDA CHRISTINA SINAMBELA	JK2	UA2	PD1	PR2	MK2	SK2	SP2	JT1	ST2	TP1	KL1	TN1	KN1
88	KUSNANTO	JK1	UA2	PD1	PR3	MK3	SK1	SP2	JT1	ST4	TP2	KL1	TN1	KN2
89	AJI SUHENDI	JK1	UA1	PD2	PR4	MK2	SK3	SP2	JT1	ST2	TP2	KL1	TN1	KN2
90	IYUS RIYAH	JK1	UA1	PD2	PR3	MK2	SK1	SP1	JT3	ST1	TP3	KL1	TN1	KN1
91	VANIE NATALIA	JK2	UA2	PD1	PR4	MK3	SK3	SP2	JT1	ST4	TP3	KL1	TN1	KN1
92	NURHAYATI	JK2	UA3	PD2	PR1	MK2	SK2	SP1	JT2	ST4	TP3	KL1	TN1	KN2

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
93	ELIS AFIYATI	JK2	UA3	PD1	PR1	MK3	SK2	SP1	JT1	ST1	TP2	KL1	TN1	KN1
94	HARJIYANTO	JK1	UA2	PD1	PR4	MK2	SK3	SP2	JT1	ST4	TP2	KL1	TN1	KN1
95	IRFAN FALIHIN	JK1	UA4	PD2	PR3	MK2	SK2	SP1	JT2	ST1	TP3	KL1	TN1	KN2
96	INDRA PURNAMA	JK1	UA1	PD2	PR3	MK1	SK1	SP2	JT1	ST1	TP2	KL1	TN1	KN1
97	AHMAD FIRLI AKMAL	JK1	UA2	PD1	PR1	MK2	SK2	SP1	JT3	ST3	TP3	KL2	TN2	KN1
98	M. NUR SIDIK	JK1	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST4	TP2	KL1	TN1	KN2
99	AHMAD TUFLIKHUN	JK1	UA3	PD2	PR2	MK3	SK1	SP1	JT2	ST2	TP3	KL1	TN1	KN2
100	HERIYANTO	JK1	UA4	PD1	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL1	TN1	KN2
101	JULIANAH	JK2	UA2	PD2	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL1	TN1	KN1
102	HERU YULYANTO	JK1	UA1	PD1	PR1	MK3	SK2	SP2	JT1	ST1	TP3	KL1	TN1	KN1
103	DEDY MULYADI	JK1	UA3	PD3	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL1	TN1	KN1
104	IDRIS MAHMUD	JK1	UA1	PD1	PR3	MK2	SK1	SP1	JT3	ST1	TP3	KL1	TN1	KN1
105	WARJAN	JK1	UA3	PD2	PR3	MK2	SK2	SP1	JT1	ST4	TP2	KL1	TN1	KN2
106	TEGUH SANTOSO	JK1	UA2	PD1	PR3	MK2	SK2	SP2	JT1	ST4	TP2	KL1	TN1	KN1
107	TRI ASMORO ADIYANTO	JK1	UA2	PD1	PR4	MK1	SK3	SP1	JT1	ST4	TP3	KL1	TN1	KN1
108	AFWAN SOFWAN	JK1	UA3	PD1	PR4	MK2	SK3	SP1	JT2	ST3	TP3	KL1	TN1	KN1
109	REZA RAMADAN	JK1	UA3	PD1	PR3	MK1	SK1	SP1	JT3	ST4	TP3	KL1	TN1	KN1
110	NANO SUKARDI	JK1	UA2	PD1	PR3	MK3	SK1	SP2	JT1	ST4	TP3	KL1	TN1	KN2
111	HERY HERMAWANTO	JK1	UA3	PD2	PR3	MK1	SK1	SP1	JT2	ST1	TP4	KL1	TN1	KN1

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
112	SHELA WATI	JK2	UA2	PD1	PR3	MK2	SK1	SP2	JT1	ST4	TP2	KL1	TN1	KN2
113	ERDI UNANDAR	JK1	UA3	PD1	PR4	MK3	SK3	SP1	JT2	ST4	TP3	KL1	TN1	KN1
114	ERYZA PERTIWI	JK2	UA3	PD2	PR3	MK2	SK1	SP1	JT2	ST1	TP3	KL1	TN1	KN1
115	ADE IRAWAN	JK1	UA2	PD2	PR1	MK2	SK2	SP1	JT2	ST3	TP2	KL1	TN1	KN1
116	ASHRI MUJAHID IDEAL	JK1	UA3	PD1	PR1	MK2	SK2	SP1	JT2	ST2	TP2	KL1	TN1	KN1
117	PUTRI NOVITRIYANI	JK2	UA1	PD3	PR4	MK1	SK3	SP2	JT1	ST4	TP3	KL1	TN1	KN1
118	MARA GURU	JK1	UA1	PD2	PR3	MK3	SK1	SP1	JT1	ST4	TP4	KL1	TN1	KN1
119	ERIK NURI EFENDI	JK1	UA3	PD3	PR4	MK2	SK3	SP1	JT2	ST4	TP2	KL1	TN1	KN1
120	AGIS SAEPUL MUHTAR	JK1	UA3	PD1	PR1	MK2	SK2	SP1	JT2	ST3	TP2	KL1	TN1	KN2

Sumber: Hasil Penelitian (2020)

4.2.5 Klasifikasi Jumlah Nilai Atribut

Berdasarkan tabel 15. Dibuat klasifikasi jumlah customer pengajuan kredit rumah subsidi yang lolos dan tidak lolos berdasarkan atribut diatas. Klasifikasi berdasarkan jenis kelamin.

Tabel IV. 16

Klasifikasi jumlah nilai atribut berdasarkan jenis kelamin.

Jenis Kelamin	Kolektabilitas	
	Tidak Lolos	Lolos
Pria	27	54
Wanita	14	25

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Usia:

Tabel IV. 17

Klasifikasi jumlah nilai atribut berdasarkan Usia.

Pendidikan	Kolektabilitas	
	Tidak Lolos	Lolos
< 25 Tahun	9	18
25 – 34 Tahun	58	13
35 – 45 Tahun	16	16
>45 Tahun	3	0

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan pendidikan:

Tabel IV. 18

Klasifikasi jumlah nilai atribut berdasarkan Pendidikan.

Pendidikan	Kolektabilitas	
	Tidak Lolos	Lolos
SMA/SLTA	20	39
Diploma	1	29
S1	8	11

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Profesi:

Tabel IV. 19

Klasifikasi jumlah nilai atribut berdasarkan profesi.

Profesi	Kolektabilitas	
	Tidak Lolos	Lolos
PNS / Instansi / Departemen / Pemda	6	14
Swasta Asing / Pma	4	7
Swasta Besar / Menengah	16	27
Wiraswasta Besar / Menengah	15	31

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Total Masa Kerja:

Tabel IV. 20

Klasifikasi jumlah nilai atribut berdasarkan Total Masa Kerja.

Total Masa Kerja	Kolektabilitas	
	Tidak Lolos	Lolos
< 1 Tahun	3	5
1 - 2 Tahun	29	52
> 2 Tahun	9	22

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Status Kerja:

Tabel IV. 21

Klasifikasi jumlah nilai atribut berdasarkan Status Kerja.

Status Kerja	Kolektabilitas	
	Tidak Lolos	Lolos
Kontrak	13	28
Tetap	13	20
Pengusaha	15	31

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Status Pernikahan:

Tabel IV. 22

Klasifikasi jumlah nilai atribut berdasarkan Status Pernikahan.

Status Pernikahan	Kolektabilitas	
	Tidak Lolos	Lolos
Menikah	22	35
Belum Menikah	19	44

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Jumlah Tanggungan:

Tabel IV. 23

Klasifikasi jumlah nilai atribut berdasarkan Jumlah Tanggungan.

Jumlah Tanggungan	Kolektabilitas	
	Tidak Lolos	Lolos
Tidak Ada	26	47
1 - 2 Orang	13	24
> 2 Orang	2	8

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Status Tempat Tinggal:

Tabel IV. 24

Klasifikasi jumlah nilai atribut berdasarkan Status Tempat Tinggal.

Status Tempat Tinggal	Kolektabilitas	
	Tidak Lolos	Lolos
Dinas	1	23
Milik Sendiri	4	4
Keluarga	3	11
Sewa / Kontrak	33	41

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Total Pendapatan:

Tabel IV. 25

Klasifikasi jumlah nilai atribut berdasarkan Total Pendapatan.

Total Pendapatan	Kolektabilitas	
	Tidak Lolos	Lolos
< 2.000.000	9	20
2.000.000 - 3.500.000	16	24
3.600.000 - 4.000.000	15	26
> 4.000.000	1	9

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Kredit Lainnya:

Tabel IV. 26

Klasifikasi jumlah nilai atribut berdasarkan Kredit Lainnya.

Kredit Lainnya	Kolektabilitas	
	Tidak Lolos	Lolos
Tidak Ada	17	76
Kredit Kendaraan	23	1
Kredit Barang	1	2

Sumber: Hasil Penelitian (2020)

Klasifikasi nilai atribut berdasarkan Tunggakan:

Tabel IV. 27

Klasifikasi jumlah nilai atribut berdasarkan Tunggakan.

Tunggakan	Kolektabilitas	
	Tidak Lolos	Lolos
Tidak Ada	17	76
Menunggak	5	1
Tidak Menunggak	19	2

Sumber: Hasil Penelitian (2020)

4.2.6 Perhitungan Manual Decision Tree Algoritma C4.5

Berdasarkan data transaksi pengajuan kredit rumah subsidi yang telah dikumpulkan, selanjutnya dilakukan proses seleksi terhadap data tersebut, lalu cleaning dan transformasi data sehingga diperoleh atribut

dan subset atribut dan subset atribut yang akan digunakan dalam klasifikasi persetujuan kredit rumah subsidi, Berikut adalah langkah – langkah untuk penyelesaian Algoritma C4.5:

- a. Mencari nilai *Entropy* dari Kriteria / Nilai dari setiap atribut.
- b. Mencari nilai *Information Gain* dari setiap atribut.
- c. Pembentukan atribut sebagai akar berdasarkan *Information Gain* tertinggi.
- d. Pembentukan cabang berdasarkan masing masing nilai.
- e. Ulangi proses untuk masing masing cabang.

1. Mencari Nilai Entropy dan Information Gain Node Akar 1

Setelah data latih sudah bersih dan nilai dari atribut juga sudah dibersihkan sudah lengkap tidak ada yang kosong maupun yang duplikat, setelah itu mencari nilai *Entropy* dari setiap nilai atribut dan mencari nilai *Information Gain* dari setiap atribut, Berdasarkan data pada tabel berikut dengan nilai atribut yang sudah di klasifikasikan, berikut adalah data dari pengajuan kredit perumahan Cahaya Darussalam 2 Bekasi yang sudah melewati tahap pre processing dan pembersihan data serta pemelopokan data:

Tabel IV. 28

Seluruh atribut dan nilai atribut untuk mencari Information Gain dan entropy

Node			Jml Kasus (S)	Tidak Lolos (S1)	Lolos (S2)	Entropy	Information Gain
1	TOTAL		120	41	79		
	JENIS KELAMIN						
		PRIA	81	27	54		
		WANITA	39	14	25		
	USIA						
		< 25 TAHUN	27	9	18		

	25 - 34 TAHUN	58	13	45		
	35 - 45 TAHUN	32	16	16		
	> 45 TAHUN	3	3	0		
PENDIDIKAN						
	SMA / SLTA	59	20	39		
	DIPLOMA	42	13	29		
	S1	19	8	11		
PROFESI						
	PNS / Instansi / Departemen / Pemda	20	6	14		
	SWASTA ASING / PMA	11	4	7		
	SWASTA BESAR / MENENGAH	43	16	27		
	WIRASWASTA BESAR / MENENGAH	46	15	31		
TOTAL MASA KERJA						
	< 1 TAHUN	8	3	5		
	1 - 2 TAHUN	81	29	52		
	> 2 TAHUN	31	9	22		
STATUS KERJA						
	KONTRAK	41	13	28		
	TETAP	33	13	20		
	PENGUSAHA	46	15	31		
STATUS PERNIKAHAN						
	MENIKAH	57	22	35		
	BELUM MENIKAH	63	19	44		
JUMLAH TANGGUNGAN						
	TIDAK ADA	73	26	47		
	1 - 2 ORANG	37	13	24		
	> 2 ORANG	10	2	8		
STATUS TEMPAT TINGGAL						
	DINAS	24	1	23		
	MILIK SENDIRI	8	4	4		
	KELUARGA	14	3	11		
	SEWA / KONTRAK	74	33	41		
TOTAL PENGHASILAN						
	< 2.000.000	29	9	20		
	2.000.000 - 3.500.000	40	16	24		
	3.600.000 - 4.000.000	41	15	26		
	> 4.000.000	10	1	9		
KREDIT LAINNYA						
	TIDAK ADA	93	17	76		
	KREDIT KENDARAAN	24	23	1		

Sumber: Hasil Penelitian (2020)

Rumus yang digunakan untuk mencari nilai Information Gain

dan entropy:

Information Gain(S, A)

$$= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Sumber: Kusrini (2009: 16)

Dimana:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut a

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Untuk menghitung nilai entropi menggunakan persamaan 2 berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

Sumber: Kusrini (2009)

Dimana:

S : himpunan kasus

A : fitur

n : jumlah partisi S

P_i : proporsi dari S_i terhadap S

Perhitungan *Entropy* Mencari node akar 1:

Hitung *Entropy* total:

$$\left(\left(\frac{41}{120} \right) x \log 2 \left(\frac{41}{120} \right) \right) + \left(\left(\frac{79}{120} \right) x \log 2 \left(\frac{79}{120} \right) \right) = 0.926404668$$

Hitung Entropy jenis kelamin pria:

$$\left(\left(\frac{27}{81} \right) x \log 2 \left(\frac{27}{81} \right) \right) + \left(\left(\frac{54}{81} \right) x \log 2 \left(\frac{54}{81} \right) \right) = 0.918295834$$

Hitung Entropy jenis kelamin wanita:

$$\left(\left(\frac{14}{39} \right) x \log 2 \left(\frac{14}{39} \right) \right) + \left(\left(\frac{25}{39} \right) x \log 2 \left(\frac{25}{39} \right) \right) = 0.941828535$$

Hitung Entropy Umur <25 Tahun:

$$\left(\left(\frac{35}{89} \right) x \log 2 \left(\frac{35}{89} \right) \right) + \left(\left(\frac{54}{89} \right) x \log 2 \left(\frac{54}{89} \right) \right) = 0.918295834$$

Hitung Entropy Umur 25 – 34 Tahun:

$$\left(\left(\frac{35}{58} \right) x \log 2 \left(\frac{35}{58} \right) \right) + \left(\left(\frac{54}{89} \right) x \log 2 \left(\frac{54}{89} \right) \right) = 0.767651587$$

Hitung Entropy Umur 35 – 45 Tahun:

$$\left(\left(\frac{35}{32} \right) x \log 2 \left(\frac{35}{32} \right) \right) + \left(\left(\frac{54}{89} \right) x \log 2 \left(\frac{54}{89} \right) \right) = 1$$

Hitung Entropy Umur <45Tahun:

$$\left(\left(\frac{35}{3} \right) x \log 2 \left(\frac{35}{3} \right) \right) + \left(\left(\frac{54}{89} \right) x \log 2 \left(\frac{54}{89} \right) \right) = 0$$

Hitung Entropy pendidikan SMA/SLTA:

$$\left(\left(\frac{20}{59} \right) x \log 2 \left(\frac{20}{59} \right) \right) + \left(\left(\frac{39}{59} \right) x \log 2 \left(\frac{39}{59} \right) \right) = 0.923842228$$

Hitung Entropy pendidikan Diploma:

$$\left(\left(\frac{13}{42} \right) x \log 2 \left(\frac{13}{42} \right) \right) + \left(\left(\frac{29}{42} \right) x \log 2 \left(\frac{29}{42} \right) \right) = 0.892623013$$

Hitung Entropy pendidikan S1:

$$\left(\left(\frac{8}{19} \right) x \log 2 \left(\frac{8}{19} \right) \right) + \left(\left(\frac{11}{19} \right) x \log 2 \left(\frac{11}{19} \right) \right) = 0.981940787$$

Hitung Entropy Profesi PNS/Instansi/Departemen/PEMDA:

$$\left(\left(\frac{6}{20} \right) x \log 2 \left(\frac{6}{20} \right) \right) + \left(\left(\frac{14}{20} \right) x \log 2 \left(\frac{14}{20} \right) \right) = 0.881290899$$

Hitung Entropy Profesi Swasta Asing/PMA:

$$\left(\left(\frac{4}{11} \right) x \log 2 \left(\frac{4}{11} \right) \right) + \left(\left(\frac{7}{11} \right) x \log 2 \left(\frac{7}{11} \right) \right) = 0.945660305$$

Hitung Entropy profesi Swasta Besar/Menengah:

$$\left(\left(\frac{16}{43} \right) x \log 2 \left(\frac{16}{43} \right) \right) + \left(\left(\frac{27}{43} \right) x \log 2 \left(\frac{27}{43} \right) \right) = 0.952265625$$

Hitung Entropy profesi Wiraswasta Besar/Menengah:

$$\left(\left(\frac{15}{46} \right) x \log 2 \left(\frac{15}{46} \right) \right) + \left(\left(\frac{31}{46} \right) x \log 2 \left(\frac{31}{46} \right) \right) = 0.910878379$$

Hitung Entropy Total Masa Kerja <1 Tahun:

$$\left(\left(\frac{3}{8} \right) x \log 2 \left(\frac{3}{8} \right) \right) + \left(\left(\frac{5}{8} \right) x \log 2 \left(\frac{5}{8} \right) \right) = 0.954434003$$

Hitung Entropy Total Masa Kerja 1-2 Tahun:

$$\left(\left(\frac{29}{81} \right) x \log 2 \left(\frac{29}{81} \right) \right) + \left(\left(\frac{52}{81} \right) x \log 2 \left(\frac{52}{81} \right) \right) = 0.941031309$$

Hitung Entropy Total Masa Kerja >2Tahun:

$$\left(\left(\frac{9}{31} \right) x \log 2 \left(\frac{9}{31} \right) \right) + \left(\left(\frac{22}{31} \right) x \log 2 \left(\frac{22}{31} \right) \right) = 0.869137581$$

Hitung Entropy Status Kerja Kontrak:

$$\left(\left(\frac{13}{41} \right) x \log 2 \left(\frac{13}{41} \right) \right) + \left(\left(\frac{28}{41} \right) x \log 2 \left(\frac{28}{41} \right) \right) = 0.901170196$$

Hitung Entropy Status Kerja Tetap:

$$\left(\left(\frac{13}{33} \right) x \log 2 \left(\frac{13}{33} \right) \right) + \left(\left(\frac{20}{33} \right) x \log 2 \left(\frac{20}{33} \right) \right) = 0.967294779$$

Hitung Entropy Status Kerja Pengusaha:

$$\left(\left(\frac{15}{46} \right) \times \log 2 \left(\frac{15}{46} \right) \right) + \left(\left(\frac{31}{46} \right) \times \log 2 \left(\frac{31}{46} \right) \right) = 0.910878379$$

Hitung Entropy Status Pernikahan Menikah:

$$\left(\left(\frac{22}{57} \right) \times \log 2 \left(\frac{22}{57} \right) \right) + \left(\left(\frac{35}{57} \right) \times \log 2 \left(\frac{35}{57} \right) \right) = 0.962146133$$

Hitung Entropy Status Pernikahan Belum Menikah:

$$\left(\left(\frac{19}{63} \right) \times \log 2 \left(\frac{19}{63} \right) \right) + \left(\left(\frac{44}{63} \right) \times \log 2 \left(\frac{44}{63} \right) \right) = 0.883222559$$

Hitung Entropy Jumlah Tanggungan Tidak Ada:

$$\left(\left(\frac{26}{73} \right) \times \log 2 \left(\frac{26}{73} \right) \right) + \left(\left(\frac{47}{73} \right) \times \log 2 \left(\frac{47}{73} \right) \right) = 0.939453207$$

Hitung Entropy Jumlah Tanggungan 1-2 Orang:

$$\left(\left(\frac{13}{37} \right) \times \log 2 \left(\frac{13}{37} \right) \right) + \left(\left(\frac{24}{37} \right) \times \log 2 \left(\frac{24}{37} \right) \right) = 0.93526914$$

Hitung Entropy Jumlah Tanggungan >2 Orang:

$$\left(\left(\frac{2}{10} \right) \times \log 2 \left(\frac{2}{10} \right) \right) + \left(\left(\frac{8}{10} \right) \times \log 2 \left(\frac{8}{10} \right) \right) = 0.721928095$$

Hitung Entropy Status tempat tinggal Dinas:

$$\left(\left(\frac{1}{24} \right) \times \log 2 \left(\frac{1}{24} \right) \right) + \left(\left(\frac{23}{24} \right) \times \log 2 \left(\frac{23}{24} \right) \right) = 0.249882293$$

Hitung Entropy Status tempat tinggal Milik Sendiri:

$$\left(\left(\frac{4}{8} \right) \times \log 2 \left(\frac{4}{8} \right) \right) + \left(\left(\frac{4}{8} \right) \times \log 2 \left(\frac{4}{8} \right) \right) = 1$$

Hitung Entropy status tempat tinggal Keluarga:

$$\left(\left(\frac{3}{14} \right) \times \log 2 \left(\frac{3}{14} \right) \right) + \left(\left(\frac{11}{14} \right) \times \log 2 \left(\frac{11}{14} \right) \right) = 0.749595257$$

Hitung Entropy Status tempat tinggal Sewa/Kontrak:

$$\left(\left(\frac{33}{74} \right) x \log 2 \left(\frac{33}{74} \right) \right) + \left(\left(\frac{41}{74} \right) x \log 2 \left(\frac{41}{74} \right) \right) = 0.99155285$$

Hitung Entropy Total pendapatan <Rp.2.000.000:

$$\left(\left(\frac{9}{29} \right) x \log 2 \left(\frac{9}{29} \right) \right) + \left(\left(\frac{20}{29} \right) x \log 2 \left(\frac{20}{29} \right) \right) = 0.893571102$$

Hitung Entropy Total pendapatan Rp.2.000.000-Rp.3.500.000:

$$\left(\left(\frac{16}{40} \right) x \log 2 \left(\frac{16}{40} \right) \right) + \left(\left(\frac{24}{40} \right) x \log 2 \left(\frac{24}{40} \right) \right) = 0.970950594$$

Hitung Entropy Total pendapatan Rp.3.600.000-Rp.4.000.000:

$$\left(\left(\frac{15}{41} \right) x \log 2 \left(\frac{15}{41} \right) \right) + \left(\left(\frac{26}{41} \right) x \log 2 \left(\frac{26}{41} \right) \right) = 0.947435136$$

Hitung Entropy Total pendapatan >Rp.4.000.000:

$$\left(\left(\frac{1}{10} \right) x \log 2 \left(\frac{1}{10} \right) \right) + \left(\left(\frac{9}{10} \right) x \log 2 \left(\frac{9}{10} \right) \right) = 0.468995594$$

Hitung Entropy Kredit Lainnya Tidak Ada:

$$\left(\left(\frac{17}{93} \right) x \log 2 \left(\frac{17}{93} \right) \right) + \left(\left(\frac{76}{93} \right) x \log 2 \left(\frac{76}{93} \right) \right) = 0.686154947$$

Hitung Entropy Kredit lainnya Kredit Kendaraan:

$$\left(\left(\frac{23}{24} \right) x \log 2 \left(\frac{23}{24} \right) \right) + \left(\left(\frac{1}{24} \right) x \log 2 \left(\frac{1}{24} \right) \right) = 0.249882293$$

Hitung Entropy Kredit Lainnya Kredit Barang:

$$\left(\left(\frac{1}{3} \right) x \log 2 \left(\frac{1}{3} \right) \right) + \left(\left(\frac{2}{3} \right) x \log 2 \left(\frac{2}{3} \right) \right) = 0.918295834$$

Hitung Entropy Tunggakan Tidak Ada:

$$\left(\left(\frac{17}{93} \right) x \log 2 \left(\frac{17}{93} \right) \right) + \left(\left(\frac{76}{93} \right) x \log 2 \left(\frac{76}{93} \right) \right) = 0.686154947$$

Hitung Entropy Tunggakan Menunggak:

$$\left(\left(\frac{5}{6} \right) x \log 2 \left(\frac{5}{6} \right) \right) + \left(\left(\frac{1}{6} \right) x \log 2 \left(\frac{1}{6} \right) \right) = 0.650022422$$

Hitung Entropy Tunggakan Tidak Menunggak:

$$\left(\left(\frac{19}{21} \right) x \log 2 \left(\frac{19}{21} \right) \right) + \left(\left(\frac{2}{21} \right) x \log 2 \left(\frac{2}{21} \right) \right) = 0.453716339$$

Perhitungan *Information Gain* dari atribut Untuk Mencari Node

Akar 1:

Hitung Information Gain Jenis Kelamin:

$$(0.901170196) - \left(\left(\frac{81}{120} x 0.918295834 \right) + \left(\frac{39}{120} x 0.941828535 \right) \right) = 0.000460706$$

Hitung Information Gain Usia:

$$(0.901170196) - \left(\left(\frac{27}{120} x 0.918295834 \right) + \left(\frac{58}{120} x 0.767651587 \right) + \left(\frac{32}{120} x 1 \right) + \left(\frac{3}{120} x 0 \right) \right) = 0.$$

082089838

Hitung Information Gain Pendidikan:

$$(0.901170196) - \left(\left(\frac{59}{120} x 0.923842228 \right) + \left(\frac{42}{120} x 0.892623013 \right) + \left(\frac{19}{120} x 0.981940787 \right) \right) = 0.$$

004290227

Hitung Information Gain Profesi:

$$(0.901170196) - \left(\left(\frac{20}{120} x 0.881290899 \right) + \left(\frac{11}{120} x 0.945660305 \right) + \left(\frac{43}{120} x 0.952265625 \right) + \left(\frac{46}{120} x 0.910878379 \right) \right) = 0.002438763$$

Hitung Information Gain Total Masa Kerja:

$$(0.901170196) - \left(\left(\frac{8}{120} x 0.954434003 \right) + \left(\frac{81}{120} x 0.941031309 \right) + \left(\frac{31}{120} x 0.869137581 \right) \right) = 0.$$

003052393

Hitung Information Gain Status Kerja:

$$(0.901170196) - \left(\left(\frac{41}{120} x 0.901170196 \right) + \left(\frac{33}{120} x 0.967294779 \right) + \left(\frac{46}{120} x 0.910878379 \right) \right) = 0.$$

003328742

Hitung Information Gain Status Pernikahan:

$$(0.901170196) - \left(\left(\frac{57}{120} x 0.962146133 \right) + \left(\frac{63}{120} x 0.883222559 \right) \right) = 0.005693411$$

Hitung Information Gain Jumlah Tanggungan:

$$(0.901170196) - \left(\left(\frac{73}{120}x 0.939453207 \right) + \left(\frac{37}{120}x 0.93526914 \right) + \left(\frac{10}{120}x 0.721928095 \right) \right) = 0.$$

006368641

Hitung Information Gain Status Tempat Tinggal:

$$(0.901170196) - \left(\left(\frac{24}{120}x 0.249882293 \right) + \left(\frac{8}{120}x 1 \right) + \left(\frac{14}{120}x 0.749595257 \right) + \left(\frac{74}{120}x 0.99155285 \right) \right) = 0.110851172$$

Hitung Information Gain Total Pendapatan:

$$(0.901170196) - \left(\left(\frac{29}{120}x 0.893571102 \right) + \left(\frac{40}{120}x 0.970950594 \right) + \left(\frac{41}{120}x 0.947435136 \right) + \left(\frac{10}{120}x 0.468995594 \right) \right) = 0.024018149$$

Hitung Information Gain Kredit Lainnya:

$$(0.901170196) - \left(\left(\frac{93}{120}x 0.686154947 \right) + \left(\frac{24}{120}x 0.249882293 \right) + \left(\frac{3}{120}x 0.918295834 \right) \right) = 0.$$

321700729

Hitung Information Gain Tunggakan:

$$(0.901170196) - \left(\left(\frac{93}{120}x 0.686154947 \right) + \left(\frac{6}{120}x 0.650022422 \right) + \left(\frac{21}{120}x 0.453716339 \right) \right) = 0.$$

282733103

Berikut adalah tabel hasil perhitungan *Information Gain* dan *Entropy* dari atribut dan masing masing nilai dari atribut ditunjukkan pada Tabel 28.

Tabel IV. 29

Hasil perhitungan Information Gain dan Entropy node akar 1.

Node			Jml Kasus (S)	Tidak Lolos (S1)	Lolos (S2)	Entropy	Information Gain
1	TOTAL		120	41	79	0.926404668	
	JENIS KELAMIN						0.000460706
	PRIA		81	27	54	0.918295834	
	WANITA		39	14	25	0.941828535	
	USIA						0.082089838
	< 25 TAHUN		27	9	18	0.918295834	
	25 - 34 TAHUN		58	13	45	0.767651587	
	35 - 45 TAHUN		32	16	16	1	
	> 45 TAHUN		3	3	0	0	
	PENDIDIKAN						0.004290227
	SMA / SLTA		59	20	39	0.923842228	

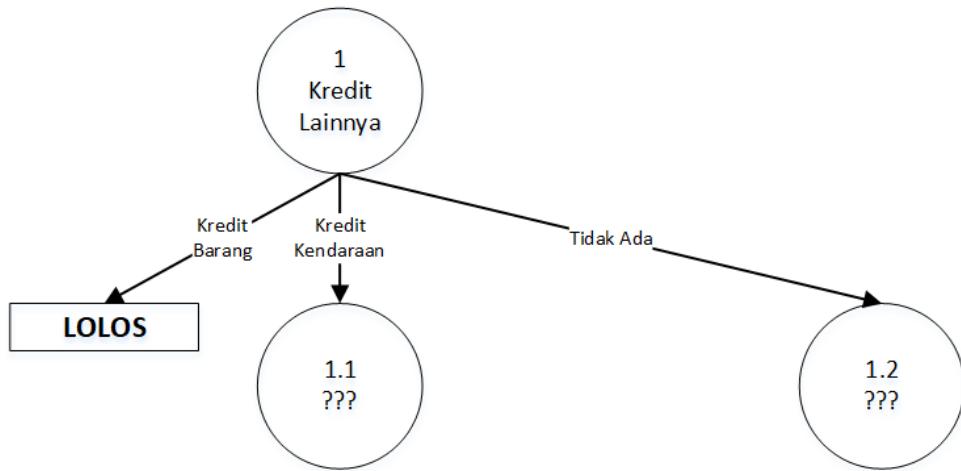
	DIPLOMA	42	13	29	0.892623013	
	S1	19	8	11	0.981940787	
PROFESI						0.002438763
	PNS / Instansi / Departemen / Pemda	20	6	14	0.881290899	
	SWASTA ASING / PMA	11	4	7	0.945660305	
	SWASTA BESAR / MENENGAH	43	16	27	0.952265625	
	WIRASWASTA BESAR / MENENGAH	46	15	31	0.910878379	
TOTAL MASA KERJA						0.003052393
	< 1 TAHUN	8	3	5	0.954434003	
	1 - 2 TAHUN	81	29	52	0.941031309	
	> 2 TAHUN	31	9	22	0.869137581	
STATUS KERJA						0.003328742
	KONTRAK	41	13	28	0.901170196	
	TETAP	33	13	20	0.967294779	
	PENGUSAHA	46	15	31	0.910878379	
STATUS PERNIKAHAN						0.005693411
	MENIKAH	57	22	35	0.962146133	
	BELUM MENIKAH	63	19	44	0.883222559	
JUMLAH TANGGUNGAN						0.006368641
	TIDAK ADA	73	26	47	0.939453207	
	1 - 2 ORANG	37	13	24	0.93526914	
	> 2 ORANG	10	2	8	0.721928095	
STATUS TEMPAT TINGGAL						0.110851172
	DINAS	24	1	23	0.249882293	
	MILIK SENDIRI	8	4	4	1	
	KELUARGA	14	3	11	0.749595257	
	SEWA / KONTRAK	74	33	41	0.99155285	
TOTAL PENGHASILAN						0.024018149
	< 2.000.000	29	9	20	0.893571102	
	2.000.000 - 3.500.000	40	16	24	0.970950594	
	3.600.000 - 4.000.000	41	15	26	0.947435136	
	> 4.000.000	10	1	9	0.468995594	
KREDIT LAINNYA						0.321700729
	TIDAK ADA	93	17	76	0.686154947	
	KREDIT KENDARAAN	24	23	1	0.249882293	
	KREDIT BARANG	3	1	2	0.918295834	
TUNGGAKAN						0.282733103
	TIDAK ADA	93	17	76	0.686154947	
	MENUNGGAK	6	5	1	0.650022422	
	TIDAK MENUNGGAK	21	19	2	0.453716339	

Sumber: Hasil Penelitian (2020)

Dari tabel diatas menunjukan bahwa atribut **KREDIT LAINNYA** memiliki nilai *Information Gain* yang paling tinggi yaitu **0.321700729** dengan demikian atribut **KREDIT LAINNYA** dapat menjadi akar node 1.

2. Pembentukan pohon keputusan (Decision tree) akar node 1

Setelah mendapatkan node akar selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai *Entropy* dari masing masing nilai atribut. Jika memenuhi ketentuan berenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari **KREDIT LAINNYA** yaitu **TIDAK ADA**, **KREDIT KENDARAAN**, dan **KREDIT BARANG**, karena kita memakai ketentuan **MINIMAL LEAF SIZE** 2,maka cari nilai atribut yang masih memungkinkan untuk dibuatkan cabang lagi, disini nilai atribut dari atribut **KREDIT LAINNYA** adalah **TIDAK ADA** dan **KREDIT KENDARAAN** yang akan dibuatkan cabang lagi karena belum memenuhi kondisi berhenti pembuatan pohon keputusan, dan untuk menentukan keputusan untuk setiap nilai atribut, Suatu variable tidak memiliki nilai (kosong).Jika kondisi ini terjadi. Buatlah sebuah *leaf* yang berisi kelas yang paling banyak muncul serta cek kondisi *leaf size* yang ditentukan, disini atribut nilai **KREDIT BARANG** menghasilkan keputusan **LOLOS**, karena jumlah kemungkinan lolos lebih besar dari pada yang tidak lolos. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1 akar ditunjukan pada Gambar IV.3.



Sumber: Penelitian (2020)

Gambar IV. 3

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node akar 1.

3. Mencari Nilai Entropy dan Information Gain node 1.1

Dari hasil perhitungan pada tabel dan pohon diatas belum terlihat keputusan yang dominan dari setiap hasil keterangan. Maka kita harus mencari kembali nilai *Entropy* dan *Information Gain* dari setiap nilai atribut (kriteria) **KREDIT LAINNYA = KREDIT KENDARAAN** dan atribut **KREDIT LAINNYA** dihilangkan karena sudah terpilih. Berikut adalah tabel perhitungan mencari nilai *Entropy* dan *Information Gain* dari atribut **KREDIT LAINNYA** dengan nilai atribut **KREDIT KENDARAAN** ditunjukan pada tabel IV.29.

Tabel IV. 30
Hasil perhitungan Information Gain dan Entropy node 1.1.

Node			Jml Kasus (S)	Tidak Lulus (S1)	Lulus (S2)	Entropy	Information Gain
1.1	KREDIT LAINNYA - KREDIT KENDARAAN		24	23	1	0.249882293	
	JENIS KELAMIN						0.068694145
	PRIA	16	16	0		0	
	WANITA	8	7	1	0.543564443		
	USIA						0.068694145
	< 25 TAHUN	5	5	0		0	
	25 - 34 TAHUN	8	7	1	0.543564443		
	35 - 45 TAHUN	10	10	0		0	
	> 45 TAHUN	1	1	0		0	
	PENDIDIKAN						0.077311066
	SMA / SLTA	9	9	0		0	
	DIPLOMA	7	6	1	0.591672779		
	S1	8	8	0			
	PROFESI						0.068694145
	PNS / Instansi / Departemen / Pemda	4	4	0		0	
	SWASTA ASING / PMA	2	2	0		0	
	SWASTA BESAR / MENENGAH	8	7	1	0.543564443		
	WIRASWASTA BESAR / MENENGAH	10	10	0		0	
	TOTAL MASA KERJA						0.017724721
	< 1 TAHUN	0	0	0		0	
	1 - 2 TAHUN	18	17	1	0.309543429		
	> 2 TAHUN	6	6	0		0	
	STATUS KERJA						0.087376687
	KONTRAK	6	5	1	0.650022422		
	TETAP	8	8	0		0	
	PENGUSAHA	10	10	0		0	
	STATUS PERNIKAHAN						0.033330102
	MENIKAH	14	13	1	0.371232327		
	BELUM MENIKAH	10	10	0		0	
	JUMLAH TANGGUNGAN						0.054467462
	TIDAK ADA	13	13	0		0	
	1 - 2 ORANG	10	9	1	0.468995594		
	> 2 ORANG	1	1	0		0	
	STATUS TEMPAT TINGGAL						0.249882293
	DINAS	1	0	1		0	
	MILIK SENDIRI	2	2	0		0	
	KELUARGA	1	1	0		0	
	SEWA / KONTRAK	20	20	0		0	
	TOTAL PENGHASILAN						0.077311066
	< 2.000.000	8	8	0		0	
	2.000.000 - 3.500.000	7	6	1	0.591672779		

		3.600.000 - 4.000.000	9	9	0	0	
		> 4.000.000	0	0	0	0	
TUNGGAKAN							0.087376687
		TIDAK ADA	0	0	0	0	
		MENUNGGAK	6	5	1	0.650022422	
		TIDAK MENUNGGAK	18	18	0	0	

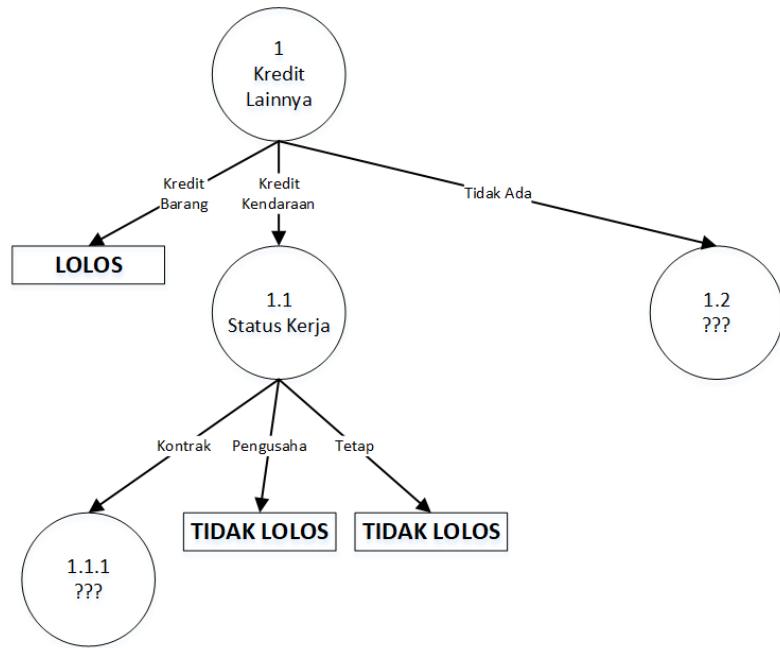
Sumber: Hasil Penelitian (2020)

Dari tabel diatas menunjukan bahwa atribut **STATUS KERJA** memiliki nilai *Information Gain* yang paling tinggi yaitu **0.087376687** dengan demikian atribut **STATUS KERJA** dapat menjadi *Internal Node/ cabang node* 1 dari nilai atribut **KREDIT LAINNYA = KREDIT KENDARAAN.**

4. Pembetukan Pohon Keputusan (Decision tree) Node 1.1.

Setelah mendapatkan node akar, cabang selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai *Entropy*, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari **STATUS KERJA** yaitu **KONTRAK**, **TETAP**, dan **PENGUSAHA**, dan untuk menentukan keputusan untuk setiap nilai atribut, apabila Data latih kosong (tidak memiliki isi). Pada kondisi ini pembentukan tree dihentikan, Jika kondisi ini terjadi. Buatlah sebuah *leaf* yang berisi kelas yang paling banyak muncul, disini nilai atribut **KONTRAK** yang akan dibuatkan cabang lagi karena kondisi nya belum memenuhi syarat berhenti, dan untuk menentukan keputusan untuk setiap nilai atribut, cari kelas pendukung terbesar, disini atribut nilai **TETAP** dan **PENGUSAHA** menghasilkan keputusan **TIDAK LOLOS** karena semua data latih kosong, maka dari itu pembuatan decision tree dihentikan. Dan berikut ini

adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1 akar ditunjukan pada Gambar IV.4.



Sumber: Hasil Penelitian (2020)

Gambar IV. 4

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.1.

5. Mencari Nilai Entropy dan Information Gain Node 1.1.1

Karena belum semua nilai atribut dari atribut **STATUS KERJA** memenuhi kriteria berhenti pembuatan pohon keputusan, maka kita harus mencari *internal node/cabang* selanjutnya dengan cara mencari nilai *entropy* dan *information gain*. Berikut adalah hasil perhitungan untuk mencari *internal node/cabang* 1.1 ditunjukan pada Tabel IV.30:

Tabel IV. 31
Hasil perhitungan Information Gain dan Entropy node 1.1.1.

Node			Jml Kasus (S)	Tidak Lulus (S1)	Lulus (S2)	Entropy	Information Gain
1.1.1	STATUS KERJA - KONTRAK		6	5	1	0.650022422	
	JENIS KELAMIN						0.650022422
		PRIA	5	5	0	0	
		WANITA	1	1	0	0	
	USIA						0.316689088
		< 25 TAHUN	2	2	0	0	
		25 - 34 TAHUN	2	1	1	1	
		35 - 45 TAHUN	2	2	0	0	
		> 45 TAHUN	0	0	0	0	
	PENDIDIKAN						0.190874505
		SMA / SLTA	0	0	0	0	
		DIPLOMA	3	2	1	0.918295834	
		S1	3	3	0	0	
	PROFESI						0.048415676
		PNS / Instansi / Departemen / Pemda	0	0	0	0	
		SWASTA ASING / PMA	1	1	0	0	
		SWASTA BESAR / MENENGAH	5	4	1	0.721928095	
		WIRASWASTA BESAR / MENENGAH	0	0	0	0	
	TOTAL MASA KERJA						0.048415676
		< 1 TAHUN	0	0	0	0	
		1 - 2 TAHUN	5	4	1	0.721928095	
		> 2 TAHUN	1	1	0	0	
	STATUS PERNIKAHAN						0.190874505
		MENIKAH	3	2	1	0.918295834	
		BELUM MENIKAH	3	3	0	0	
	JUMLAH TANGGUNGAN						0.316689088
		TIDAK ADA	4	4	0	0	
		1 - 2 ORANG	2	1	1	1	
		> 2 ORANG	0	0	0	0	
	TOTAL PENGHASILAN						0.316689088
		< 2.000.000	1	1	0	0	
		2.000.000 - 3.500.000	2	1	1	1	
		3.600.000 - 4.000.000	3	3	0	0	
		> 4.000.000	0	0	0	0	
	TUNGGAKAN						0.316689088
		TIDAK ADA	0	0	0	0	
		MENUNGGAK	2	1	1	1	
		TIDAK MENUNGGAK	4	4	0	0	

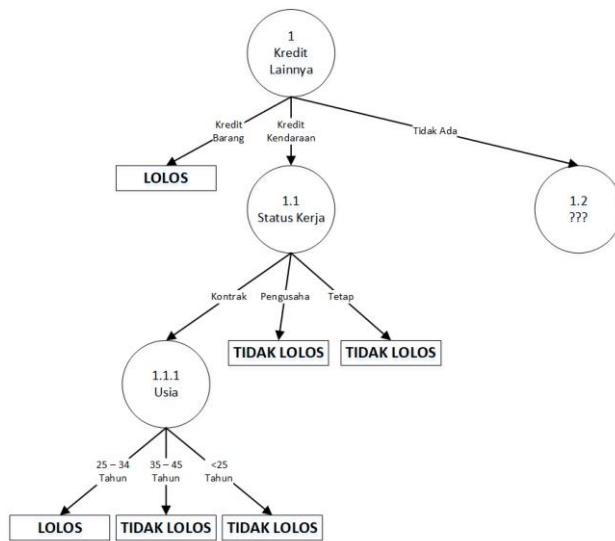
Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukan bahwa atribut **USIA** memiliki nilai *Information Gain* yang paling tinggi yaitu **0.316689088** dengan demikian atribut **USIA** dapat menjadi *Internal Node/ cabang node* 1.1 dari nilai atribut **STATUS KERJA = KONTRAK**. Atribut jenis kelamin tidak terpilih walaupun nilai gainnya paling besar karena semakin nilai information gain mendekati nilai entropy dari internal node / cabang maka tingkat pengaruhnya semakin kecil dan sebaliknya, maka dari itu cari nilai information gain tertinggi lainnya.

6. Pembentukan Pohon Keputusan Node 1.1.1.

Setelah mendapatkan node 1.1, cabang selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai *Entropy*, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada empat nilai atribut dari **USIA** yaitu **<25 TAHUN, 25 – 34 TAHUN, 35 – 45 TAHUN, >45 TAHUN**, dan untuk menentukan keputusan untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini pembentukan tree dihentikan. dan Suatu variable tidak memiliki nilai (kosong). Jika kondisi ini terjadi. Buatlah sebuah leaf yang berisi kelas yang paling banyak muncul berdasarkan pengetahuan pakar atau memilih pada hasil acak. Disini nilai atribut **<25 TAHUN** bernilai **TIDAK LOLOS** karena memiliki data latih kosong dan semua kelas berada di **TIDAK LOLOS, 25 – 34 TAHUN** bernilai **LOLOS** karena jumlah setiap kelasnya sama maka menggunakan pengetahuan pakar atau memilih secara acak, **35 – 45 TAHUN** bernilai **TIDAK LOLOS** karena memiliki data latih kosong dan semua kelas berada di **TIDAK LOLOS, >45 TAHUN** tidak mempunyai

nilai karena tidak memiliki sampel. Karena semua nilai sudah memenuhi syarat berhenti, maka dari itu pembuatan decision tree dihentikan. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1 akar ditunjukan pada Gambar IV.5.



Sumber: Hasil Penelitian (2020).

Gambar IV. 5

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.1.1.

7. Mencari Nilai Entropy dan Information Gain Node 1.2.

Setelah node 1.1.1 menemukan kondisi akhir berhenti membuat pohon keputusan, maka kita akan melanjutkan untuk mencari internal node/cabang dari node 1, yaitu atribut kredit lainnya dengan nilai atribut tidak ada dengan cara mencari nilai atribut dan information gain dari setiap atribut dan nilai atribut lainnya, Berikut hasil perhitungan nilai entropy dari setiap nilai atribut dan nilai information gain berdasarkan atribut. Berikut adalah hasil perhitungan untuk mencari internal node/cabang 1.2 ditunjukan pada Tabel IV.31:

Tabel IV. 32
Hasil perhitungan Information Gain dan Entropy node 1.2.

Node		Jml Kasus (S)	Tidak Lelos (S1)	Lelos (S2)	Entropy	Information Gain
1.2	KREDIT LAINNYA - TIDAK ADA	93	17	76	0.686154947	
	JENIS KELAMIN					0.000672529
	PRIA	63	11	52	0.668127334	
	WANITA	30	6	24	0.721928095	
	USIA					0.072026536
	< 25 TAHUN	21	3	18	0.591672779	
	25 - 34 TAHUN	48	6	42	0.543564443	
	35 - 45 TAHUN	22	6	16	0.845350937	
	> 45 TAHUN	2	2	0	0	
	PENDIDIKAN					0.039566759
	SMA / SLTA	48	11	37	0.776555785	
	DIPLOMA	34	6	28	0.672294817	
	S1	11	0	11	0	
	PROFESI					0.012200695
	PNS / Instansi / Departemen / Pemda	15	2	13	0.566509507	
	SWASTA ASING / PMA	8	2	6	0.811278124	
	SWASTA BESAR / MENENGAH	34	8	26	0.787126586	
	WIRASWASTA BESAR / MENENGAH	36	5	31	0.581321499	
	TOTAL MASA KERJA					0.018322698
	< 1 TAHUN	8	3	5	0.954434003	
	1 - 2 TAHUN	60	11	49	0.687315093	
	> 2 TAHUN	25	3	22	0.529360865	
	STATUS KERJA					0.00613884
	KONTRAK	34	7	27	0.733537929	
	TETAP	23	5	18	0.755375413	
	PENGUSAHA	36	5	31	0.581321499	
	STATUS PERNIKAHAN					0.003925821
	MENIKAH	42	9	33	0.749595257	
	BELUM MENIKAH	51	8	43	0.626751137	
	JUMLAH TANGGUNGAN					0.005587638
	TIDAK ADA	58	12	46	0.735508582	
	1 - 2 ORANG	26	4	22	0.619382195	
	> 2 ORANG	9	1	8	0.503258335	
	STATUS TEMPAT TINGGAL					0.046476955
	DINAS	23	1	22	0.258018669	
	MILIK SENDIRI	5	2	3	0.970950594	
	KELUARGA	12	2	10	0.650022422	
	SEWA / KONTRAK	53	12	41	0.77170947	
	TOTAL PENGHASILAN					0.042923217
	< 2.000.000	20	1	19	0.286396957	

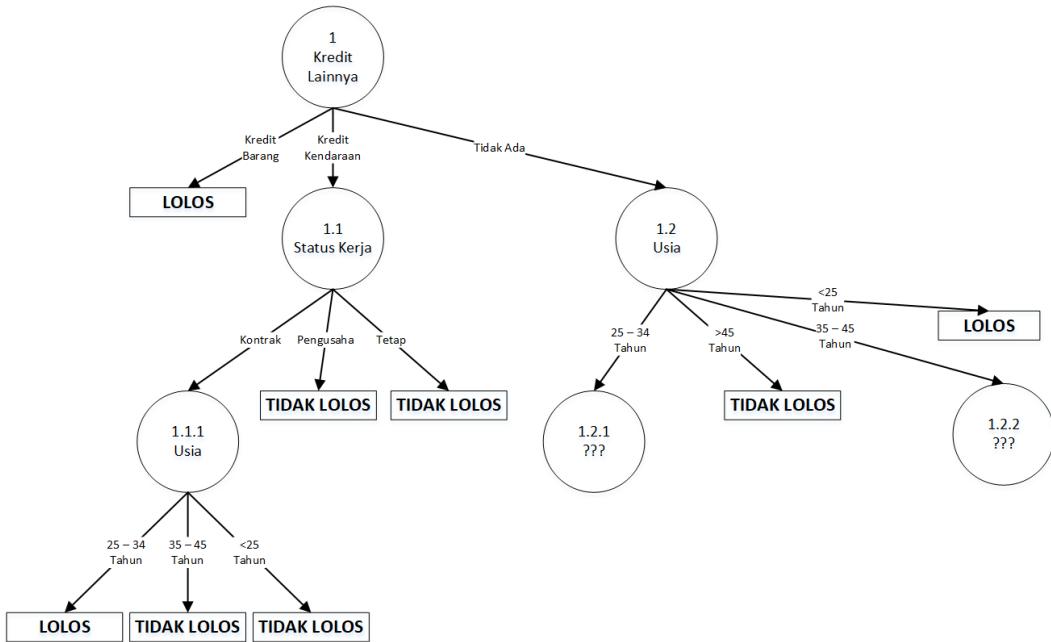
		2.000.000 3.500.000	-	32	9	23	0.857148437	
		3.600.000 4.000.000	-	31	6	25	0.708835673	
		> 4.000.000		10	1	9	0.468995594	
TUNGGAKAN								0
	TIDAK ADA		93	17	76		0.686154947	
	MENUNGGAK		0	0	0		0	
	TIDAK MENUNGGAK		0	0	0		0	

Sumber: Hasil Penelitian (2020)

Dari tabel diatas menunjukan bahwa atribut **USIA** memiliki nilai *Information Gain* yang paling tinggi yaitu **0.072026536** dengan demikian atribut **USIA** dapat menjadi *Internal Node/cabang node 1* dari nilai atribut **KREDIT LAINNYA = TIDAK ADA.**

8. Pembentukan Pohon Keputusan Node 1.2.

Setelah mendapatkan internal node/cabang 1.2, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai *Entropy*, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada empat nilai atribut dari **USIA** yaitu **<25 TAHUN, 25 – 34 TAHUN, 35 – 45 TAHUN, >45 TAHUN,** dan untuk menentukan keputusan untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan. dan Suatu variable tidak memiliki nilai (kosong).Jika kondisi ini terjadi. Buatlah sebuah leaf yang berisi kelas yang paling banyak muncul berdasarkan pengetahuan pakar atau memilih pada hasil acak. Disini nilai atribut **<25 TAHUN** bernilai **LOLOS** karena jumlah kelasnya lebih banyak di **LOLOS, 25 – 34** dan **35 – 45 TAHUN** menjadi cabang karena belum memenuhi syarat untuk berhenti. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2 akar ditunjukan pada Gambar IV.6.



Sumber: Hasil Penelitian (2020).

Gambar IV. 6

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.

9. Mencari Nilai Entropy dan Information Gain Node 1.2.1.

Setelah menemukan internal node / cabang dari node 1.2 dan ada 2 nilai atribut dari atribut usia yaitu USIA = 25 – 35 TAHUN yang belum memenuhi kriteria maka kita harus mencari kembali internal node / cabang 1.2.1 dengan cara mencari nilai entropy dari nilai atribut dan nilai information gain dari setiap atribut, berikut adalah tabel hasil perhitungan node 1.2.1 ditunjukkan pada tabel IV.32:

Tabel IV. 33
Hasil perhitungan Information Gain dan Entropy node 1.2.1.

Node			Jml Kasus (S)	Tidak Lolos (S1)	Lolos (S2)	Entropy	Information Gain
1.2.1	USIA - 25-34 TAHUN		48	6	42	0.543564443	
	JENIS KELAMIN						0.009225439
		PRIA	31	3	28	0.458685816	
		WANITA	17	3	14	0.672294817	
	PENDIDIKAN						0.106300543
		SMA / SLTA	28	6	22	0.749595257	
		DIPLOMA	14	0	14	0	
		S1	6	0	6	0	
	PROFESI						0.176359902
		PNS / Instansi / Departemen / Pemda	6	0	6	0	
		SWASTA ASING / PMA	5	0	5	0	
		SWASTA BESAR / MENENGAH	20	6	14	0.881290899	
		WIRASWASTA BESAR / MENENGAH	17	0	17	0	
	TOTAL MASA KERJA						0.090711461
		< 1 TAHUN	3	2	1	0.918295834	
		1 - 2 TAHUN	31	2	29	0.345117315	
		> 2 TAHUN	14	2	12	0.591672779	
	STATUS KERJA						0.104813832
		KONTRAK	20	5	15	0.811278124	
		TETAP	11	1	10	0.439496987	
		PENGUSAHA	17	0	17	0	
	STATUS PERNIKAHAN						0.00856653
		MENIKAH	14	1	13	0.371232327	
		BELUM MENIKAH	34	5	29	0.602430802	
	JUMLAH TANGGUNGAN						0.017775438
		TIDAK ADA	35	5	30	0.591672779	
		1 - 2 ORANG	9	1	8	0.503258335	
		> 2 ORANG	4	0	4	0	
	STATUS TEMPAT TINGGAL						0.073288019
		DINAS	9	0	9	0	
		MILIK SENDIRI	1	0	1	0	
		KELUARGA	5	0	5	0	
		SEWA / KONTRAK	33	6	27	0.684038436	
	TOTAL PENGHASILAN						0.120458289
		< 2.000.000	12	0	12	0	
		2.000.000 - 3.500.000	18	5	13	0.852405179	
		3.600.000 - 4.000.000	12	1	11	0.41381685	
		> 4.000.000	6	0	6	0	
	TUNGGAKAN						0
		TIDAK ADA	48	6	42	0.543564443	
		MENUNGGAK	0	0	0	0	

	TIDAK MENUNGGAK	0	0	0	0	
--	--------------------	---	---	---	---	--

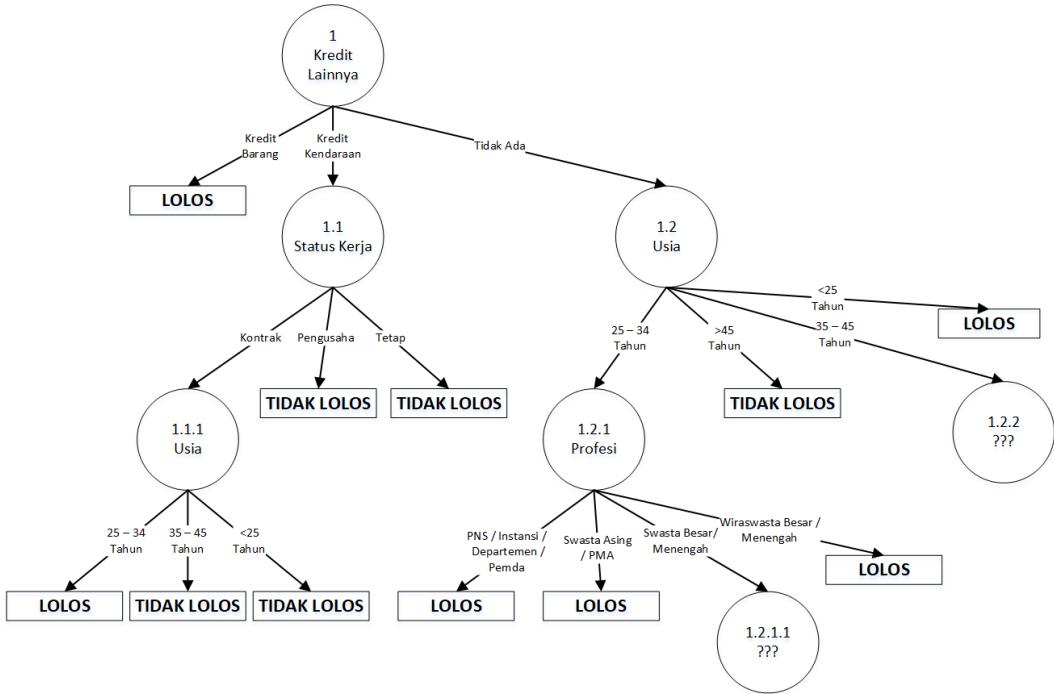
Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukan bahwa atribut PROFESI memiliki nilai Information Gain yang paling tinggi yaitu 0. 176359902 dengan demikian atribut PROFESI dapat menjadi Internal Node/cabang node 1.2 dari nilai atribut USIA = 25 – 34 TAHUN.

10. Pembentukan Pohon Keputusan Node 1.2.1.

Setelah mendapatkan internal node/cabang 1.2.1, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai Entropy, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada empat nilai atribut dari PROFESI yaitu

PNS / Instansi / Departemen / Pemda, SWASTA ASING / PMA, SWASTA BESAR / MENENGAH, WIRASWASTA BESAR / MENENGAH, dan untuk menentukan keputusan hasil untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan. Disini nilai atribut PNS / Instansi / Departemen / Pemda, SWASTA ASING / PMA dan WIRASWASTA BESAR / MENENGAH bernilai LOLOS karena jumlah kelasnya lebih banyak di LOLOS, SWASTA BESAR / MENENGAH menjadi cabang karena belum memenuhi syarat untuk berhenti. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2.1 akar ditunjukan pada Gambar IV.7.



Sumber: Hasil Penelitian (2020).

Gambar IV. 7

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.

11. Mencari Nilai Entropy dan Information Gain Node 1.2.1.1.

Setelah menemukan internal node / cabang dari node 1.2.1 dan ada 1 nilai atribut dari atribut PROFESI = SWASTA BESAR/MENENGAH yang belum memenuhi kriteria maka kita harus mencari kembali internal node / cabang 1.2.1.1 dengan cara mencari nilai entropy dari nilai atribut dan nilai information gain dari setiap atribut, berikut adalah tabel hasil perhitungan node 1.2.1.1 ditunjukkan pada Tabel IV.33:

Tabel IV. 34
Hasil perhitungan Information Gain dan Entropy node 1.2.1.1.

Node			Jml Kasus (S)	Tidak Lelos (S1)	Lelos (S2)	Entropy	Information Gain
1.2.1.1	PROFESI - SWASTA BESAR / MENENGAH		20	6	14	0.881290899	
	JENIS KELAMIN						0.012750423
		PRIA	12	3	9	0.811278124	
		WANITA	8	3	5	0.954434003	
	PENDIDIKAN						0.234068055
		SMA / SLTA	13	6	7	0.995727452	
		DIPLOMA	5	0	5	0	
		S1	2	0	2	0	
	TOTAL MASA KERJA						0.235954824
		< 1 TAHUN	2	2	0	0	
		1 - 2 TAHUN	13	2	11	0.619382195	
		> 2 TAHUN	5	2	3	0.970950594	
	STATUS KERJA						0.000662633
		KONTRAK	17	5	12	0.873981048	
		TETAP	3	1	2	0.918295834	
		PENGUSAHA	0	0	0	0	
	STATUS PERNIKAHAN						0.012087
		MENIKAH	5	1	4	0.721928095	
		BELUM MENIKAH	15	5	10	0.918295834	
	JUMLAH TANGGUNGJAN						0.026715938
		TIDAK ADA	16	5	11	0.896038233	
		1 - 2 ORANG	3	1	2	0.918295834	
		> 2 ORANG	1	0	1	0	
	STATUS TEMPAT TINGGAL						0.153077953
		DINAS	5	0	5	0	
		MILIK SENDIRI	0	0	0	0	
		KELUARGA	0	0	0	0	
		SEWA / KONTRAK	15	6	9	0.970950594	
	TOTAL PENGHASILAN						0.098362919
		< 2.000.000	0	0	0	0	
		2.000.000 - 3.500.000	-	12	5	0.979868757	
		3.600.000 - 4.000.000	-	6	1	0.650022422	
		> 4.000.000	2	0	2	0	
	TUNGGAKAN						0
		TIDAK ADA	20	6	14	0.881290899	
		MENUNGGAK	0	0	0	0	
		TIDAK MENUNGGAK	0	0	0	0	

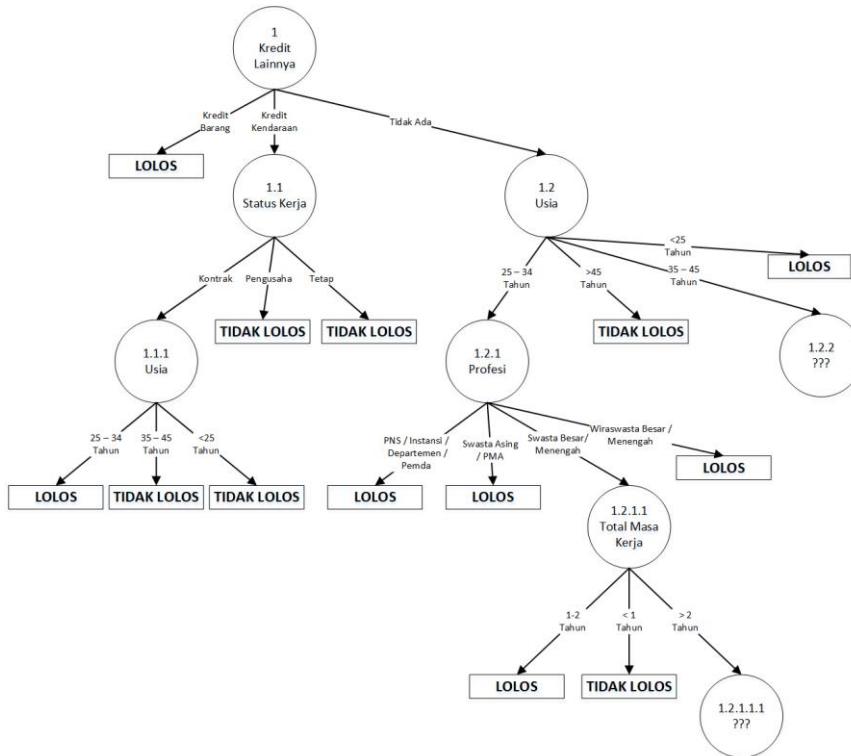
Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukkan bahwa atribut TOTAL MASA KERJA memiliki nilai Information Gain yang paling tinggi yaitu 0.235954824 dengan

demikian atribut TOTAL MASA KERJA dapat menjadi Internal Node/cabang node 1.2.1 dari nilai atribut PROFESI = SWASTA BESAR/MENENGAH.

12. Pembentukan Pohon Keputusan Node 1.2.1.1.

Setelah mendapatkan internal node/cabang 1.2.1.1, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai Entropy, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari TOTAL MASA KERJA yaitu < 1 TAHUN, 1 – 2 TAHUN, > 2 TAHUN, dan untuk menentukan keputusan hasil untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan. Dan Suatu variable tidak memiliki nilai (kosong). Jika kondisi ini terjadi. Buatlah sebuah leaf yang berisi kelas yang paling banyak muncul atau pemilihan berdasarkan pengetahuan pakar atau memilih pada hasil acak. Disini nilai atribut < 1 TAHUN bernilai TIDAK LOLOS karena terdapat nilai kosong pada data latih dan semua kelas berada di LOLOS, 1 – 2 TAHUN bernilai LOLOS karena semua kelas berada di LOLOS, > 2 TAHUN tidak bernilai dan akan dijadikan cabang, karena belum memenuhi syarat untuk berhenti. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2.1.1 akar ditunjukan pada Gambar IV.8:



Sumber: Hasil Penelitian (2020).

Gambar IV. 8

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.1.

13. Mencari Nilai Entropy dan Information Gain Node 1.2.1.1.1.

Setelah menemukan internal node / cabang dari node 1.2.1.1 dan ada 1 nilai atribut dari atribut TOTAL MASA KERJA = < 2 TAHUN yang belum memenuhi kriteria maka kita harus mencari kembali internal node / cabang 1.2.1.1 dengan cara mencari nilai entropy dari nilai atribut dan nilai information gain dari setiap atribut, berikut adalah tabel hasil perhitungan node 1.2.1.1 ditunjukkan pada Tabel IV.34:

Tabel IV. 35
Hasil perhitungan Information Gain dan Entropy node 1.2.1.1.1.

Node			Jml Kasus (S)	Tidak Lelos (S1)	Lelos (S2)	Entropy	Information Gain
1.2.1.1.1	TOTAL MASA KERJA - > 2 TAHUN		5	2	3	0.970950594	
	JENIS KELAMIN						0.419973094
		PRIA	3	2	1	0.918295834	
		WANITA	2	0	2	0	
	PENDIDIKAN						0.419973094
		SMA / SLTA	3	2	1	0.918295834	
		DIPLOMA	1	0	1	0	
		S1	1	0	1	0	
	STATUS KERJA						0.170950594
		KONTRAK	4	2	2	1	
		TETAP	1	0	1	0	
		PENGUSAHA	0	0	0	0	
	STATUS PERNIKAHAN						0.170950594
		MENIKAH	1	0	1	0	
		BELUM MENIKAH	4	2	2	1	
	JUMLAH TANGGUNGAN						0
		TIDAK ADA	5	2	3	0.970950594	
		1 - 2 ORANG	0	0	0	0	
		> 2 ORANG	0	0	0	0	
	STATUS TEMPAT TINGGAL						0.170950594
		DINAS	1	0	1	0	
		MILIK SENDIRI	0	0	0	0	
		KELUARGA	0	0	0	0	
		SEWA / KONTRAK	4	2	2	1	
	TOTAL PENGHASILAN						0.019973094
		< 2.000.000	0	0	0	0	
		2.000.000 - 3.500.000	3	1	2	0.918295834	
		3.600.000 - 4.000.000	2	1	1	1	
		> 4.000.000	2	0	2	0	
	TUNGGAKAN						0
		TIDAK ADA	5	2	3	0.970950594	
		MENUNGGAK	0	0	0	0	
		TIDAK MENUNGGAK	0	0	0	0	

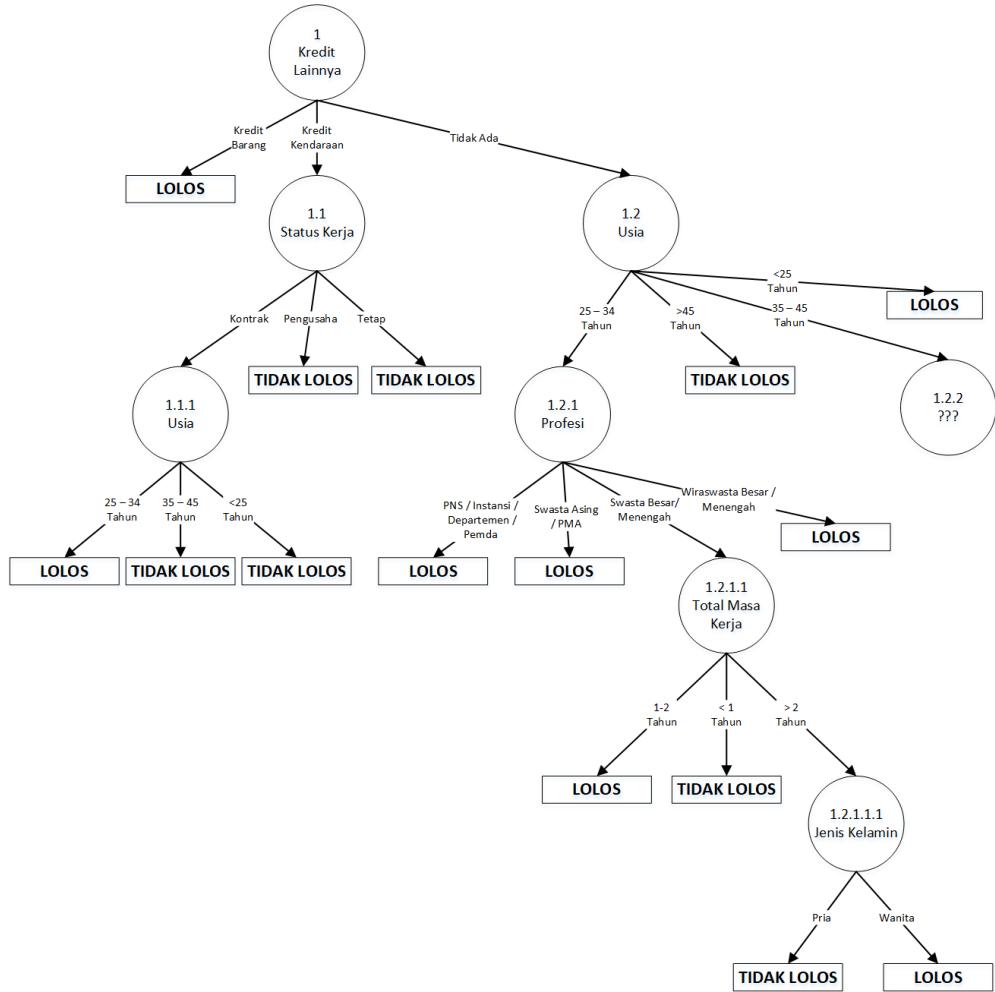
Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukan bahwa atribut JENIS KELAMIN memiliki nilai Information Gain yang paling tinggi yaitu 0. 419973094 dengan demikian atribut TOTAL MASA KERJA dapat menjadi Internal Node/cabang node 1.2.1.1 dari nilai atribut TOTAL MASA KERJA > 2 TAHUN. Dan internal

node/cabang ini merupakan cabang terakhir, karena nilai atribut sudah memenuhi syarat berhenti membuat pohon keputusan.

14. Pembentukan Pohon Keputusan Node 1.2.1.1.1.

Setelah mendapatkan internal node/cabang 1.2.1.1.1, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai Entropy, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari JENIS KELAMIN yaitu PRIA dan WANITA, dan untuk menentukan keputusan hasil untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan. Dan Suatu variable tidak memiliki nilai (kosong). Jika kondisi ini terjadi. Buatlah sebuah leaf yang berisi kelas yang paling banyak muncul atau pemilihan berdasarkan pengetahuan pakar atau memilih pada hasil acak. Disini nilai atribut PRIA bernilai TIDAK LOLOS karena semua kelas berada pada TIDAK LOLOS dan sudah tidak bisa dijadikan untuk membuat cabang karena jumlah sample tidak cukup, sedangkan nilai atribut WANITA bernilai LOLOS karena semua kelas berada pada LOLOS. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2.1.1.1 akan ditunjukkan pada Gambar IV.9:



Sumber: Hasil Penelitian (2020).

Gambar IV. 9

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.1.1.

15. Mencari Nilai Entropy dan Information Gain Node 1.2.2.

Setelah menemukan internal node/cabang dari node 1.2 dan ada 2 nilai atribut dari atribut USIA = 35 – 45 TAHUN yang belum memenuhi kriteria maka kita harus mencari kembali internal node / cabang 1.2.2 dengan cara mencari nilai entropy dari nilai atribut dan nilai information gain dari setiap atribut, berikut adalah tabel hasil perhitungan node 1.2.2 ditunjukkan pada Tabel IV.35:

Tabel IV. 36
Hasil perhitungan Information Gain dan Entropy node 1.2.2.

Node			Jml Kasus (S)	Tidak Lelos (S1)	Lelos (S2)	Entropy	Information Gain
1.2.2	USIA - 35-45 TAHUN		22	6	16	0.845350937	
	JENIS KELAMIN						0.004886164
		PRIA	16	4	12	0.811278124	
		WANITA	6	2	4	0.918295834	
	PENDIDIKAN						0.139695355
		SMA / SLTA	11	2	9	0.684038436	
		DIPLOMA	8	4	4	1	
		S1	3	0	3	0	
	PROFESI						0.111636871
		PNS / Instansi / Departemen / Pemda	6	2	4	0.918295834	
		SWASTA ASING / PMA	1	1	0	0	
		SWASTA BESAR / MENENGAH	7	1	6	0.591672779	
		WIRASWASTA BESAR / MENENGAH	8	2	6	0.811278124	
	TOTAL MASA KERJA						0.055165574
		< 1 TAHUN	2	0	2	0	
		1 - 2 TAHUN	15	5	10	0.918295834	
		> 2 TAHUN	5	1	4	0.721928095	
	STATUS KERJA						0.025994957
		KONTRAK	6	1	5	0.650022422	
		TETAP	8	3	5	0.954434003	
		PENGUSAHA	8	2	6	0.811278124	
	STATUS PERNIKAHAN						0
		MENIKAH	22	6	16	0.845350937	
		BELUM MENIKAH	0	0	0	0	
	JUMLAH TANGGUNGAN						0.147514774
		TIDAK ADA	5	3	2	0.970950594	
		1 - 2 ORANG	15	2	13	0.566509507	
		> 2 ORANG	2	1	1	1	
	STATUS TEMPAT TINGGAL						0.12833105
		DINAS	5	0	5	0	
		MILIK SENDIRI	3	1	2	0.918295834	
		KELUARGA	2	1	1	1	
		SEWA / KONTRAK	12	4	8	0.918295834	
	TOTAL PENGHASILAN						0.058844846
		< 2.000.000	2	0	2	0	
		2.000.000 - 3.500.000	8	2	6	0.811278124	
		3.600.000 - 4.000.000	10	3	7	0.881290899	
	TUNGGAKAN						0
		TIDAK ADA	22	6	16	0.845350937	
		MENUNGGAK	0	0	0	0	

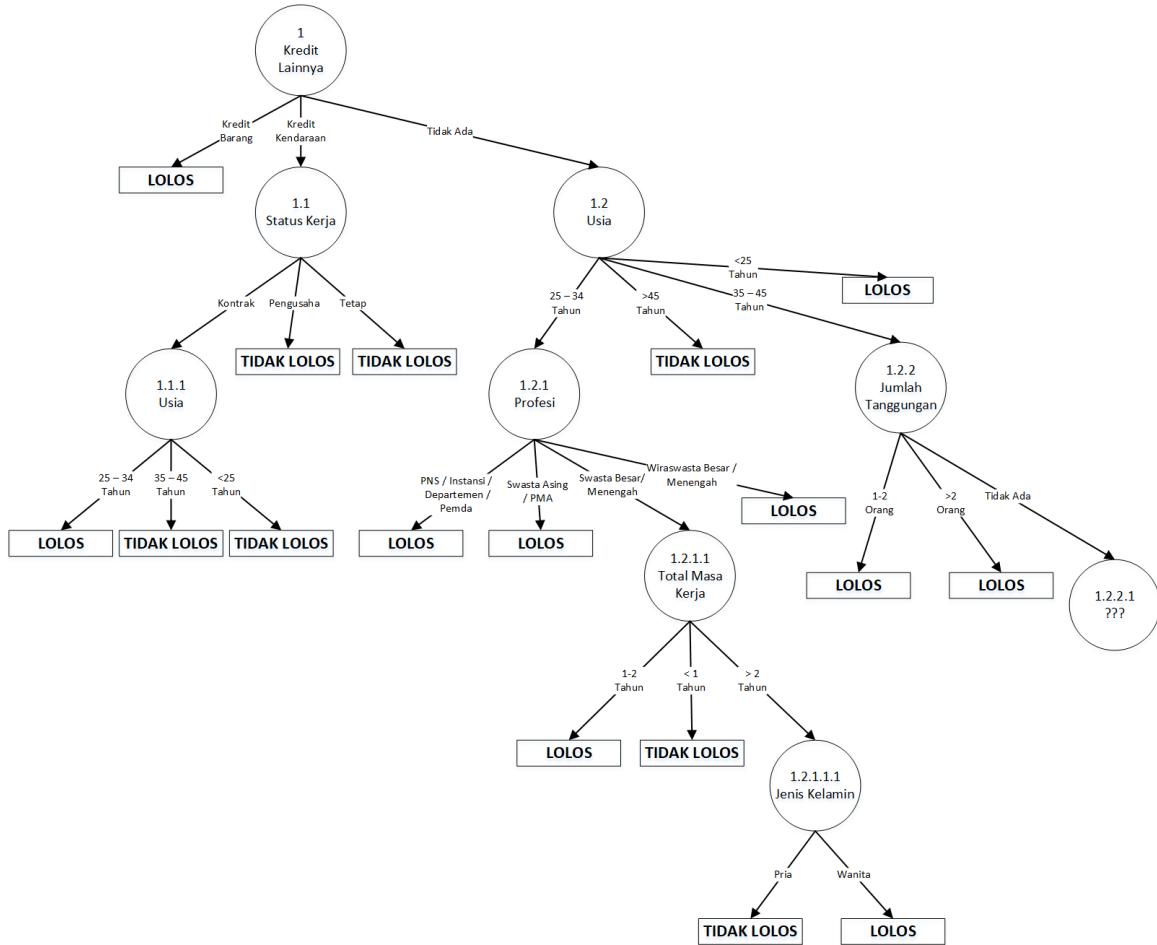
	TIDAK MENUNGGAK	0	0	0	0	
--	--------------------	---	---	---	---	--

Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukan bahwa atribut JUMLAH TANGGUNGAN memiliki nilai Information Gain yang paling tinggi yaitu 0.147514774 dengan demikian atribut JUMLAH TANGGUNGAN dapat menjadi Internal Node/cabang node 1.2.2 dari nilai atribut USIA = 35 – 45 TAHUN.

16. Pembentukan Pohon Keputusan Node 1.2.2.

Setelah mendapatkan internal node/cabang 1.2.2, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai Entropy, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari JUMLAH TANGGUNGAN yaitu TIDAK ADA, 1 – 2 ORANG, > 2 ORANG, dan untuk menentukan keputusan hasil untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan. Dan Suatu variable tidak memiliki nilai (kosong).Jika kondisi ini terjadi. Buatlah sebuah leaf yang berisi kelas yang paling banyak muncul atau pemilihan berdasarkan pengetahuan pakar atau memilih pada hasil acak. Disini nilai atribut TIDAK ADA tidak bernilai dan akan dibuatkan cabang karena belum memenuhi syarat berhenti, 1 – 2 ORANG, > 2 ORANG bernilai LOLOS karena semua kelas berada pada LOLOS. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2.2 akan ditunjukan pada Gambar IV.10:



Sumber: Hasil Penelitian (2020).

Gambar IV. 10

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.2.

17. Mencari Nilai Entropy dan Information Gain Node 1.2.2.1

Setelah menemukan internal node/cabang dari node 1.2.2 dan ada 1 nilai atribut dari atribut JUMLAH TANGGUNGAN = TIDAK ADA yang belum memenuhi kriteria maka kita harus mencari kembali internal node / cabang 1.2.2.1 dengan cara mencari nilai entropy dari nilai atribut dan nilai information gain dari setiap atribut, berikut adalah tabel hasil perhitungan node 1.2.2.1 ditunjukkan pada Tabel IV.36:

Tabel IV. 37
Hasil perhitungan Information Gain dan Entropy node 1.2.2.1.

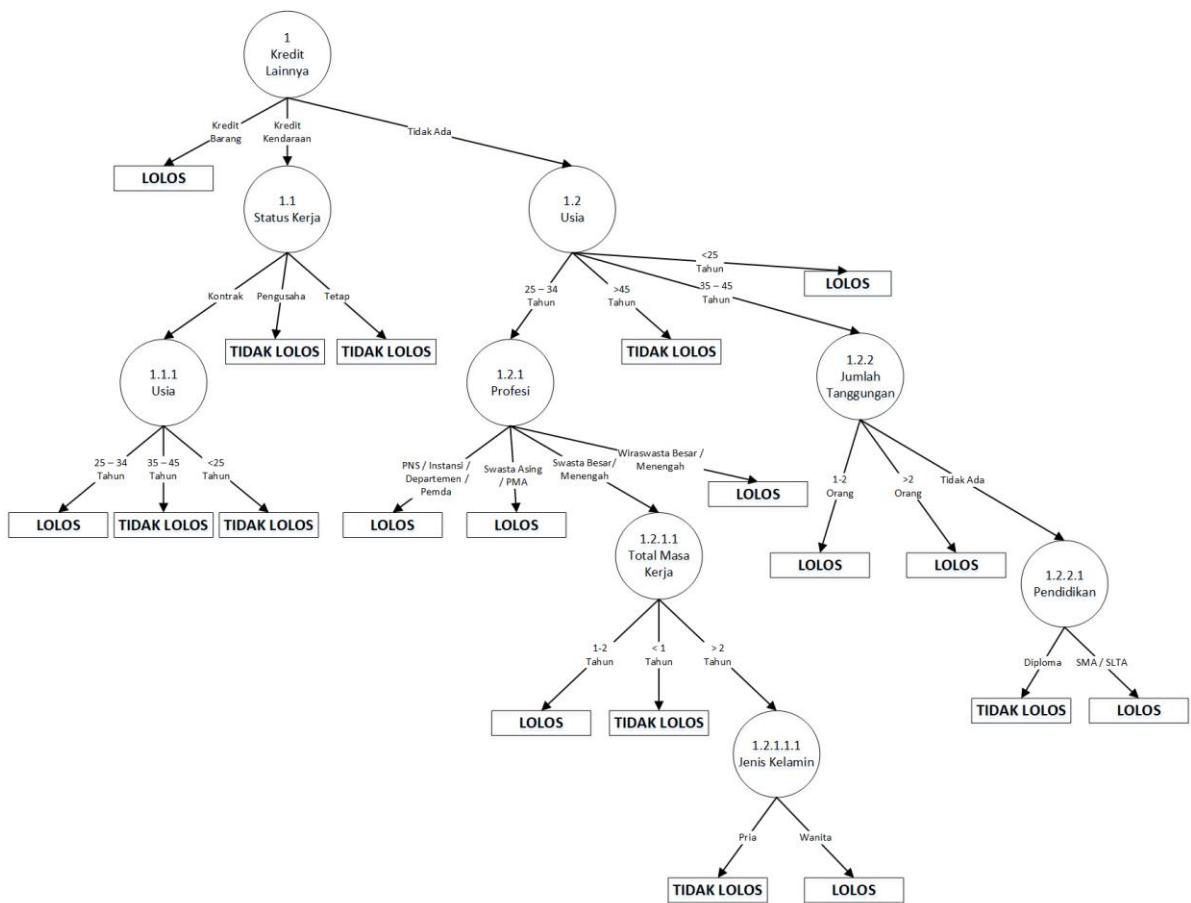
Node			Jml Kasus (S)	Tidak Lulus (S1)	Lolos (S2)	Entropy	Information Gain
1.2.2.1	JUMLAH TANGGUNGAN - TIDAK ADA		5	3	2	0.970950594	
	JENIS KELAMIN						0.170950594
		PRIA	4	2	2	1	
		WANITA	1	1	0	0	
	PENDIDIKAN						0.970950594
		SMA / SLTA	2	0	2	0	
		DIPLOMA	3	3	0	0	
		S1	0	0	0	0	
	PROFESI						0.170950594
		PNS / Instansi / Departemen / Pemda	1	1	0	0	
		SWASTA ASING / PMA	0	0	0	0	
		SWASTA BESAR / MENENGAH	2	1	1	1	
		WIRASWASTA BESAR MENENGAH	/ 2	1	1	1	
	TOTAL MASA KERJA						0.321928095
		< 1 TAHUN	1	0	1	0	
		1 - 2 TAHUN	4	3	1	0.811278124	
		> 2 TAHUN	0	0	0	0	
	STATUS KERJA						0.570950594
		KONTRAK	1	0	1	0	
		TETAP	2	2	0	0	
		PENGUSAHA	2	1	1	1	
	STATUS PERNIKAHAN						0
		MENIKAH	5	3	2	0.970950594	
		BELUM MENIKAH	0	0	0	0	
	STATUS TEMPAT TINGGAL						0.970950594
		DINAS	0	0	0	0	
		MILIK SENDIRI	1	1	0	0	
		KELUARGA	0	0	0	0	
		SEWA / KONTRAK	4	4	0	0	
	TOTAL PENGHASILAN						0.419973094
		< 2.000.000	0	0	0	0	
		2.000.000 - 3.500.000	1	1	0	0	
		3.600.000 - 4.000.000	3	1	2	0.918295834	
		> 4.000.000	1	1	0	0	
	TUNGGAKAN						0
		TIDAK ADA	5	3	2	0.970950594	
		MENUNGGAK	0	0	0	0	
		TIDAK MENUNGGAK	0	0	0	0	

Sumber: Hasil Penelitian (2020).

Dari tabel diatas menunjukan bahwa atribut PENDIDIKAN memiliki nilai Information Gain yang paling tinggi yaitu 0. 970950594 dengan demikian atribut PENDIDIKAN dapat menjadi Internal Node/cabang node 1.2.2.1 dari nilai atribut JUMLAH TANGGUNGJAN = TIDAK ADA.

18. Pembentukan Pohon Keputusan Node 1.2.2.1.

Setelah mendapatkan internal node/cabang 1.2.2.1, selanjutnya menentukan hasil keputusan dari setiap nilai atribut berdasarkan nilai Entropy, Jika memenuhi ketentuan berhenti pembuatan pohon keputusan maka harus berenti, jika tidak maka lanjutnya dengan pembuatan cabang. Ada tiga nilai atribut dari PENDIDIKAN yaitu SMA/SLTA, DIPLOMA, S1, dan untuk menentukan keputusan hasil untuk setiap nilai atribut, apabila data latih kosong (tidak memiliki isi) Pada kondisi ini, pembentukan tree dihentikan dan menulis kelas yang yang paling banyak muncul. Disini nilai atribut SMA/SLTA bernilai LOLOS karena kelas paling banyak muncul LOLOS, sedangkan nilai atribut DIPLOMA bernilai TIDAK LOLOS karena kelas terbanyak pada TIDAK LOLOS, S1 tidak memiliki nilai karena tidak memiliki data sample dan tidak ditulis. Dan berikut ini adalah tree dari rekapitulasi nilai Entropy dan Information Gain Node 1.2.2.1 akan ditunjukan pada Gambar IV.11:



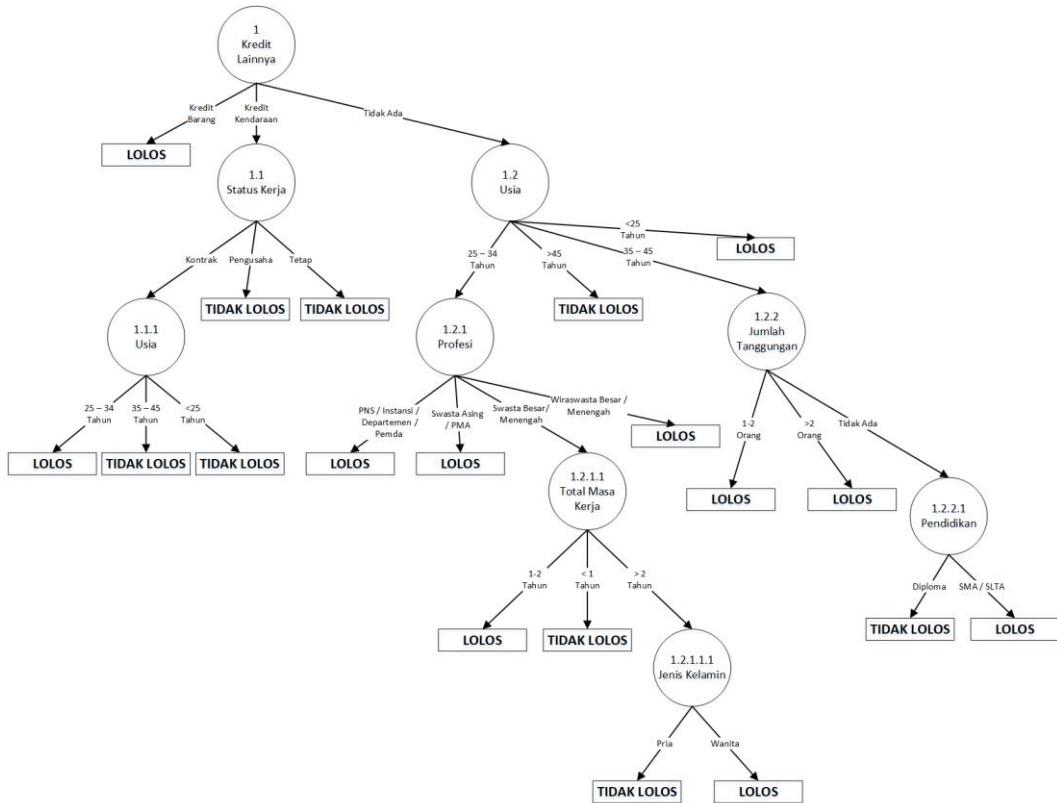
Sumber: Hasil Penelitian (2020).

Gambar IV. 11

Pembentukan Decision Tree dari nilai Entropy dan Information Gain node 1.2.2.1.

4.2.7 Hasil Akhir Pohon Keputusan (Decision Tree)

Berikut adalah hasil akhir dari perhitungan manual algoritma C4.5 adalah pohon keputusan / Decision tree:



Sumber: Hasil Penelitian (2020).

Gambar IV. 12
Pohon Keputusan / Decision Tree akhir.

4.2.8 Mengubah Pohon Keputusan Menjadi Aturan (Tree to Rules)

Setelah mendapatkan pohon keputusan, pohon keputusan akan dikonversikan menjadi klasifikasi *rules*, berikut ini adalah rules yang berasal dari pohon keputusan ditunjukkan pada Tabel IV.38:

Tabel IV. 38

Mengubah Pohon Keputusan/decision tree menjadi klasifikasi aturan.

No	Klasifikasi Aturan
1	If Kredit Lainnya = Kredit Barang Then LOLOS
2	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = 25-34 Then LOLOS
3	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = 35-45 Then TIDAK LOLOS

4	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = <25 Then TIDAK LOLOS
5	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Pengusaha Then TIDAK LOLOS
6	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Tetap Then TIDAK LOLOS
7	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = PNS / Instansi / Departemen / Pemda Then LOLOS
8	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Asing / PMA Then LOLOS
9	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = 1-2 Tahun Then LOLOS
10	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = < 1 Tahun Then TIDAK LOLOS
11	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = > 2 Tahun And Jenis Kelamin = PRIA Then TIDAK LOLOS
12	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = > 2 Tahun And Jenis Kelamin = WANITA Then LOLOS
13	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Wiraswasta Besar / Menengah Then LOLOS
14	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = 1-2 Orang Then LOLOS
15	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = > 2 Orang Then LOLOS
16	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = Tidak Ada And Pendidikan = Diploma Then TIDAK LOLOS
17	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = Tidak Ada And Pendidikan = SMA / SLTA Then LOLOS
18	If Kredit Lainnya = Tidak Ada And Usia = <25 Then LOLOS
19	If Kredit Lainnya = Tidak Ada And Usia = >45 Then TIDAK LOLOS

Sumber: Hasil Penelitian (2020).

4.3. Hasil

Setelah melakukan proses algoritma C4.5 hasil yang diperoleh dapat dijelaskan dalam uraian pada 4.3.1, yaitu:

4.3.1 Analisa Hasil

Uraian hasil pengujian 120 data pengajuan kredit rumah subsidi dapat dilihat pada tabel IV.39:

Tabel IV. 39
Hasil Analisis Menggunakan Algoritma C4.5

No	Klasifikasi Aturan	Prediksi (Ket)
1	If Kredit Lainnya = Kredit Barang Then LOLOS	LOLOS
2	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = 25-34 Then LOLOS	LOLOS
3	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = 35-45 Then TIDAK LOLOS	TIDAK LOLOS
4	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Kontrak And Usia = <25 Then TIDAK LOLOS	TIDAK LOLOS
5	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Pengusaha Then TIDAK LOLOS	LOLOS
6	If Kredit Lainnya = Kredit Kendaraan And Status Kerja = Tetap Then TIDAK LOLOS	LOLOS
7	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = PNS / Instansi / Departemen / Pemda Then LOLOS	LOLOS
8	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Asing / PMA Then LOLOS	LOLOS
9	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = 1-2 Tahun Then LOLOS	LOLOS
10	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = < 1 Tahun Then TIDAK LOLOS	TIDAK LOLOS

11	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = > 2 Tahun And Jenis Kelamin = PRIA Then TIDAK LOLOS	TIDAK LOLOS
12	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Swasta Besar / Menengah And Total Masa Kerja = > 2 Tahun And Jenis Kelamin = WANITA Then LOLOS	LOLOS
13	If Kredit Lainnya = Tidak Ada And Usia = 25-34 And Profesi = Wiraswasta Besar / Menengah Then LOLOS	LOLOS
14	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = 1-2 Orang Then LOLOS	LOLOS
15	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = > 2 Orang Then LOLOS	LOLOS
16	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = Tidak Ada And Pendidikan = Diploma Then TIDAK LOLOS	TIDAK LOLOS
17	If Kredit Lainnya = Tidak Ada And Usia = 35-45 And Jumlah Tanggungan = Tidak Ada And Pendidikan = SMA / SLTA Then LOLOS	LOLOS
18	If Kredit Lainnya = Tidak Ada And Usia = <25 Then LOLOS	LOLOS
19	If Kredit Lainnya = Tidak Ada And Usia = >45 Then TIDAK LOLOS	TIDAK LOLOS

Sumber: Hasil Penelitian (2020).

Berdasarkan hasil pada tabel IV.38 berdasarkan jumlah sample 120 pengajuan kredit rumah subsidi, dapat dilihat bahwa ada *rules* / aturan yang membuat pengajuan kredit rumah subsidi disetujui pihak bank,yaitu:

1. Customer dengan kriteria kredit lainnya = kredit barang
2. Customer dengan kriteria kredit lainnya = kredit kendaraan dan status kerja = kontrak dan usia = 25-34 tahun.
3. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 24 – 34 tahun dan profesi = PNS/Instansi/Departemen/Pemda.

4. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = swasta asing / PMA.
5. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = swasta besar / menengah dan maa kerja = 1-2 Tahun.
6. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = swasta besar/menengah dan total masa kerja = > 2 Tahun dan jenis kelamin = wanita.
7. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = wiraswasta besar / menengah.
8. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 35 – 45 tahun dan jumlah tanggungan 1 – 2 orang.
9. Customer dengan kriteria kredit lainnya = tidak ada usia = 35 – 45 tahun dan jumlah tanggungan = > 2 orang.
10. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 35 – 45 tahun dan jumlah tanggungan tidak ada dan pendidikan = SMA/SLTA.
11. Customer dengan kriteria kredit lainnya = tidak ada dan usia = < 25 tahun.

Dan berikut rules / aturan yang membuat pengajuan kredit rumah subsidi tidak disetujui / tidak lolos.

1. Customer dengan kriteria kredit lainnya = kredit kendaraan dan status kerja = kontrak dan usia = 35 – 45 tahun.

2. Customer dengan kriteria kredit lainnya = kredit kendaraan dan status kerja = kontrak dan usia = < 25 tahun.
3. Customer dengan kriteria kredit lainnya = kredit kendaraan dan status kerja = pengusaha.
4. Customer dengan kriteria kredit lainnya = kredit kendaraan dan status kerja = tetap.
5. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = swasta besar/menengah dan total masa kerja = < 1 tahun.
6. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 25 – 34 tahun dan profesi = swasta besar/menengah dan total masa kerja = > 2 Tahun dan jenis kelamin = pria
7. Customer dengan kriteria kredit lainnya = tidak ada dan usia = 35 – 45 tahun dan jumlah tanggungan = tidak ada dan pendidikan = diploma.
8. Customer dengan kriteria kredit lainnya = tidak ada dan usia = >45 tahun.

4.3.2 Pengujian Algoritma C4.5

Pengujian terhadap klasifikasi rules pengajuan kredit perumahan Cahaya Darussalam 2 Bekasi menggunakan metode *confusion matrix* dengan *RapidMiner Studio 8.2*, ada 2 indikator yang menjadi tolak ukur, yaitu presisi dan akurasi, ditunjukkan pada gambar IV.12:

PerformanceVector

```

PerformanceVector:
accuracy: 90.83%
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 78 10
TIDAK LOLOS: 1 31
precision: 96.88% (positive class: TIDAK LOLOS)
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 78 10
TIDAK LOLOS: 1 31
recall: 75.61% (positive class: TIDAK LOLOS)
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 78 10
TIDAK LOLOS: 1 31
AUC (optimistic): 0.960 (positive class: TIDAK LOLOS)
AUC: 0.943 (positive class: TIDAK LOLOS)
AUC (pessimistic): 0.926 (positive class: TIDAK LOLOS)

```

Sumber: Hasil Penelitian (2020).

Gambar IV. 13

Hasil Performa Pengujian Algoritma C4.5.

accuracy: 90.83%			
	true LOLOS	true TIDAK LOLOS	class precision
pred. LOLOS	78	10	88.64%
pred. TIDAK LOLOS	1	31	96.88%
class recall	98.73%	75.61%	

Sumber: Hasil Penelitian (2020).

Gambar IV. 14

Hasil pengujian Akurasi Algoritma C4.5.

Untuk hasil pengujian presisi ditunjukan pada gambar IV.13.

precision: 96.88% (positive class: TIDAK LOLOS)			
	true LOLOS	true TIDAK LOLOS	class precision
pred. LOLOS	78	10	88.64%
pred. TIDAK LOLOS	1	31	96.88%
class recall	98.73%	75.61%	

Sumber: Hasil Penelitian (2020).

Gambar IV. 15

Hasil pengujian Presisi Algoritma C4.5.

Dari hasil diatas mempunyai arti hasil pengujian berdasarkan indikator Akurasi 91% dan Presisi 97% bahwa klasifikasi rules menggunakan Algoritma C4.5 pengajuan kredit rumah subsidi menggunakan algoritma C4.5 memiliki tingkat akurasi dan presisi yang sangat tinggi.

Perbandingan Hasil performa yang didapatkan dengan menggunakan algoritma klasifikasi Naïve Bayes untuk klasifikasi rules pengajuan kredit rumah subsidi, ditunjukan pada gambar IV.15:

PerformanceVector

```

PerformanceVector:
accuracy: 85.00%
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 76 15
TIDAK LOLOS: 3 26
precision: 89.66% (positive class: TIDAK LOLOS)
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 76 15
TIDAK LOLOS: 3 26
recall: 63.41% (positive class: TIDAK LOLOS)
ConfusionMatrix:
True: LOLOS TIDAK LOLOS
LOLOS: 76 15
TIDAK LOLOS: 3 26
AUC (optimistic): 0.906 (positive class: TIDAK LOLOS)
AUC: 0.906 (positive class: TIDAK LOLOS)
AUC (pessimistic): 0.906 (positive class: TIDAK LOLOS)

```

Sumber: Hasil Penelitian (2020).

Gambar IV. 16
Hasil Performa Pengujian Algoritma Naïve Bayes.

accuracy: 85.00%			
	true LOLOS	true TIDAK LOLOS	class precision
pred. LOLOS	76	15	83.52%
pred. TIDAK LOLOS	3	26	89.66%
class recall	96.20%	63.41%	

Sumber: Hasil Penelitian (2020).

Gambar IV. 17
Hasil Pengujian Akurasi Algoritma Naïve Bayes.

precision: 89.66% (positive class: TIDAK LOLOS)			
	true LOLOS	true TIDAK LOLOS	class precision
pred. LOLOS	76	15	83.52%
pred. TIDAK LOLOS	3	26	89.66%
class recall	96.20%	63.41%	

Sumber: Hasil Penelitian (2020).

Gambar IV. 18

Hasil Pengujian Presisi Algoritma Naïve Bayes.

Dari hasil diatas mempunyai arti hasil pengujian berdasarkan indikator Akurasi 85% dan Presisi 89% bahwa klasifikasi rules menggunakan Algoritma Naïve Bayes pengajuan kredit rumah subsidi menggunakan algoritma Naïve Bayes memiliki tingkat akurasi dan presisi yang cukup tinggi.

Dapat simpulkan hasil pengujian performa dengan indikator dan akurasi bahwa **Algoritma C4.5** memiliki akurasi dan presisi yang **lebih tinggi** dibandingkan dengan algoritma **Naïve Bayes**.

4.4. Implementasi Algoritma C4.5 dengan *RapidMiner Studio 8.2*

Dalam pengimplementasian Algoritma C4.5 pada *RapidMiner Studio 8.2* terdapat langkah – langkah yang harus dilakukan sampai terdapat aturan klasifikasi sesuai dengan yang diharapkan, adapun langkah – langkahnya adalah sebagai berikut:

1. Add Data Sample

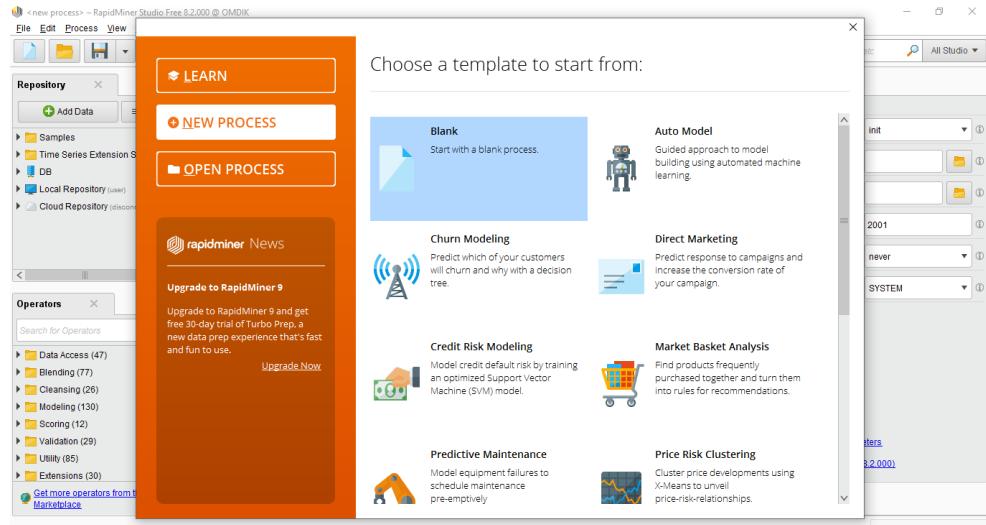
Siapkan data sample dalam akan dijadikan aturan klasifikasi dalam format *excel*, data yang akan diuji sudah dilakukan Data Selection, Data Preprocessing

/ Data Cleaning, Data Transformation, dan Data Reduction. Jika sudah melewati tahap itu, buka *Software RapidMiner Studio 8.2* lalu **CREATE NEW PROCESS -> KLIK ADD DATA -> PILIH MY COMPUTER -> CARI FILE ANDA -> PILIH ATRIBUT YANG AKAN DIJADIKAN SAMPLE -> GANTI ROLE KEPUTUSAN MENJADI LABEL.** Ditunjukan pada gambar IV.19:



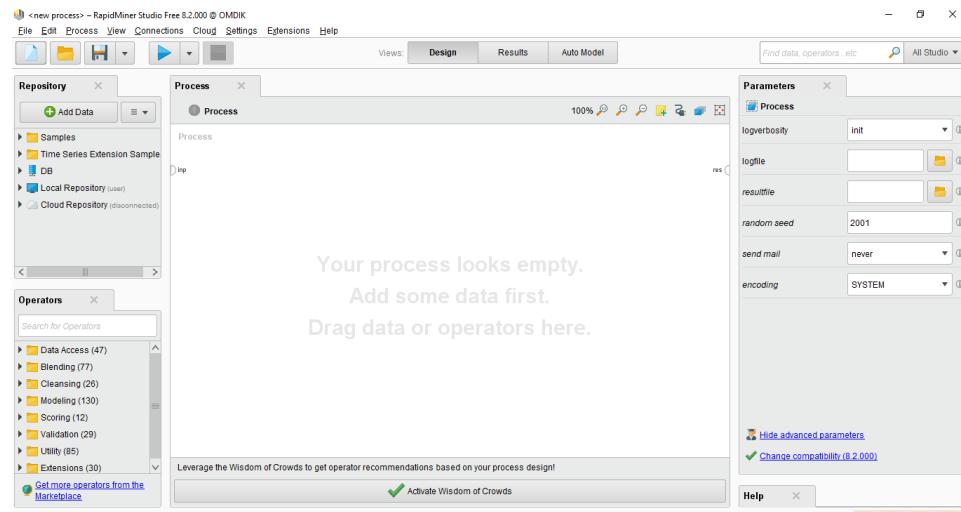
Sumber: Hasil Penelitian (2020).

Gambar IV. 19
Splash Screen RapidMiner Studio 8.2.



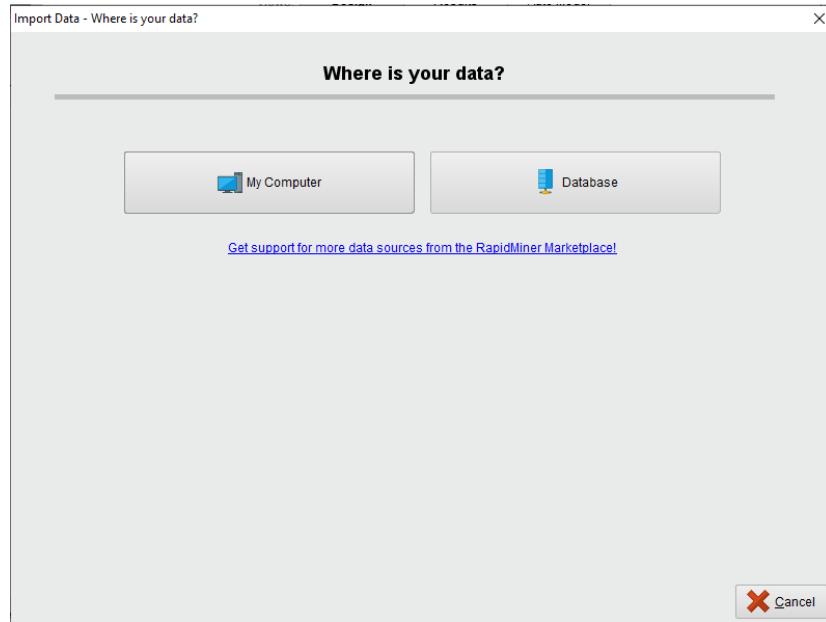
Sumber: Hasil Penelitian (2020).

Gambar IV. 20 Landing Page RapidMiner Studio 8.2.



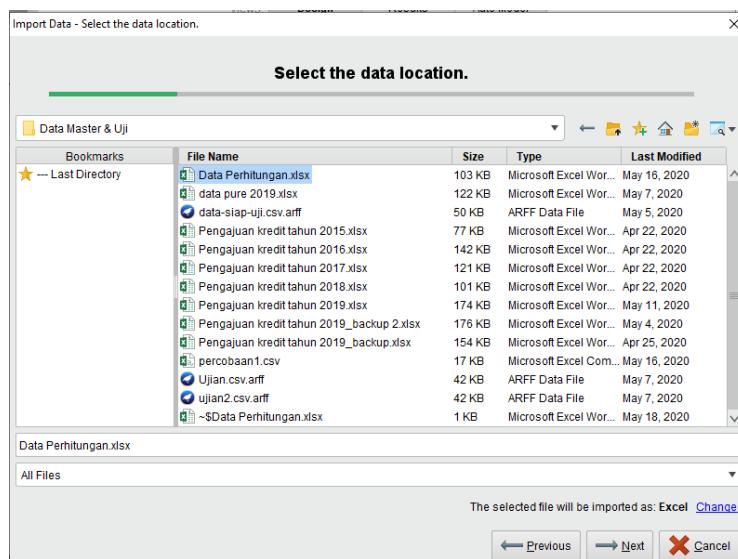
Sumber: Hasil Penelitian (2020).

Gambar IV. 21 Home Page RapidMiner Studio 8.2.



Sumber: Hasil Penelitian (2020).

Gambar IV. 22
Import Data Sample dari My Computer.



Sumber: Hasil Penelitian (2020).

Gambar IV. 23
Mencari File Data Sample.

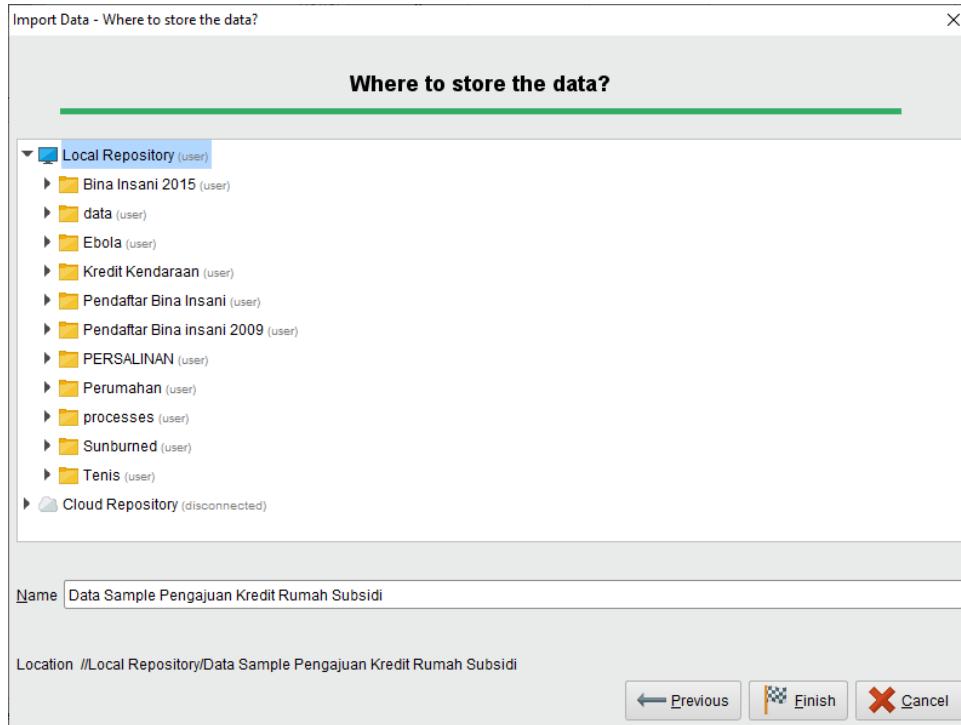
Import Data - Select the cells to import.										
Select the cells to import.										
Sheet:	Data Sample		Cell range:		C:O	Select All		<input checked="" type="checkbox"/> Define hea...		
A	B	C	D	E	F	G	H	I	J	
1	No	Nama	Jenis Ke...	Usia	Pendidik...	Profesi	Total Ma...	Status K...	Status P...	Jumlah
2	1.000	AR.REZ...	PRIA	25.000	SMA / SL...	Swasta...	1,2	Tetap	BELUM ...	0.000
3	2.000	MASNITA...	WANITA	31.000	SMA / SL...	Wiraswa...	> 2	Pengusa...	MENIKAH	1.000
4	3.000	RIKO SE...	PRIA	41.000	S1	Wiraswa...	1,2	Pengusa...	MENIKAH	1.000
5	4.000	DINI DWI...	WANITA	25.000	Diploma	Wiraswa...	1,2	Pengusa...	MENIKAH	2.000
6	5.000	NURYANA	PRIA	22.000	Diploma	Wiraswa...	> 2	Pengusa...	BELUM ...	0.000
7	6.000	SANTHI...	WANITA	23.000	Diploma	Swasta...	1,2	Kontrak	BELUM ...	0.000
8	7.000	RICHAR...	PRIA	26.000	SMA / SL...	Wiraswa...	1,2	Pengusa...	BELUM ...	0.000
9	8.000	EDI MUL...	PRIA	26.000	SMA / SL...	Wiraswa...	> 2	Pengusa...	MENIKAH	2.000
10	9.000	MUH AN...	PRIA	28.000	Diploma	Wiraswa...	> 2	Pengusa...	MENIKAH	3.000
11	10.000	NURIMAN	PRIA	30.000	SMA / SL...	Swasta...	1,2	Kontrak	BELUM ...	0.000
12	11.000	AHGY M...	PRIA	21.000	Diploma	Wiraswa...	1,2	Pengusa...	BELUM ...	0.000
13	12.000	IVAN PR...	PRIA	35.000	SMA / SL...	PNS / In...	> 2	Tetap	MENIKAH	1.000
14	13.000	JUNAEDI	PRIA	28.000	Diploma	Wiraswa...	1,2	Pengusa...	BELUM ...	0.000

Sumber: Hasil Penelitian (2020).

Gambar IV. 24
Pilih Atribut yang akan Digunakan untuk sampling.

Sumber: Hasil Penelitian (2020).

Gambar IV. 25
Change Role Atribut keputusan menjadi label.



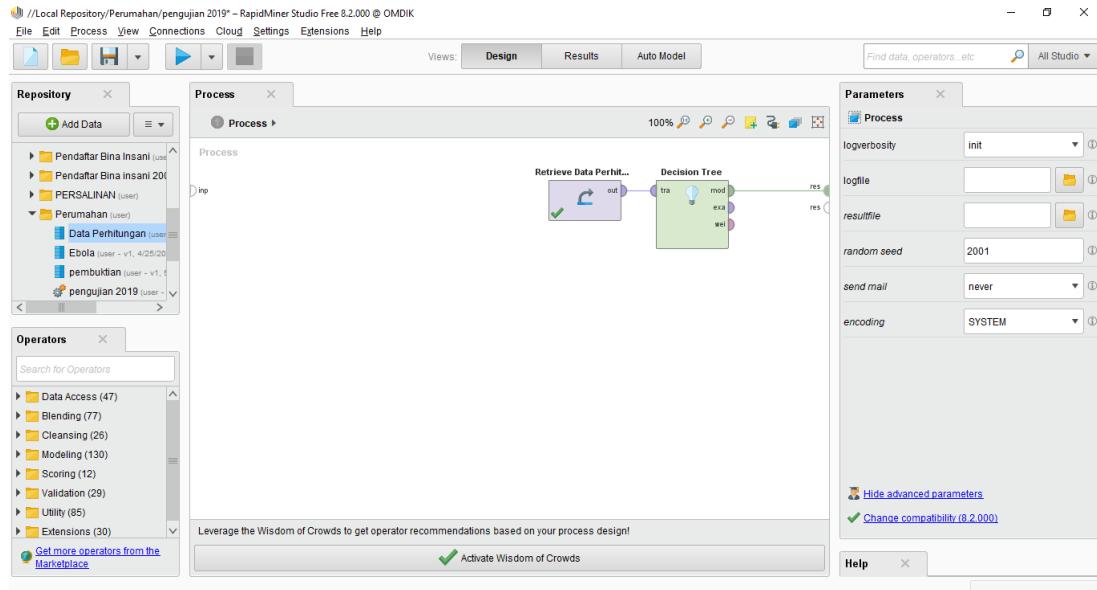
Sumber: Hasil Penelitian (2020).

Gambar IV. 26

Simpan Data Sample.

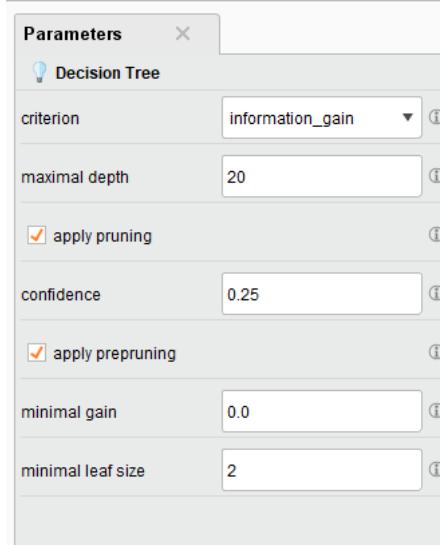
2. Pembuatan Pohon Keputusan Menggunakan Algoritma C4.5

Setelah membuat **NEW PROCESS** selanjutnya memasukan beberapa operator kedalam halaman proses, masukan **DATA SAMPLE -> OPERATOR DECISION TREE -> KLIK 2X OPERATOR DECISION TREE** lalu ganti **PARAMETER CRITERION** menjadi **INFORMATION GAIN, MAXIMAL DEPT = 20, CONFIDENCE = 0.25, MINIMAL GAIN = 0,0** dan **MINIMAL LEAF SIZE = 2** dan sambungkan operator satu sama lain, ditunjukan pada gambar berikut:



Sumber: Hasil Penelitian (2020).

Gambar IV. 27
Pembuatan Pohon Keputusan (Decision Tree).

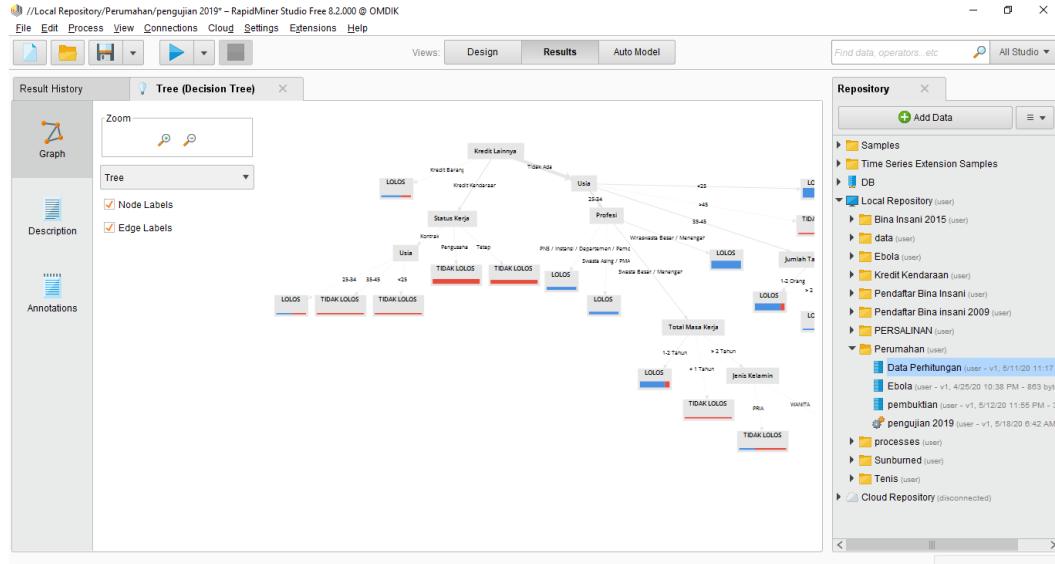


Sumber: Hasil Penelitian (2020).

Gambar IV. 28
Parameter Operator Pohon Keputusan (Decision Tree).

Setelah itu klik **KLIK TOMBOL RUN PROCESS -> KLIK BUTTON**

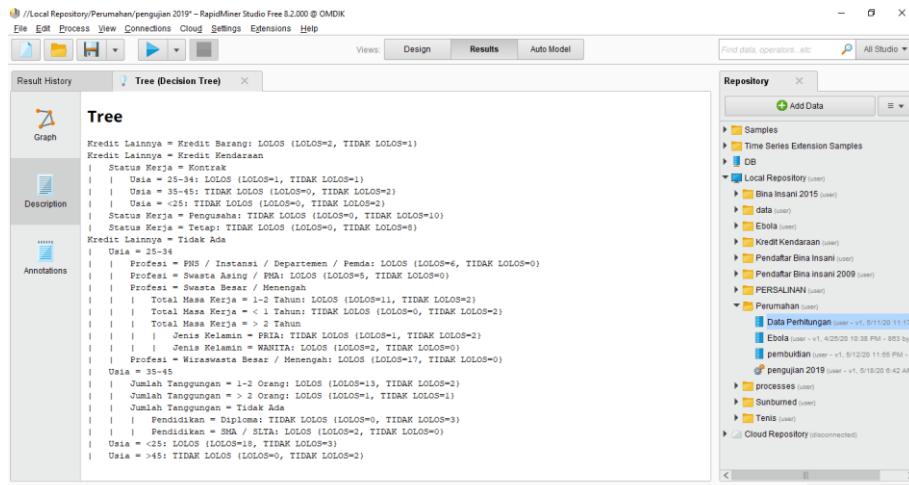
RESULT untuk melihat hasilnya:



Sumber: Hasil Penelitian (2020).

Gambar IV. 29

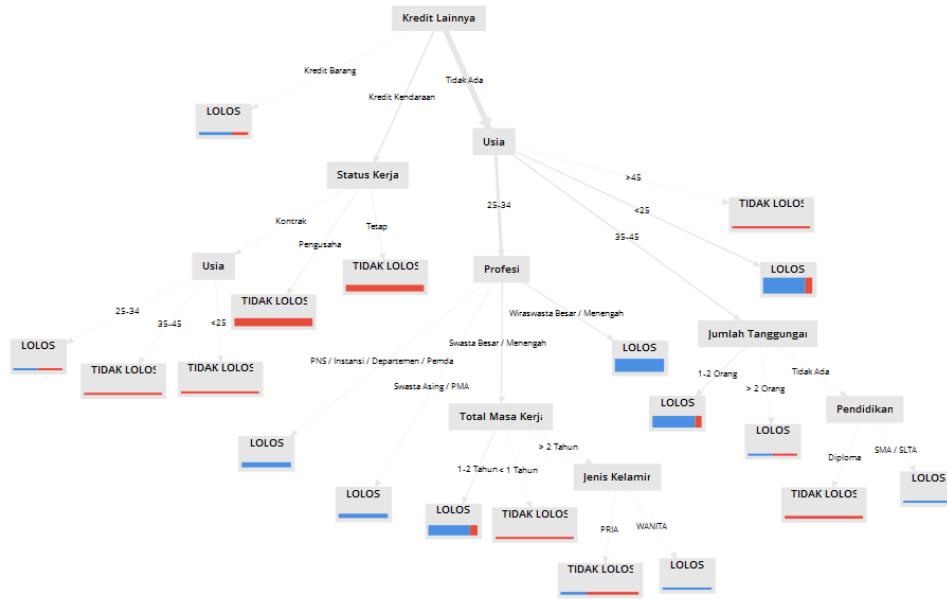
Pohon Keputusan (Decision Tree) Pengajuan Kredit Rumah Subsidi.



Sumber: Hasil Penelitian (2020).

Gambar IV. 30

Deskripsi dari Pohon Keputusan (Decision Tree) Pengajuan Kredit Rumah Subsidi.



Sumber: Hasil Penelitian (2020).

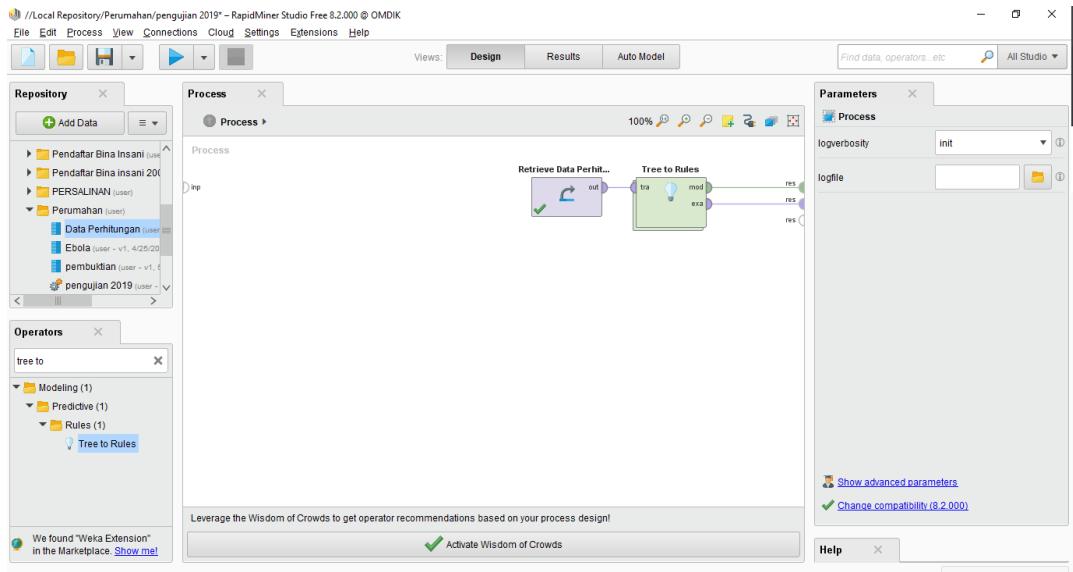
Gambar IV. 31
Hasil Pohon Keputusan Pengajuan Kredit Rumah Subsidi.

3. Konversi pohon keputusan menjadi aturan klasifikasi

Setelah pohon keputusan terbentuk maka langkah selanjutnya adalah merubah/mengkonversi pohon tersebut menjadi aturan klasifikasi,

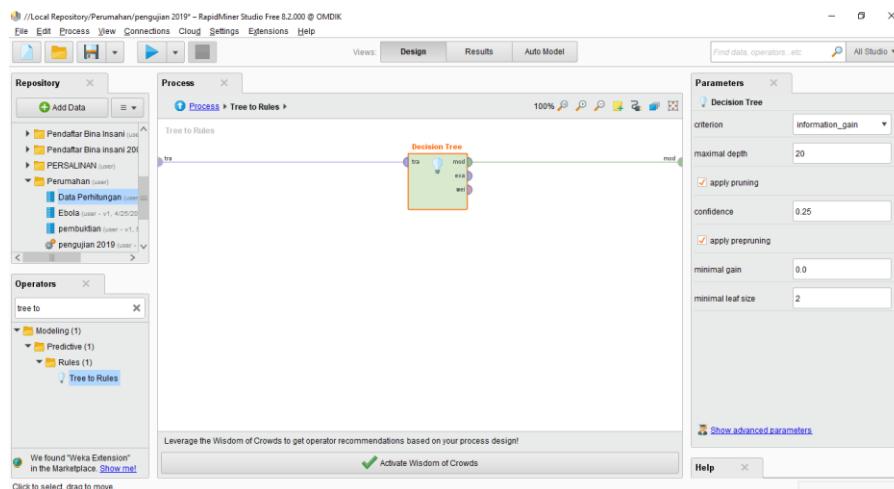
TAMBAHKAN OPERATOR TREE TO RULES KEDALAM HALAMAN PROSES -> KLIK 2X PADA OPERATOR TREE TO RULES

-> TAMBAHKAN OPERATOR DECISION TREE DIDALAMNYA dan sambungkan kembali dan klik tombol **RUN PROCESS -> KLIK BUTTON RESULT** untuk melihat hasil nya, ditunjukan pada gambar berikut ini:



Sumber: Hasil Penelitian (2020).

Gambar IV. 32
Operator Konversi Tree To Rules.

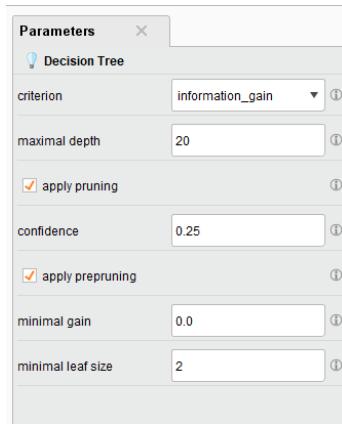


Sumber: Hasil Penelitian (2020).

Gambar IV. 33
Operator Decision Tree didalam Operator Tree To Rules.

KLIK 2X PADA OPERATOR DECISION TREE dan ganti parameter menjadi, **PARAMETER CRITERION** menjadi **INFORMATION GAIN**,

MAXIMAL DEPT = 20, CONFIDENCE = 0.25, MINIMAL GAIN = 0,0 dan MINIMAL LEAF SIZE = 2 dan sambungkan operator satu sama lain, ditunjukan pada gambar berikut:



Sumber: Hasil Penelitian (2020).

Gambar IV. 34
Parameter Operator Decision Tree.

Berikut adalah hasil dari konversi pohon keputusan menjadi aturan klasifikasi:

```

//Local Repository/Perumahan/pengujian 2019 - RapidMiner Studio Free 8.2.00 © OMDIK
File Edit Process View Connections Cloud Settings Eghensions Help
Views Design Results Auto Model Find data, operators, etc. All Studio ▾
Result History ExampleSet (Retrieve Data Perhitungan) RuleModel (Tree to Rules) ▾
RuleModel
if Kredit Lainnya = Tidak Banyak then LOLOS (3 / 3)
if Kredit Lainnya = Kredit Kendaraan dan Status Kerja = Kontak and Usia = 25-34 then LOLOS (1 / 1)
if Kredit Lainnya = Kredit Kendaraan dan Status Kerja = Kontak and Usia = 35-45 then TIDAK LOLOS (0 / 2)
if Kredit Lainnya = Kredit Kendaraan and Status Kerja = Kontak and Usia <25 then TIDAK LOLOS (0 / 10)
if Kredit Lainnya = Kredit Kendaraan and Status Kerja = Pengusaha then TIDAK LOLOS (0 / 10)
if Kredit Lainnya = Kredit Kendaraan and Status Kerja = Tetap then TIDAK LOLOS (0 / 8)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = PNS / Instansi Pemerintah / Penda then LOLOS (6 / 6)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = Swasta / Menengah and Total Masa Kerja = < 1 Tahun then LOLOS (5 / 5)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = Swasta / Menengah and Total Masa Kerja = 1-2 Tahun then LOLOS (11 / 2)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = Swasta / Menengah and Total Masa Kerja = < 1 Tahun then TIDAK LOLOS (0 / 2)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = Swasta / Menengah and Total Masa Kerja = > 1 Tahun and Jenis Kelamin = PRIA then TIDAK LOLOS (0 / 2)
if Kredit Lainnya = Tidak Ada and Usia = 25-34 and Profesi = Swasta / Menengah and Total Masa Kerja = > 2 Tahun and Jenis Kelamin = WANITA then LOLOS (17 / 9)
if Kredit Lainnya = Tidak Ada and Usia = 35-45 and Jumlah Tanggungan = 1 Orang then LOLOS (17 / 17)
if Kredit Lainnya = Tidak Ada and Usia = 35-45 and Jumlah Tanggungan = > 2 Orang then LOLOS (1 / 1)
if Kredit Lainnya = Tidak Ada and Usia = 35-45 and Jumlah Tanggungan = Tidak Ada and Pendidikan = Diploma then TIDAK LOLOS (0 / 3)
if Kredit Lainnya = Tidak Ada and Usia = <25 then LOLOS (18 / 3)
if Kredit Lainnya = Tidak Ada and Usia = >45 then TIDAK LOLOS (0 / 2)
correct: 109 out of 120 training examples.

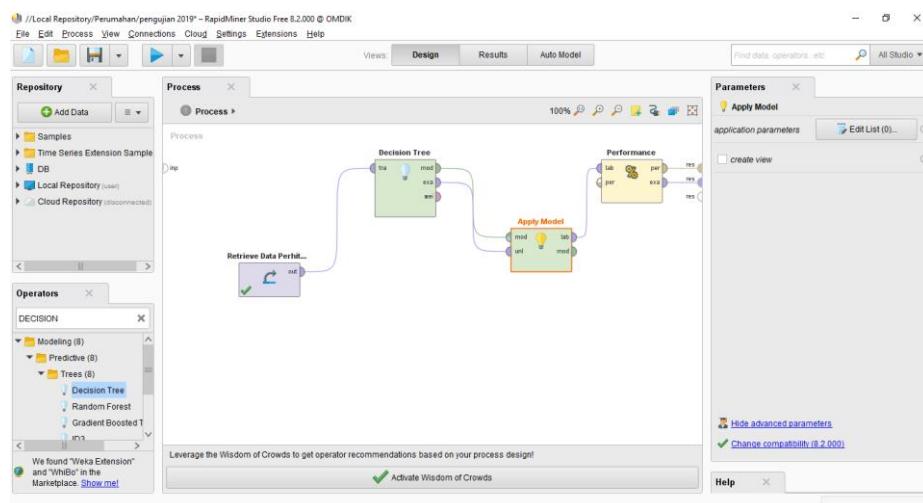
```

Sumber: Hasil Penelitian (2020).

Gambar IV. 35
Hasil Konversi Pohon Keputusan Menjadi Aturan Klasifikasi.

4. Pengujian Algoritma C4.5

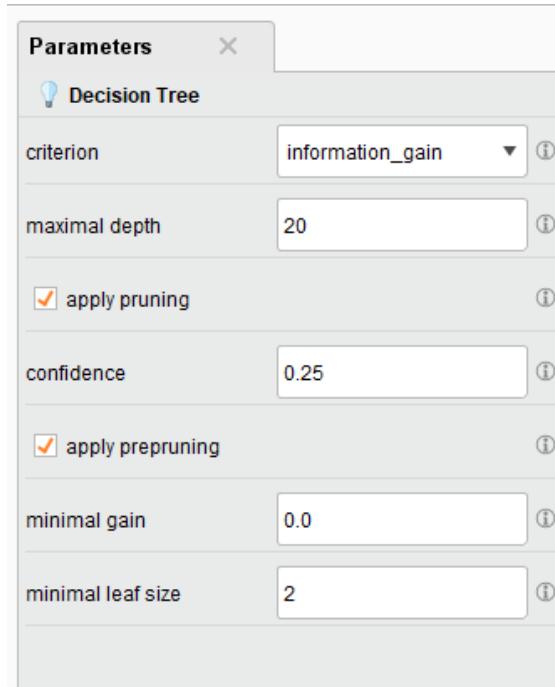
Selanjutnya adalah pengujian terhadap aturan klasifikasi persetujuan kredit rumah subsidi memakai indikator akurasi dan presisi menggunakan metode *Confusion Matrix* dengan *RapidMiner Studio 8.2*, masukan DATA SAMPLE operator DECISION TREE, APPLY MODEL, dan PERFORMANCE kedalam proses, kemudian **KLIK 2X PADA OPERATOR DECISION TREE** dan ganti parameter menjadi, **PARAMETER CRITERION** menjadi **INFORMATION GAIN, MAXIMAL DEPT = 20, CONFIDENCE = 0.25, MINIMAL GAIN = 0,0** dan **MINIMAL LEAF SIZE = 2**, kemudian sambungkan semua operator dan setelah itu **KLIK TOMBON RUN PROCESS -> KLIK RESULT** untuk melihat hasilnya, ditunjukan pada gambar IV.34:



Sumber: Hasil Penelitian (2020).

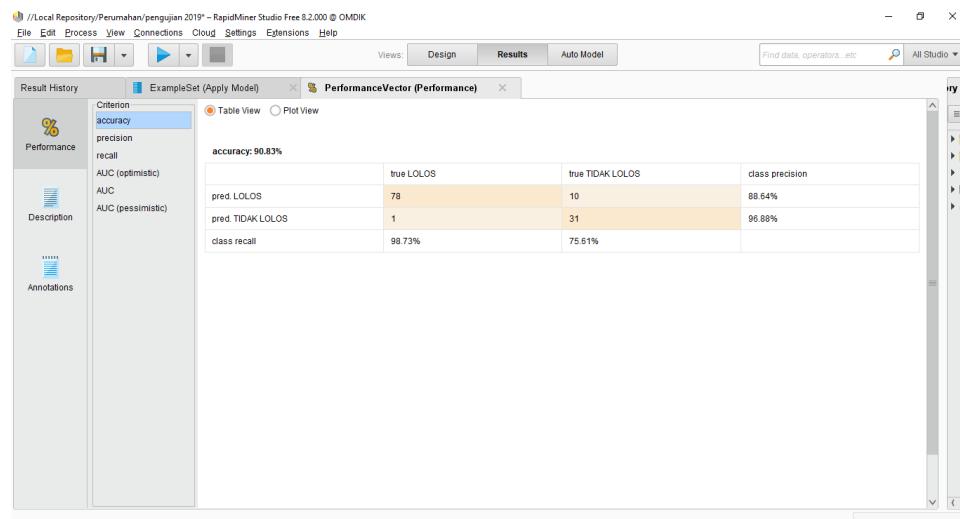
Gambar IV. 36

Pengujian Algoritma C4.5 Menggunakan Confusion Matrix.



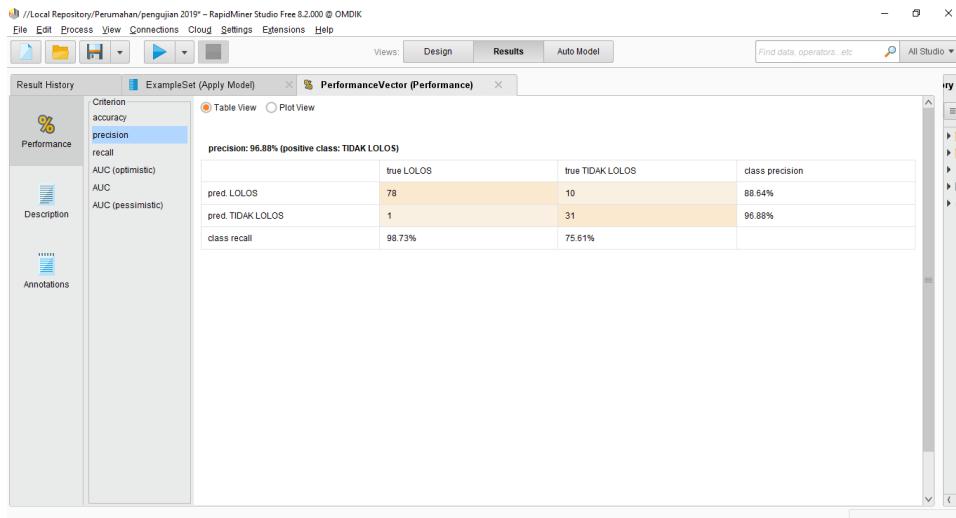
Sumber: Hasil Penelitian (2020).

Gambar IV. 37
Parameter Operator Decision Tree.



Sumber: Hasil Penelitian (2020).

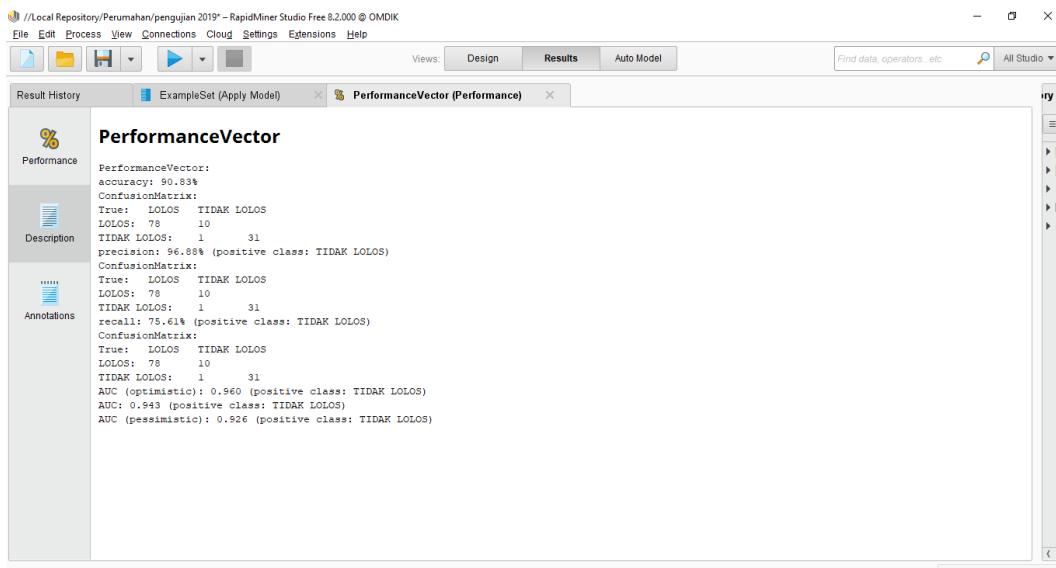
Gambar IV. 38
Hasil Pengujian Akurasi Algoritma C4.5 Menggunakan Confusion Matrix.



Sumber: Hasil Penelitian (2020).

Gambar IV. 39

Hasil Pengujian Presisi Algoritma C4.5 Menggunakan Confusion Matrix.



Sumber: Hasil Penelitian (2020).

Gambar IV. 40

Hasil Pengujian Performa Algoritma C4.5 Menggunakan Confusion Matrix.

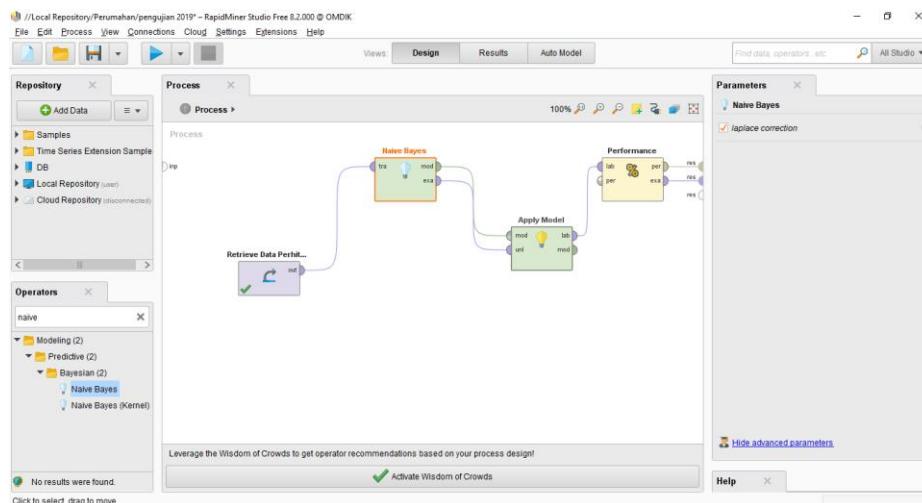
Dari hasil diatas mempunyai arti hasil pengujian berdasarkan indikator Akurasi 91% dan Presisi 97% bahwa klasifikasi rules menggunakan Algoritma

C4.5 pengajuan kredit rumah subsidi menggunakan algoritma C4.5 memiliki tingkat akurasi dan presisi yang sangat tinggi.

5. Perbandingan dengan Algoritma Naïve Bayes

Perbandingan Hasil performa yang didapatkan dengan menggunakan algoritma klasifikasi Naïve Bayes untuk klasifikasi rules pengajuan kredit rumah subsidi.

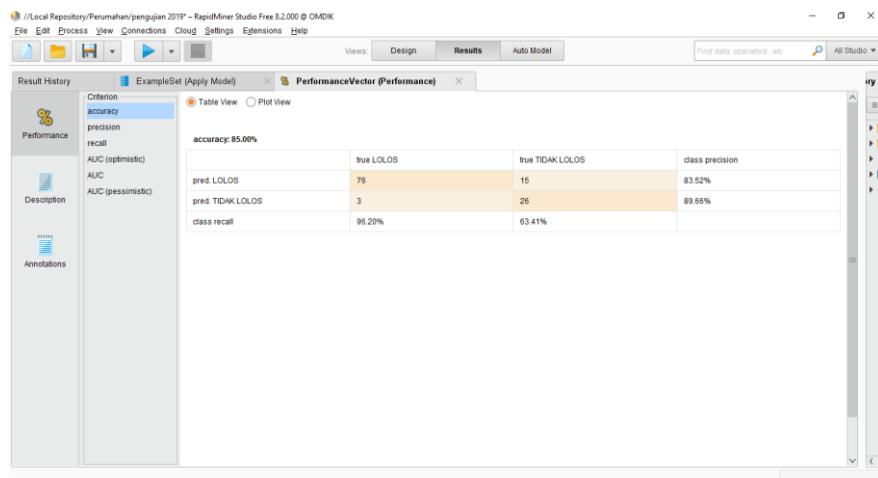
Langkah selanjutnya adalah untuk melakukan perbandingan hasil performa dari algoritma **NAÏVE BAYES**, **masukan DATA SAMPLE operator NAÏVE BAYES, APPLY MODEL, dan PERFORMANCE** kedalam proses, kemudian sambungkan semua operator dan setelah itu **KLIK TOMBON RUN PROCESS -> KLIK RESULT** untuk melihat hasilnya, ditunjukkan pada gambar IV.39:



Sumber: Hasil Penelitian (2020).

Gambar IV. 41

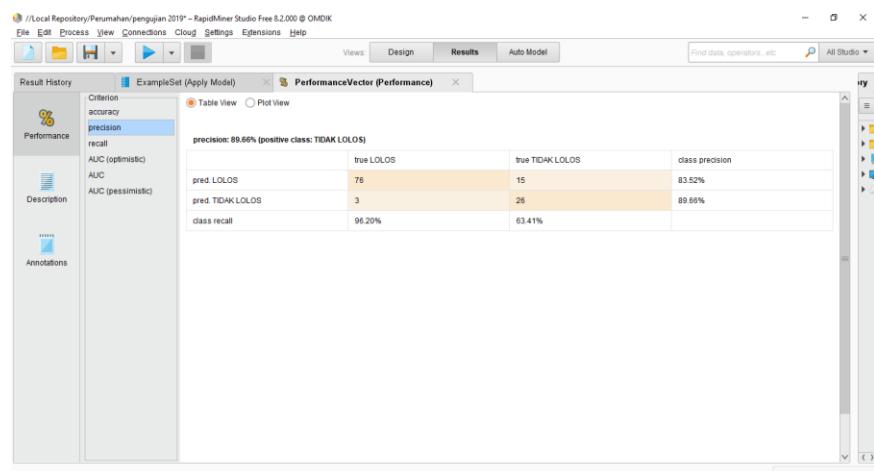
Pengujian Perbandingan Performa dengan algoritma Naïve Bayes Menggunakan Confusion Matrix.



Sumber: Hasil Penelitian (2020).

Gambar IV. 42

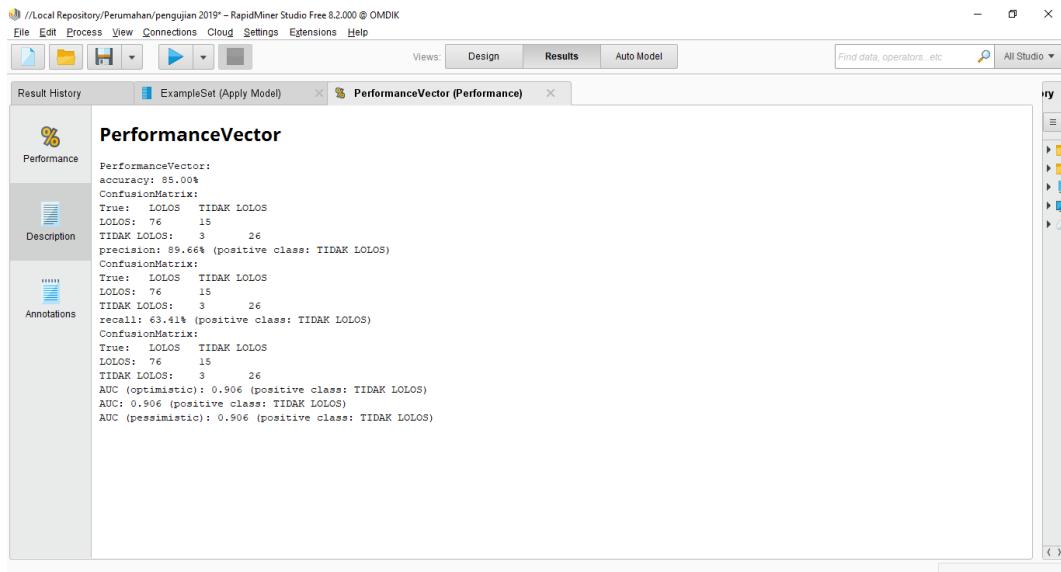
Hasil Pengujian Akurasi Algoritma Naïve Bayes Menggunakan Confusion Matrix.



Sumber: Hasil Penelitian (2020).

Gambar IV. 43

Hasil Pengujian Presisi Algoritma Naïve Bayes Menggunakan Confusion Matrix.



Sumber: Hasil Penelitian (2020).

Gambar IV. 44

Hasil Pengujian Performa Algoritma Naïve Bayes Menggunakan Confusion Matrix.

Dari hasil diatas mempunyai arti hasil pengujian berdasarkan indikator Akurasi 85% dan Presisi 89% bahwa klasifikasi rules menggunakan Algoritma Naïve Bayes pengajuan kredit rumah subsidi menggunakan algoritma Naïve Bayes memiliki tingkat akurasi dan presisi yang cukup tinggi.

Dapat simpulkan hasil pengujian performa dengan indikator dan akurasi bahwa **Algoritma C4.5** memiliki akurasi dan presisi yang **lebih tinggi** dibandingkan dengan algoritma **Naïve Bayes**.

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang dilakukan pada Perumahan Cahaya Darussalam 2 Bekasi mengenai Algoritma C4.5 dalam persetujuan kredit rumah subsidi untuk menentukan aturan klasifikasi persetujuan kredit dan menghasilkan prediksi persetujuan berdasarkan hasil analisa data pengajuan kredit customer yang selama ini hanya menumpuk saja di penyimpanan tanpa ada pemanfaatan lebih untuk mendapatkan informasi yang dibutuhkan. Aturan klasifikasi ini dibuat sesuai dengan kebutuhan dan permintaan dari Perumahan Cahaya Darussalam 2 Bekasi untuk menangi pemanfaatan data pengajuan kredit customer dan menemukan informasi. Penelitian tersebut dapat disimpulkan:

1. Dengan hasil aturan klasifikasi dengan algoritma C4.5, ditemukan kriteria pengajuan kredit rumah subsidi yang mempengaruhi di setujui oleh pihak bank.
2. Dengan hasil aturan klasifikasi dengan algoritma C4.5, ditemukan kriteria pengajuan kredit rumah subsidi yang membuat tidak disetujui oleh pihak bank.
3. Adanya aturan klasifikasi kriteria persetujuan kredit rumah subsidi, marketing dapat memberikan prediksi kepada konsumen bahwa pengajuan

kredit rumah subsidi akan disetujui atau tidak di setujui oleh pihak bank, berdasarkan hasil analisis yang dilakukan.

4. Aturan klasifikasi ini sekaligus menambah proses validasi terhadap kelengkapan data pengajuan kredit yang dilakukan oleh customer perumahan, dan dapat memangkas waktu persetujuan yang diberikan pihak bank sampai dengan 3 bulan, dengan adanya aturan klasifikasi, prediksi persetujuan kredit rumah subsidi didapatkan dalam kurun waktu 2 hari.

5.2 Saran – Saran

Berdasarkan hasil penelitian dan kesimpulan yang didapatkan, selanjutnya peneliti dapat juga memberikan beberapa saran yang relevan dengan hasil penelitian. Saran ini berupa masukan masukan yang ditunjukkan ke organisasi/obyek penelitian dan untuk penelitian selanjutnya:

1. Aspek Manejerial

Menambah jumlah karyawan yang padam Data Mining untuk melakukan pengolahan data pada tahun berikutnya.

2. Aspek Sistem dan Program

Melakukan analisis Data Mining menggunakan aplikasi dan algoritma data mining lainnya, seperti *Naïve Bayes*, *K-Nearest Neightbor*, *Apriori* untuk aplikasinya seperti *WEKA*, *Tanagra*, *Matlab* dan yang lainnya.

3. Aspek Penelitian Selanjutnya

Melakukan perancangan dan pengembangan sistem informasi Data Mining agar pengolahan data perusahaan dapat diolah sendiri tidak perlu menggunakan software dari luar.

DAFTAR PUSTAKA

- Adinugroho S, Sari Arum Y. 2018. Aplikasi Data Mining Menggunakan Weka. Malang: UB Press. Diambil dari: <https://books.google.co.id/books?id=p91qDwAAQBAJ&pg=PR7&dq=aplikasi+data+mining+menggunakan+weka&hl=en&sa=X&ved=0ahUKEwiW3tfui7HhAhVj7nMBHaYoBHUQ6AEIKjAA#v=onepage&q=aplikasi%20data%20mining%20menggunakan%20weka&f=false>. (15 Januari 2020).
- Elisa E. 2017. Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. Jurnal Online Informatika. Diambil Dari: <http://join.if.uinsgd.ac.id/index.php/join/article/view/v2i17>. (16 Januari 2020)
- Marsipah S. 2016. Komparasi Algoritma Klasifikasi Data Mining untuk Evaluasi Pemberian Kredit. Information and Communication Technologies. Diambil dari:<https://media.neliti.com/media/publications/234336-komparasi-algoritma-klasifikasi-data-min-7606386e.pdf>(16 Januari 2020)
- Pratama Az, Dkk. Algoritma C4.5 Untuk Klasifikasi Nasabah Dalam Memprediksi Kredit Macet. Information System For Educators And Professionals. Diambil Dari: <http://ejournal-binainsani.ac.id/index.php/ISBI/article/view/1083> (16 januari 2020)
- Takalamingan HF. 2018. Analisis Penerapan Sistem Dan Prosedur Pemberian Kredit Pemilikan Rumah (Kpr) Subsidi Pada Bank Tabungan Negara Cabang Manado. Jurnal Riset Akuntansi Going Concern. Diambil dari: <https://ejournal.unsrat.ac.id/index.php/gc/article/view/22060> (16 Januari 2020)
- Daniel T.L, Chantal D.L. 2015. Data Mining and Predictive Analytics. Second Edition. New Jersey: Jon Willey& Son, Inc. Diambil dari: <https://archive.org/details/DataMiningAndPredictiveAnalytics>. (19 Maret 2019).
- Han J, Kamber M, Pei J. 2012. Data Mining: Concepts and Techniques. Third Edition. USA: Morgan Kaufmann Publishers. Diambil dari: https://archive.org/details/dataminingconcep00hanj_0. (19 Maret 2019).
- Himayati. 2008. Eksplorasi Zahir Accounting. Jakarta: Elex Media Komputindo. Diambil dari: https://books.google.co.id/books?id=vZbEJi4ULiIC&pg=PA123&lpg=PA123&dq=penjualan&source=bl&ots=IWqSPV7ao_&sig=hdP2payj5Qnol1oS

CSu69C0OC3c&hl=id&sa=X&ei=tpIRUL6tCIbUrQfe3oDIDw&ved=0CEc
Q6AEwBg#v=onepage&q=penjualan&f=false . (19 Maret 2019).

Kusrini, Luthfi TE. 2009. Algoritma Data Mining. Yogyakarta: ANDI.

Nofriansyah D. 2014. Konsep Data Mining Vs Sistem Pendukung Keputusan. Yogyakarta: Deepublish. Diambil dari:
<https://books.google.co.id/books?id=PoJyCAAAQBAJ&printsec=frontcover&dq=.+Konsep+Data+Mining+Vs+Sistem+Pendukung+Keputusan.&hl=en&sa=X&ved=0ahUKEwiGsILjj7HhAhX87XMBHZHiCVsQ6AEIKjAA#v=onepage&q=.%20Konsep%20Data%20Mining%20Vs%20Sistem%20Pendukung%20Keputusan.&f=false>. (26 Maret 2019).

Sukamto AR , Shalahuddin M. 2013. Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek. Bandung: Informatika Bandung.

Amin RK Dkk.2015. Implementasi Klasifikasi Decision Tree Dengan Algoritma C4.5 Dalam Pengambilan Keputusan Permohonan Kredit Oleh Debitur (Studi Kasus: Bank Pasar Daerah Istimewa Yogyakarta). Jurnal Tugas Akhir. Diambil dari:
<https://openlibrary.telkomuniversity.ac.id/pustaka/100441/implementasi-klasifikasi-decision-tree-dengan-algoritma-c4-5-dalam-pengambilan-keputusan-permohonan-kredit-oleh-debitur-studi-kasus-bank-pasar-daerah-istimewa-yogyakarta-.html> (16 januari 2020)

Brian SV Dkk.2011. *Confusion Matrix-Based Feature Selection. Midwest Artificial Intelligence And Cognitive Science Conference*. Vol 7.

DAFTAR RIWAT HIDUP

I. Biodata Mahasiswa

NPM : 2016320010.
Nama Lengkap : Dicki Rizki Amarullah.
Tempat & Tanggal Lahir : Sumedang, 01 Juni 1998.
Alamat Lengkap : Perumahan Villa Bekasi Indah 2 Blok C7/27 Tambun Selatan Bekasi

II. Pendidikan

a. Formal

1. SDN Sumber Jaya 06 di Bekasi, lulus tahun 2010.
2. SMPI Al Munir di Bekasi, lulus tahun 2013.
3. SMKS Telkom Telesandi di Bekasi, lulus tahun 2016.

b. Tidak Formal

III. Riwayat Pengalaman Berorganisasi / Pekerjaan

- a. Pengurus HIMA STMIK Bina Insani Periode 2016-2017.
- b. Ketua HIMA STMIK Bina Insani Periode 2017-2018.
- c. Dewan Pengawas Mahasiswa Periode 2018-2019.

IV. Kemampuan

- a. Microsoft Office.
- b. Vb.net.
- c. HTML
- d. Bahasa Pemrograman PHP.
- e. Bahasa Pemrograman Javascript.
- f. Bahasa Pemrograman Jquery.



Bekasi, 15 Mei 2020

Dicki Rizki Amarullah