

**ANALISA ALGORITMA C.45 UNTUK PENYEDIAAN
LAPANGAN PEKERJAAN DI KELURAHAN SUKARAMAI II**

SKRIPSI

OLEH:

WIDYA RIZKI UTAMI

178160050



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MEDAN AREA

MEDAN

2022

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 23/12/22

Access From (repository.uma.ac.id)23/12/22

**ANALISA ALGORITMA C.45 UNTUK PENYEDIAAN
LAPANGAN PEKERJAAN DI KELURAHAN SUKARAMAI II**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana (S1) di Fakultas Teknik
Universitas Medan Area



OLEH:

WIDYA RIZKI UTAMI

178160050

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MEDAN AREA

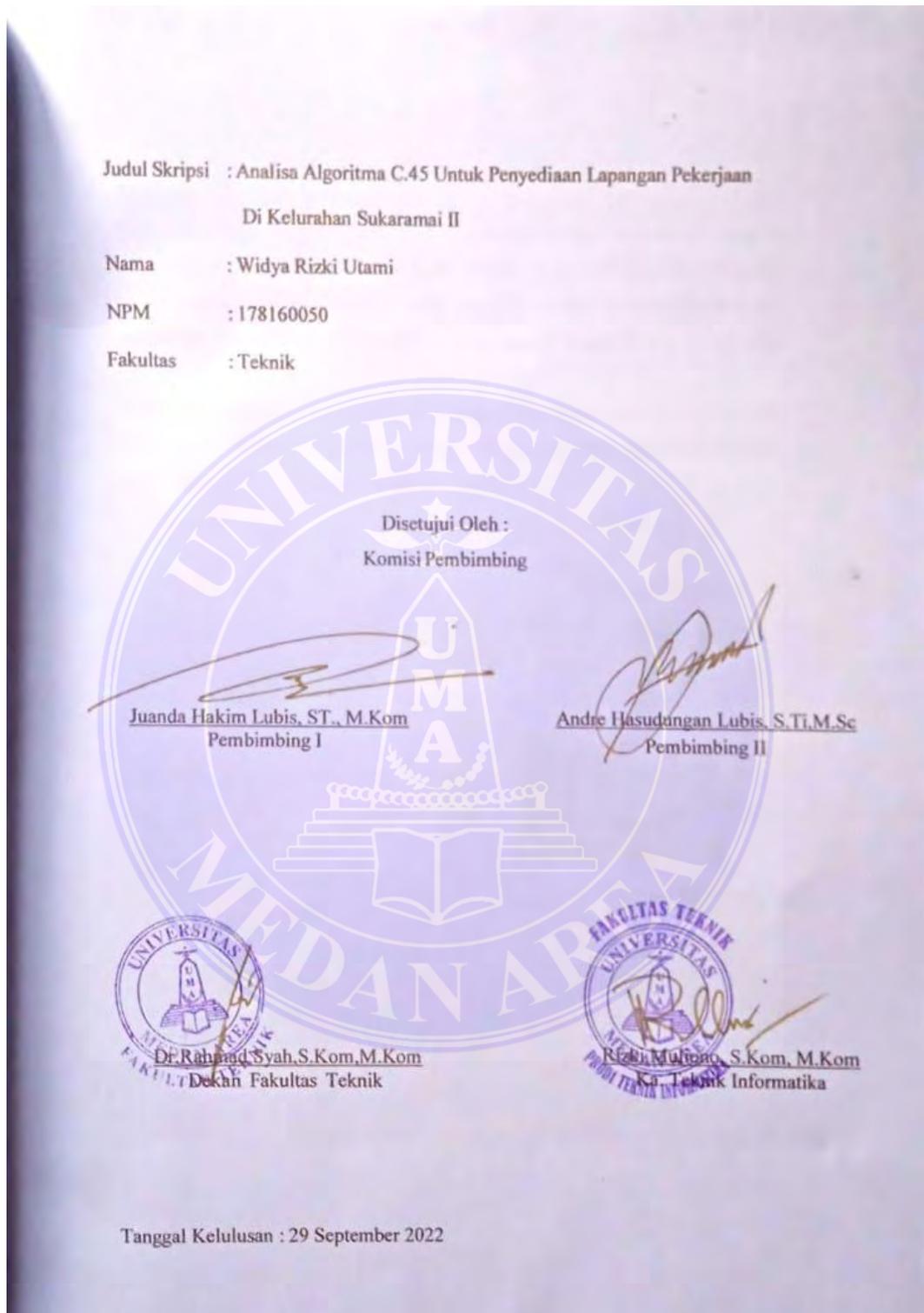
MEDAN

2022

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area



HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai kondisi memperoleh gelar sarjana ialah yang akan terjadi karya tulis sendiri. Adapun bagian-bagian tertentu pada penulisan skripsi ini yang saya kutip berasal dari karya orang lain yang telah dituliskan sumbernya secara kentara sesuai dengan norma, kaidah serta etika penulisan ilmiah.

Saya bersedia mendapatkan sanksi pencabutan gelar akademik yang saya peroleh dan sanksi hukuman lainnya dengan peraturan yang berlaku, apabila dikemudian hari ditemukan adanya plagiat dalam skripsi ini.

Medan, 29 September 2022

Hormat Saya,
Penulis



Widya Rizki Utami
178160050

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini:

Nama : Widya Rizki Utami
NPM : 178160050
Program Studi : Informatika
Fakultas : Teknik
Jenis Karya : Tugas Akhir

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area Hak Bebas Royalti Noneklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah yang berjudul :

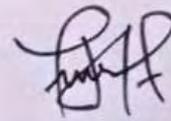
Analisa Algoritma C.45 Untuk Penyediaan Lapangan Pekerjaan Di Kelurahan Sukaramai II

Beserta dengan perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-eksklusif ini Universitas Medan Area berhak menyimpan, mengalihkan media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas/akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Medan

Pada Tanggal : 29 September 2022

Yang menyatakan



(Widya Rizki Utami)

RIWAYAT HIDUP

Widya Rizki Utami, lahir di Medan pada tanggal 20 Maret 2000 dari pasangan orang tua ayah Darwinsyah dan ibu Gustina. Penulis merupakan putri pertama dari satu bersaudara. Penulis memiliki satu saudara kandung yang bernama Muhammad Aditya Gemilang.

Rincian riwayat hidup yang pernah ditempuh oleh penulis sebagai berikut :

1. Tahun 2005 masuk SD Tamansiswa di Sei Rengas Permata, Kecamatan Medan Area, Kota Medan, Sumatera Utara dan tamat (lulus) pada tahun 2011.
2. Tahun 2011 masuk SMP Negeri 13 Medan di Pandau Hulu II, Kecamatan Medan Area, Kota Medan, Sumatera Utara dan tamat (lulus) pada tahun 2014.
3. Tahun 2014 masuk SMK Negeri 1 Medan jurusan Akuntansi di Pusat PASar, Kecamatan Medan Kota, Kota Medan, Sumatera Utara dan tamat (lulus) pada tahun 2017.
4. Tahun 2017 melanjutkan kuliah Program Sarjana di Universitas Medan Area Fakultas Teknik Program Studi Informatika di Jalan Kolam Nomor 1 Medan Estate, Kota Medan, Sumatera Utara.

Selama mengikuti perkuliahan penulis melakukan rangkaian kegiatan untuk menambah wawasan penulis, diantaranya :

1. Pada bulan Februari 2019 penulis melaksanakan kegiatan Pengabdian Kepada Masyarakat di SDN 057751 Simp, UPL dan SDN 056004 Basilam.
2. Pada Tahun 2019 penulis melaksanakan magang di PT. Kereta Api Indonesia (Persero) Divisi Regional I Sumatera Utara pada tanggal 01 Agustus – 30 Agustus 2019.
3. Pada tahun 2020 penulis melaksanakan Kerja Praktek di Kantor Kelurahan Sukaramai II, Kecamatan Medan Area, Kota Medan, Sumatera Utara pada tanggal 10 Agustus – 10 September 2020.

ABSTRAK

Pengangguran adalah seseorang yang tidak bekerja sama sekali atau seseorang yang sedang mencari sebuah pekerjaan atau bisa juga dikatakan sebagai seseorang yang dalam masa pemecatan dalam dua hari selama seminggu dan berusaha untuk memperoleh sebuah pekerjaan. Penyebab terjadinya peningkatan pengangguran ialah dikarenakan kurangnya ketersediaan lapangan pekerjaan, banyaknya jumlah tenaga kerja yang tidak sebanding dengan lapangan pekerjaan dan juga kurangnya keahlian yang dimiliki oleh para pencari kerja, termasuk pencari kerja yang terdidik. Kelurahan Sukaramai II yaitu memiliki jumlah penduduk sebesar lebih dari 6000 jiwa. Dimana hal ini tidak sebanding dengan tersedianya lapangan pekerjaan yang ada di Medan. Dengan jumlah penduduk yang begitu banyak mengakibatkan sulitnya untuk mengklasifikasikan penduduk yang sesuai untuk diberikan pekerjaan secara akurat dan efisien. Penelitian ini membangun suatu sistem yang dapat membantu pemerintah untuk penyediaan lapangan pekerjaan bagi penduduk. Adapun klasifikasi dalam penelitian ini akan dilakukan menggunakan beberapa klasifikasi yaitu : status pernikahan, luas tanah yang dimiliki, penghasilan dan umur dengan metode algoritma c.45. Penelitian ini menggunakan 3 aturan (*rule*) sebagai acuan dalam mengklasifikasi data yang berjumlah sebanyak 378 data sebagai sampel. Hasil pengujian akurasi menunjukkan sebanyak 100%, pengujian kedua 100%, dan pengujian ketiga 100%, atau dapat dikatakan rasio prediksi dari keseluruhan data bernilai positif. Dalam pengujian ini dapat menentukan penduduk tersebut tersedia atau tidak tersedia nya pekerjaan yang akan diberikan. Dengan adanya sistem ini maka menentukan penyediaan lapangan pekerjaan dapat dilakukan dengan cepat dan meminimalisirkan waktu serta data-datanya tersimpan dengan baik di *database*.

Kata Kunci : Penduduk, Pengangguran, Lapangan Pekerjaan, Klasifikasi, Algoritma C.45

ABSTRACT

Unemployment is an individual circumstance who does not work at all or who is looking for a job it can also be said as who is being fired two days or a week and looking for a new occupation. This is due to the inadequate employment opportunities, and also the lack of expertise possessed by job seekers, even though they have a sufficient educational level. The urban village of Sukaramai II has a population of more than 6000 inhabitants, which is not comparable to the availability of jobs in Medan. A large population causes difficulties to classify the appropriate population to be given a job accurately and efficiently. The research built a system that can assist the government to determine the provision of jobs for the population. The classification in this study will be carried out using several classifications, namely: marital status, a land area owned, income and age by using the C.45 algorithm. The study used 3 rules as a reference in classifying data, which amounted to 378 data as samples. The results of the accuracy test show value of first test 100%, second test 100%, and third test 100% or it can be said that the correct prediction ratio of the overall data is positive. The test allows determining whether the population is available or not for the job to be given. The system proposes the capability to determine the provision of jobs that can be done quickly and minimize time and the data is stored properly in the database.

Key Word: Population, Unemployment, Jobs, Classification, C.45 Algorithm

KATA PENGANTAR

Syukur alhamdulillah penulis ucapkan atas kehadiran Allah SWT yang mana telah memberikan rahmat dan karunia – Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul “Analisa Algoritma C4.5 Untuk Penyediaan Lapangan Pekerjaan Di Kelurahan Sukaramai II” ini dengan sebaik–baiknya. Laporan tugas akhir/skripsi disusun berdasarkan hasil pengamatan pada lapangan.

Penyusunan laporan tugas akhir/skripsi ini merupakan syarat yang harus di tempuh untuk memenuhi kelulusan dalam menempuh Gelar Sarjana Jenjang Strata (S-1) sesuai dengan kurikulum Jurusan Teknik Informatika Universitas Medan Area yang berlaku pada saat ini.

Penyusunan laporan tugas akhir/skripsi ini tidak akan selesai tanpa bimbingan, nasehat serta petunjuk dari berbagai pihak. Untuk itu, saya sebagai penulis menyampaikan ucapan terima kasih yang sebesar – besarnya kepada :

1. Kedua Orang Tua, adik saya dan Almh. Bukde Darwati yang saya sayangi. Terima kasih atas segala curahan kasih sayang melalui perhatian, doa, dukungan serta pengorbanan yang telah diberikan selama ini, sebagai motivasi utama bagi penulis untuk dapat terus berusaha menjadi yang terbaik.
2. Bapak Prof. Dr. Dadan Ramdan, M.Eng, M.Sc. selaku Rektor Universitas Medan Area.
3. Bapak Dr. Rahmad Syah, S.Kom, M.Kom selaku Dekan Fakultas Teknik Universitas Medan Area.
4. Bapak Rizki Muliono, S.Kom, M.Kom selaku Kepala Program Studi Teknik Informatika.
5. Bapak Juanda Hakim Lubis, ST., M.Kom dan Bapak Andre Hasudungan Lubis, S.T, M.Sc. selaku Komisi Pembimbing tugas akhir yang dengan sabar telah membimbing saya serta memberikan masukan-masukan yang berguna bagi saya.

6. Seluruh dosen dan staff Universitas Medan Area khususnya dosen prodi Teknik Informatika yang telah membagi ilmunya kepada penulis selama perkuliahan.
7. Sahabat-sahabat saya yaitu Nurani El Furqani, Dinda Rizky Aprillya, Eka Pirdia Wanti S.Kom, Diah Ayu Larasati S.Kom, Ayu Pariyandani S.Kom serta teman-teman angkatan 2017 terutama Teknik Informatika. Terima kasih selalu mendukung serta memeberikan bantuan selama penyusunan skripsi ini.

Penulis menyadari bahwa laporan tugas akhir/skripsi ini masih memiliki kekurangan, oleh karena itu kritik dan saran yang bersifat membangun sangat penulis harapkan demi kesempurnaan laporan tugas akhir/skripsi ini. Penulis berharap laporan tugas akhir/skripsi ini dapat bermanfaat baik untuk kalangan pendidikan maupun masyarakat.

Akhir kata penulis sangat mengharapkan laporan tugas akhir/skripsi ini dapat berguna bagi para pembacanya, dapat menjadi sumber informasi, menambah wawasan khususnya bagi semua pihak.

Medan, 29 September 2022

Hormat Saya,
Penulis

Widya Rizki Utami

178160050

DAFTAR ISI

HALAMAN PERNYATAAN	iv
RIWAYAT HIDUP.....	vi
ABSTRAK	vii
ABSTRACT.....	viii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN.....	xviii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	5
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian.....	5
1.5. Manfaat Penelitian.....	5
1.6. Metodologi Penelitian	6
1.7. Sistematika Penulisan.....	7
BAB II LANDASAN TEORI.....	9
2.1. Data Mining.....	9
2.1.1. Tahapan Data Mining	9
2.1.2. Aktivitas Data Mining	10
2.2. Klasifikasi.....	13
2.3. Algoritma C.45.....	13
2.3.1. Perhitungan Algoritma C.45.....	13
2.4. Penduduk.....	14
2.5. DFD (<i>Data Flow Diagram</i>).....	15
2.5.1. Levelisasi DFD.....	17
2.6. Data Model.....	17
2.6.1. <i>Entity Relationship Diagram</i> (ERD).....	18

2.7. Penelitian Terdahulu	20
BAB III DESAIN DAN PERANCANGAN SISTEM.....	24
3.1. Analisis Sistem.....	24
3.1.1. Deskripsi Sistem.....	24
3.1.1.1. Sistem Yang Berjalan	24
3.1.1.2. Sistem Yang Diusulkan	24
3.1.2. Analisis Kebutuhan Sistem	25
3.1.3. Analisis User	25
3.1.4. Analisis Kebutuhan Fungsional	26
3.1.5. Batasan-Batasan Sistem	27
3.2. Desain Sistem.....	27
3.2.1. Diagram Konteks.....	27
3.2.2. <i>Data Flow Diagram</i>	28
3.2.3. Kamus Data	29
3.2.4. Spesifikasi Proses (<i>Process Spesification</i>).....	31
3.3. Desain Basis Data.....	32
3.3.1. <i>Diagram Entry Relationship</i> (ERD).....	32
3.3.2. Struktur Tabel.....	33
3.4. Contoh Perhitungan Algoritma C.45.....	34
3.4.1. Pilih Atribut Akar	35
3.4.2. Cabang Untuk Setiap Nilai.....	39
3.4.3. Pembagian Setiap Kasus Cabang	40
3.4.4. Pembagian Setiap Kasus Cabang Kedua.....	44
3.4.5. Pembagian Setiap Kasus Cabang Ketiga	46
3.4.6. Pohon Keputusan.....	50
3.4.7. <i>General Rule</i>	50
BAB IV HASIL DAN PEMBAHASAN	53
4.1. Hasil	53
4.1.1. Data Kasus.....	53
4.1.2. <i>Generate Rule</i> 100 Data dan Pengujian	54
4.1.3. <i>Generate Rule</i> 200 Data dan Pengujian	63

4.1.4. <i>Generate Rule 378 Data dan Pengujian</i>	71
4.2. <i>Pembahasan</i>	80
BAB V KESIMPULAN DAN SARAN.....	82
5. 1. <i>Kesimpulan</i>	82
5. 2. <i>Saran</i>	82
DAFTAR PUSTAKA	83
LAMPIRAN	85



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	20
Table 3.1 Tabel Data Kasus	33
Tabel 3.2 Tabel Data Mining	33
Tabel 3.3 Tabel Pohon Keputusan	33
Tabel 3.4 Tabel Operator	34
Tabel 3.5 Tabel Data Penduduk	34
Tabel 3.6 Tabel Perhitungan <i>Node</i> 1	39
Tabel 3.7 Tabel Perhitungan <i>Node</i> 1.1	43
Tabel 3.8 Tabel Perhitungan <i>Node</i> 1.1.1	45
Tabel 3.9 Tabel Perhitungan <i>Node</i> 1.2	49
Tabel 4.1 Data Penduduk	53
Tabel 4.2 Perhitungan <i>Node</i> 1 (100 data)	54
Tabel 4.3 Perhitungan <i>Node</i> 1.1 (100 data)	55
Tabel 4.4 Perhitungan <i>Node</i> 1.1.1 (100 data)	56
Tabel 4.5 Perhitungan <i>Node</i> 1.1.1.1 (100 data)	57
Tabel 4.6 Perhitungan <i>Node</i> 1.2 (100 data)	58
Tabel 4.7 Perhitungan <i>Node</i> 1.2.1 (100 data)	59
Tabel 4.9 Perhitungan <i>Node</i> 1 (200 data)	63
Tabel 4.10 Perhitungan <i>Node</i> 1.1 (200 data).....	64
Tabel 4.11 Perhitungan <i>Node</i> 1.1 (200 data).....	65
Tabel 4.12 Perhitungan <i>Node</i> 1.1.1.1 (200 data).....	66
Tabel 4.12 Perhitungan <i>Node</i> 1.1.2 (200 data).....	67
Tabel 4.13 Perhitungan <i>Node</i> 1.1.2.1 (200 data).....	68
Tabel 4.14. <i>Confusion Matrix</i> 200 Data	70
Tabel 4.15 Perhitungan <i>Node</i> 1 (378 data).....	72
Tabel 4.16 Perhitungan <i>Node</i> 1.1 (378 data).....	73
Tabel 4.17 Perhitungan <i>Node</i> 1.1.1 (378 data).....	73
Tabel 4.18 Perhitungan <i>Node</i> 1.1.1.1 (378 data).....	74
Tabel 4.19 Perhitungan <i>Node</i> 1.1.2 (378 data).....	75

Tabel 4.20 Perhitungan <i>Node</i> 1.1.2.1 (378 data).....	76
Tabel 4.21 <i>Confusion Matrix</i> 378 Data.....	79
Tabel 4.22 Akurasi	80
Tabel 4.23 Waktu Proses Pengujian.....	80



DAFTAR GAMBAR

Gambar 2.1. Bidang Ilmu Data Mining	11
Gambar 2.2. <i>Data Flow Diagram</i>	16
Gambar 2.3. <i>Entity Relationship Diagram</i>	18
Gambar 3.1. Diagram Konteks.....	27
Gambar 3.2. <i>Data Flow Diagram</i> (DFD) Level 1	28
Gambar 3.3. <i>Data Flow Diagram</i> Level 2 Proses <i>Training</i>	29
Gambar 3.6. <i>Entity Relationship Diagram</i> i(ERD).....	32
Gambar 3.7 Pohon Keputusan Hasil Perhitungan <i>Node</i> 1	39
Gambar 3.8 Pohon Keputusan <i>Node</i> 1.1	43
Gambar 3.9 Pohon Keputusan <i>Node</i> 1.1.1	46
Gambar 3.10 Pohon Keputusan <i>Node</i> 1.2	49
Gambar 3.11 Pohon Keputusan Cabang Akhir	50
Gambar 4.1 Pohon Keputusan <i>Node</i> 1 (100 data).....	55
Gambar 4.2 Pohon Keputusan <i>Node</i> 1.1	55
Gambar 4.3 Pohon Keputusan <i>Node</i> 1.1.1	56
Gambar 4.4 Pohon Keputusan <i>Node</i> 1.1.1	57
Gambar 4.5 Pohon Keputusan <i>Node</i> 1.2	58
Gambar 4.6 Pohon Keputusan <i>Node</i> 1.2.1	59
Gambar 4.7 Pengujian 20 data	61
Gambar 4.8 Pohon Keputusan <i>Node</i> 1 (200 data).....	64
Gambar 4.9 Pohon Keputusan <i>Node</i> 1.1 (200 Data).....	64
Gambar 4.10 Pohon Keputusan <i>Node</i> 1.1.1 (200 Data).....	65
Gambar 4.11 Pohon Keputusan <i>Node</i> 1.1.1.1 (200 Data).....	66
Gambar 4.12 Pohon Keputusan <i>Node</i> 1.1.2 (200 Data).....	67
Gambar 4.13 Pohon Keputusan <i>Node</i> 1.1.2.1 (200 Data).....	68
Gambar 4.14 Pengujian 200 data	70
Gambar 4.15 Pohon Keputusan <i>Node</i> 1 (378 Data).....	72
Gambar 4.16 Pohon Keputusan <i>Node</i> 1.1 (378 Data).....	73
Gambar 4.17 Pohon Keputusan <i>Node</i> 1.1.1 (378 Data).....	74

Gambar 4.18 Pohon Keputusan <i>Node</i> 1.1.1.1 (378 Data).....	75
Gambar 4.19 Pohon Keputusan <i>Node</i> 1.1.2 (378 Data).....	76
Gambar 4.20 Pohon Keputusan <i>Node</i> 1.1.2.1 (378 Data).....	77
Gambar 4.21 Pengujian 378 data	78
Gambar 4.22 Waktu Proses Pengujian.....	81



DAFTAR LAMPIRAN

Lampiran Kode Program.....	85
Lampiran SK Pembimbing Skripsi.....	97
Lampiran Surat Pengantar Riset.....	98
Lampiran Surat Selesai Riset	99
Lampiran Turnitin	100



BAB I

PENDAHULUAN

1.1. Latar Belakang

Tumbuhnya perekonomian merupakan suatu kenaikan kualitas, kuantitas produk bahkan jasa yang dibuat oleh negara dari tahun ke tahun. Hal ini dapat diukur dengan indikator seperti penghasilan nasional suatu negara, penghasilan per kapita, angkatan kerja yang lebih besar daripada total pengangguran serta kemiskinan yang menurun. Meningkatnya ekonomi juga dapat diartikan sebagai cara penyempurnaan berkelanjutan dalam ekonomi di suatu negara. Ketika aktivitas ekonomi penduduk di beberapa negara berpengaruh langsung di kenaikan produksi barang dan jasa, ekonomi negara itu bisa disebut mengalami peningkatan. Pemerintah dapat melakukan rancangan penerimaan negara dan pembangunan ke depan berdasarkan seberapa besarnya pertumbuhan ekonomi. Pada saat itu juga, para pelaku bisnis dapat menggunakan ukuran pertumbuhan ekonomi untuk merancang suatu strategi pertumbuhan, produk dan sumber daya bisnis mereka (Fajri & Iriani, 2022).

Dilihat dari sisi positifnya tenaga kerja merupakan sebagai sumber daya yang sangat penting dalam pembangunan dan kemajuan perekonomian suatu negara. Tetapi jika dilihat dalam sudut pandang negatifnya maka tenaga kerja juga dapat menjadi beban atau masalah yang justru dapat mempersulit pemecahan permasalahan perekonomian oleh pemerintah. Sebagai akibatnya dari kurangnya pemerintah dalam penyediaan lapangan pekerjaan sebagai dampak dari bertambahnya jumlah penduduk, sehingga tenaga kerja yang ada tidak terserap secara penuh menjadikan konsekuensi yang terciptalah pengangguran (Sholeh, 2017).

Badan Pusat Statistika mencatat sebanyak 9,77 juta orang di Indonesia yang tidak memiliki pekerjaan/pengangguran sehingga ini dapat menurunkan perekonomian di Indonesia. Medan merupakan salah satu kota yang ada di pulau Sumatera yang memiliki tingkat permasalahan yang tinggi, salah satunya yaitu

pengangguran. Medan menempati peringkat kedua pengangguran tertinggi di provinsi Sumatera Utara. Hal tersebut berarti Medan menyumbang 10,74% pengangguran di Sumatera Utara. Dalam hal ini penggunaan tenaga kerja secara maksimal harus dilaksanakan oleh pemerintah, jika pemerintah ingin meningkatkan dalam pembangunan, jika tidak perlahan tapi pasti peningkatan jumlah angkatan kerja yang tidak terserap. Maka bisa menjadi beban dan penghambat dalam perekonomian dan pada akhirnya menjadi masalah pada daerah tersebut (Sholeh, 2017).

Jumlah pengangguran secara nasional ini tentu saja dapat disebabkan dengan jumlah pengangguran di daerah baik itu perkotaan maupun di perdesaan. Ketersediaannya lapangan kerja yang relatif sedikit, sehingga tidak bisa menyerap seluruh para tenaga kerja yang mengalami kenaikan setiap tahunnya seiring dengan meningkatnya jumlah penduduk. Besarnya angka pengangguran tidak hanya menciptakan masalah-masalah di bidang ekonomi saja, bahkan juga diberbagai masalah bidang sosial, seperti kemiskinan dan rawannya kejahatan sosial (Mulyadi, 2016).

Pengangguran ialah seseorang yang tidak bekerja sama sekali atau seseorang yang sedang mengharapkan sebuah pekerjaan atau bisa juga dikatakan sebagai seseorang yang dalam masa pemecatan dalam dua hari selama seminggu dan berusaha untuk memperoleh sebuah pekerjaan (Kalsum,2017). Peningkatan jumlah pengangguran di suatu negara memiliki pengaruh terhadap kenaikan perekonomian di suatu negara. Penyebab terjadinya peningkatan pengangguran ialah dikarenakan kurangnya ketersediaan lapangan pekerjaan, banyaknya jumlah tenaga kerja yang tidak sebanding dengan lapangan pekerjaan dan juga kurangnya keahlian yang dimiliki oleh para pencari kerja, termasuk pencari pekerja yang terdidik. Dengan pemerintah harus segera mencari solusi tingginya tingkat pengangguran demi kemajuan perekonomian negara. Dalam hal ini, ketersediaan tenaga kerja yang tidak sebanding dengan lapangan pekerjaan ada serta kurangnya keahlian yang dimiliki para pencari kerja menjadi aspek yang penting dalam pengurangan jumlah pengangguran yang ada di Sumatera Utara.

Salah satunya yaitu di kelurahan Sukaramai II yang memiliki jumlah penduduk sebesar lebih dari 6000 jiwa. Dimana hal ini tidak sebanding dengan tersedianya lapangan pekerjaan yang ada di Medan. Dengan begitu banyak penduduknya yang sangat sulit mencari kerja. Sehingga banyak dari mereka yang menjadi pengangguran karena kesulitan mencari pendapatan sedangkan mereka harus memenuhi kebutuhan sehari-harinya. Hal ini menjadi penyebab banyaknya pengangguran di kelurahan Sukaramai II. Dimulai dari umur 18 tahun sampai dengan umur 40 tahun ke atas sangat sulit untuk mencari pekerjaan yang layak. Banyak dari mereka yang berjualan untuk mencukupi kehidupan sehari-hari, tetapi hal itu tidak cukup untuk memenuhi kebutuhan lainnya. Dimana pendapatan yang dimiliki tidak sebanding dengan pengeluaran yang dikeluarkan oleh warga kelurahan Sukaramai II. Seperti untuk pembayaran pendidikan anak-anak mereka, membangun rumah yang layak, sulit juga untuk kebutuhan air dan listrik serta sulit juga untuk memenuhi kebutuhan kesehatan mereka. Sehingga kemiskinan ini sudah sangat menjadi permasalahan yang sangat serius untuk ditangani oleh pemerintah. Untuk melakukan penyediaan lapangan pekerjaan bagi mereka yang kurang mampu yang sulit untuk mendapatkan penghasilan bagi mereka.

Sehingga di kelurahan Sukaramai II melakukan program penyediaan lapangan pekerjaan bagi penduduk yang membutuhkan pekerjaan. dengan adanya penyediaan lapangan pekerjaan bagi penduduk di kelurahan Sukaramai II masalah tersebut akan terselesaikan. Tetapi ada permasalahan dalam penentuan penyediaan lapangan pekerjaan bagi penduduk tersebut, yaitu dikarenakan dalam mengklasifikasi siapa saja yang membutuhkan pekerjaan masih sulit dilakukan. Yang disebabkan oleh jumlah penduduk yang begitu banyak sehingga sulit untuk mengetahui siapa saja yang perlu di berikan pekerjaan. Hal itu dikarenakan dalam penentuan penerimaan penyediaan lapangan pekerjaan masih dilakukan secara manual serta belum ada pola yang jelas dalam penentuan penerimaan lapangan pekerjaan bagi penduduk. Tentu hal ini kurang efektif dan akurat dalam menentukan siapa saja yang membutuhkan pekerjaan serta dibutuhkan waktu yang banyak dalam mengklasifikasi penduduk yang jumlahnya sangat banyak.

Algoritma c.45 artinya salah satu solusi pemecahan perkara yang seringkali digunakan dalam pemecahan persoalan pada teknik klasifikasi. Keluaran dari algoritma c.45 itu berupa sebuah *decision tree* layaknya teknik klasifikasi lain. Sebuah pohon keputusan artinya sebuah struktur yang dapat digunakan buat membagi perpaduan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan pembagian, anggota himpunan hasil sebagai mirip satu dengan yang lain (Nofriansyah & Nurcahyo, 2019). Dalam penelitian sebelumnya telah berhasil melakukan klasifikasi data dengan algoritma c.45 untuk kasus keterlambatan pembayaran SPP mahasiswa dengan variabel-variabel yang telah di dapatkan, yang selanjutnya dilakukan analisa algoritma c.45 untuk mencari pola pembayaran SPP mahasiswa yang sudah ada, kemudian dijadikan dasar untuk memprediksi pembayaran pada semester berikutnya.

Algoritma c.45 juga telah berhasil menyelesaikan masalah memprediksi masyarakat dalam menerima bantuan sosial dengan acuan variabel penghasilan, pekerjaan, kepemilikan rumah, status dan kelayakan dalam menerima bantuan sosial, dari hasil penelitiannya ditemukan akurasi yang didapatkan sebanyak 100% dengan demikian algoritma c.45 yang baik dalam memprediksi masyarakat dalam menerima bantuan sosial.

Pada penelitian ini akan dilakukan pengklasifikasian data penduduk serta pembuatan pola dalam penentuan penyediaan lapangan pekerjaan. Dengan melakukan pendataan penduduk dengan menentukan variabel sebagai parameter dalam pengklasifikasian. Adapun variabel yang telah ditentukan yaitu status pernikahan, luas tanah yang dimiliki, penghasilan dan umur. Hasil klasifikasi nantinya akan dibandingkan dengan prediksi awal secara manual untuk dilanjutkan dengan analisa klasifikasi terhadap presentase ketepatan prediksi algoritma c.45 dalam penyelesaian penentuan lapangan pekerjaan di kelurahan Sukaramai II.

1.2. Perumusan Masalah

Berdasarkan pada latar belakang masalah yang telah dijabarkan diatas, adapun rumusan masalah pada penelitian tugas akhir ini yaitu mencari seberapa akurat algoritma c.45 dalam mengklasifikasi data penduduk dalam penentuan penyediaan lapangan pekerjaan bagi penduduk di kelurahan Sukaramai II.

1.3. Batasan Masalah

Adapun batasan masalah dalam penelitian ini ialah sebagai berikut :

1. Data yang digunakan dalam penelitian adalah data sekunder dari kantor kelurahan Sukaramai II. Dengan jumlah populasi sebanyak 6844 penduduk dan diambil sampel sebanyak 378 data penduduk.
2. Menerapkan metode klasifikasi data mining dengan algoritma c.45 untuk klasifikasi data penduduk di kelurahan Sukaramai II.
3. Ada 4 klasifikasi yang digunakan untuk menentukan tersedianya lapangan pekerjaan bagi masyarakat yaitu status perkawinan, luas tanah yang dimiliki, penghasilan, dan umur.
4. Bahasa pemrograman yang akan digunakan ialah bahasa pemrograman *PHP*.

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini merupakan untuk mencari akurasi algoritma c.45 data penduduk penentuan penyediaan lapangan pekerjaan bagi penduduk kelurahan Sukaramai II.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Dapat memberikan informasi atau hasil yang tepat, cepat dan akurat bagi pemerintah untuk menentukan penduduk yang sangat membutuhkan pekerjaan. Sehingga penyediaan lapangan pekerjaan didapatkan secara merata dan tepat kepada penduduk Sukaramai II.
2. Membantu kepala daerah khususnya untuk memilih dan menentukan penduduk yang akan diberikan pekerjaan. Sehingga dapat memberikan

lapangan pekerjaan yang layak kepada masyarakat di kelurahan Sukaramai II serta dapat meningkatkan perekonomian dan penurunan angka pengangguran di kelurahan Sukaramai II.

1.6. Metodologi Penelitian

Adapun metodologi penyelesaian masalah yang dilakukan dalam membuat analisa algoritma c.45 untuk penyediaan lapangan pekerjaan di kelurahan Sukaramai II yaitu sebagai berikut :

1. Tahap Pengumpulan Data

Dalam mengumpulkan data penelitian, penulis menggunakan teknik pengumpulan data Studi lapangan, yaitu pengumpulan data atau informasi melalui aktivitas penelitian dengan turun langsung ke lokasi penelitian buat mencari keterangan fakta yang berkaitan dengan masalah yang diteliti.

Adapun studi lapangan pada penelitian ini terdiri dari:

a. Studi Literatur

Pada tahap ini saya melakukan pengumpulan data dengan mengumpulkan data-data dari berbagai jurnal yang ada, buku-buku dan literatur lainnya.

b. Observasi

Pada tahap ini saya mengumpulkan data-data dalam penulisan tugas akhir saya dengan melakukan observasi langsung ke kantor kelurahan Sukaramai II.

2. Metode Pengembangan Perangkat Lunak

a. Analisa Kebutuhan Perangkat Lunak

Di proses ini dilakukan analisa dan pengumpulan kebutuhan yaitu informasi perihal data-data kriteria melalui *e-book* serta jurnal ataupun informasi lainnya.

b. Desain Sistem

Desain sistem ini dirancang dengan permodelan diagram konteks, dfd, dan erd yang digunakan buat perancangan desain sistem. Proses

ini nantinya akan fokus pada struktur data, arsitektur aplikasi, representasi *interface*, dan detail algoritma prosedural.

c. Penulisan Kode Program

Penulisan kode program memakai bahasa pemrograman *PHP*. Hal ini akan memudahkan proses pada saat perancangan kode program. Selanjutnya pengkodean selesai maka akan melakukan *testing* terhadap sistem yang telah dibuat tersebut. Tujuan *testing* ialah menemukan kesalahan-kesalahan terhadap sistem tadi serta kemudian bisa diperbaiki.

d. Pengujian Program

Berisi beberapa tahap yang dilakukan pada pembuatan sistem dan tahapan-tahapan pengujian yang dilakukan buat masing-masing blok sistem yang dirancang untuk :

1. Analisis apakah sistem aplikasi ini sudah sesuai dengan aturan.
2. Melakukan tes pada sistem aplikasi.
3. Melakukan pemeliharaan pada sistem aplikasi.

1.7. Sistematika Penulisan

Sistematika penulisan skripsi ialah urutan berfikir yang menjelaskan proses penulisan skripsi, untuk mempermudah mencari laporan penulisan perlu adanya sistematika penulisan. Sehingga pembaca mengetahui dari awal tentang permasalahan yang diteliti hingga penutup. Agar mempermudah pembaca dalam memahami isi dari penelitian ini, penulis membagi sistematika penulisan dalam lima bab yang terdiri dari beberapa bab, masing-masing penjelasan yang secara garis besar dapat dijelaskan sebagai berikut :

1) BAB I PENDAHULUAN

Dalam bab ini merupakan pendahuluan yang materinya sebagian besar menyempurnakan usulan penelitian yang berisikan tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, hipotesis penelitian, manfaat penelitian, dan sistematika penelitian.

2) BAB II LANDASAN TEORI

Pada bab ini membahas mengenai studi literatur, prinsip kerja, dan komponen-komponen yang ada.

3) BAB III ANALISIS PERANCANGAN SISTEM

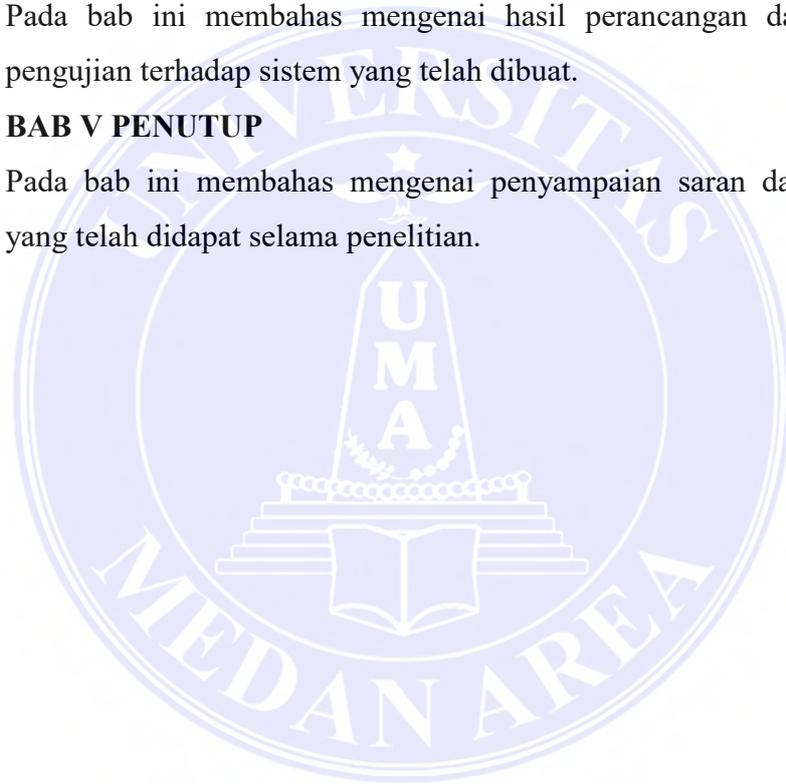
Pada bab ini membahas mengenai menganalisis sistem yang digunakan dan menganalisa sistem yang akan dibuat, mendesain sistem yang akan dibuat nantinya, dan mendesain basis data pada sistem yang akan digunakan.

4) BAB IV HASIL DAN ANALISA

Pada bab ini membahas mengenai hasil perancangan dan melakukan pengujian terhadap sistem yang telah dibuat.

5) BAB V PENUTUP

Pada bab ini membahas mengenai penyampaian saran dan kesimpulan yang telah didapat selama penelitian.



BAB II

LANDASAN TEORI

2.1. Data Mining

Data mining merupakan istilah yang dimanfaatkan buat berbagi inovasi ilmu di dalam *database*. Data mining artinya sebuah proses penggalian data yang belum diketahui sebelumnya tetapi dapat dimengerti serta bermanfaat dari *database* yang besar dan dimanfaatkan sebagai suatu keputusan bisnis yang sangat penting. Dengan kata lain bahwasanya data mining dimanfaatkan buat penggalian dari info penting yang tersembunyi dari *database* yang besar (Jollyta, Ramdhan, & Zarlis, 2020).

2.1.1. Tahapan Data Mining

Data Mining ialah bagian dari tahapan proses *Knowledge Discovery in Database* (KDD). Cara ini merupakan cara untuk menghasilkan pengetahuan dari *database*. Pada *database* ada tabel-tabel yang saling berkaitan, hasil pengetahuan yang didapatkan tadi bisa dimanfaatkan menjadi basis pengetahuan (*knowledge base*) buat kepentingan pengambilan keputusan. Proses KDD secara garis besar dapat dijabarkan sebagai berikut :

- 1 *Data Selection*

Pemilihan (seleksi) data ialah sejumlah data operasional harus dilaksanakan sebelum tahap ekstrasi informasi pada *Knowledge Discovery in Database* (KDD) dimulai. Data hasil seleksi yang akan dimanfaatkan untuk mengolah data mining, yang di *save* dalam sebuah berkas tersendiri.

- 2 *Pre-processing/Cleaning*

Sebelum pengolahan data mining dilakukan, harus diadakan proses *cleaning* pada data yang sebagai fokus *Knowledge Discovery in Database* (KDD). Proses *cleaning* melingkupi antara lain menghapus duplikasi data, memeriksa data yang inkosisten, serta memperbaiki kesalahan di data, seperti kesalahan cetak. Perlu dijalankan proses *enrichment*, yaitu proses “memperkaya” data yang sudah terdapat dengan data atau info lain yang

relevan dan dibutuhkan buat *Knowledge Discovery in Database* (KDD), seperti data atau info eksternal lainnya yang diharapkan.

3 *Transformation*

Coding merupakan proses transformasi pada data yang telah diputuskan, sehingga data itu sesuai buat proses data mining. Proses *ending* dalam *Knowledge Discovery in Database* (KDD) merupakan proses kreatif dan sangat bergantung pada tipe atau skema informasi yang akan dicari dalam basis data.

4 Data Mining

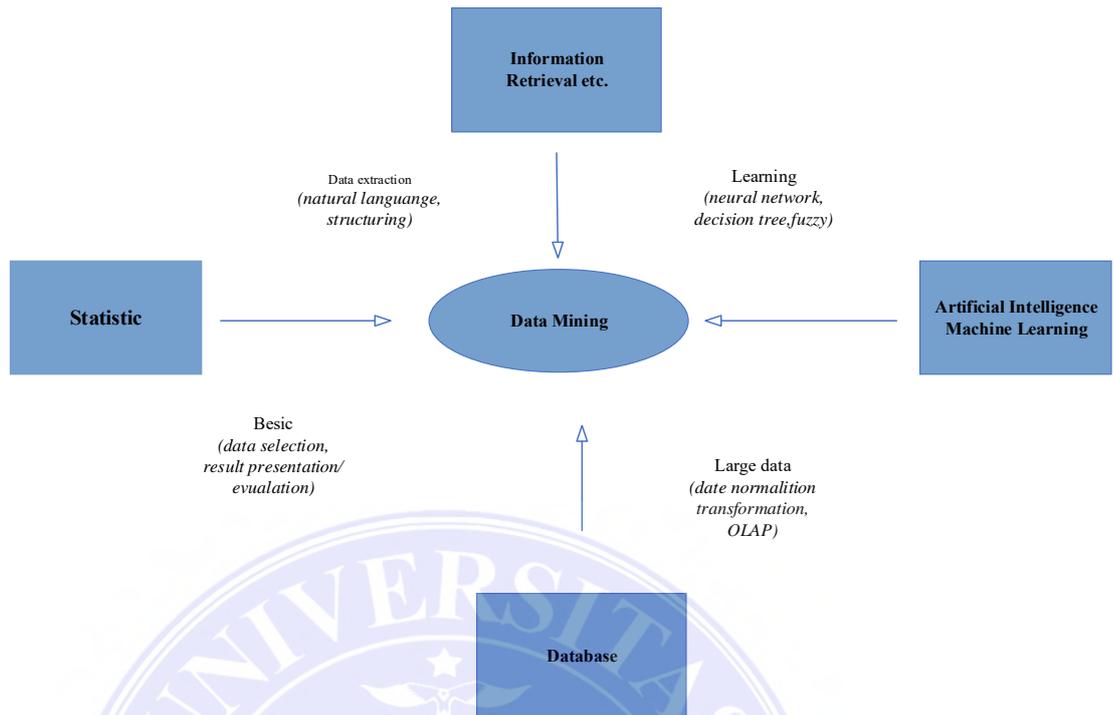
Data mining ialah proses mencari skema atau informasi menarik dalam data pilihan yang menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat beragam. Pemilihan atau metode atau algoritma yang akurat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) semuanya.

5 *Interpretation/ Evaluation*

Skema informasi yang dihasilkan pada proses data mining perlu digambarkan pada bentuk yang mudah dipahami oleh pihak yang berkepentingan. Tahap ini adalah bagian dari proses *Knowledge Discovery in Database* (KDD) yang dianggap *interpretation*. Tahap ini mencakup investigasi apakah pola atau info yang ditemukan bertentangan dengan informasi atau hipotesis yang ada sebelumnya.

2.1.2. **Aktivitas Data Mining**

Data mining dibagi dalam beberapa kelompok berlandaskan tugas yang bisa dilakukan, yaitu :



Gambar 2.1. Bidang Ilmu Data Mining (Mardi, 2017)

1 *Description* (Deskripsi)

Terkadang peneliti serta analis secara sederhana ingin mencoba serta mencari cara untuk mendeskripsikan pola dan kecenderungan yang ada pada data. Deskripsi dari pola dan kesamaan seringkali memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2 *Estimation* (Estimasi)

Estimasi nyaris sama seperti klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kelompok. Model diciptakan menggunakan *record* lengkap yang menyajikan nilai dari variabel target sebagai nilai prediksi. Berikutnya, pada peninjauan selanjutnya estimasi nilai dari variabel target dibuat berlandaskan nilai variabel prediksi.

3 *Prediction* (prediksi)

Prediksi hampir sama dengan klasifikasi serta perkiraan, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Contoh prediksi harga sembako pada 2 bulan yang akan tiba. Beberapa metode

dan teknik yang digunakan pada klasifikasi dan perkiraan dapat pula dipergunakan (untuk keadaan yang sempurna) untuk prediksi.

4 *Classification* (klasifikasi)

Dalam klasifikasi, terdapat target variabel kategori. Sebagai misalnya yaitu menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang. Atau contoh lainnya mendiagnosis penyakit seseorang pasien termasuk golongan penyakit apa. Dengan kata lain klasifikasi merupakan proses pengumpulan data beserta sama yang didasarkan atas sekumpulan kecenderungan yang awalnya sudah dipengaruhi oleh seorang analis sebelum di analisis dimulai. Teknik ini menilik data yang telah diklasifikasikan serta dikumpulkan pada kelompok beserta sesuai dengan keanggotaannya.

5 *Clustering* (Pengklusteran)

Pengklusteran ialah pengelompokkan *record*, pengamatan, atau memperhatikan dan membuat kelas objek-objek yang memiliki kemiripan. Kluster artinya kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record-record* dalam kluster lain. Pengklusteran adalah suatu contoh yang dilakukan dengan cara melakukan proses *segmentasi* terhadap populasi yang tidak sejenis ke dalam sejumlah *cluster homogeny*. Proses *clustering* ini tidak selaras dengan klasifikasi di mana pada clustering artinya tidak diketahui saat di mana prosedur pemecahan dimulai (Jollyta, Ramdhan , & Zarlis, 2020).

6 *Association* (Asosiasi)

Tugas asosiasi dalam data mining merupakan menemukan atribut yang timbul dalam satu waktu. Menurut (Jollyta, Ramdhan, & Zarlis, 2020) asosiasi digunakan buat mengenali tingkah laku dari insiden khusus atau proses di mana link asosiasi muncul di setiap peristiwa. Teknik ini mencoba untuk menyiapkan nilai-nilai yang timbul di saat bersamaan pada setiap barisnya dan menampilkan yang akan terjadi di keluaran yang disimpulkan dalam sebuah *rule*.

2.2. Klasifikasi

Klasifikasi merupakan salah satu teknik dalam data mining. Klasifikasi (taksonomi) ialah proses penempatan objek atau konsep tertentu ke pada satu set kategori berdasarkan objek yang dipergunakan. Salah satu teknik klasifikasi yang populer digunakan ialah *decision tree*.

Klasifikasi sendiri terbagi sebagai 2 tahap, yaitu pengklasifikasian serta pembelajaran. Tahap pembelajaran, sebuah algoritma klasifikasi akan membentuk sebuah contoh klasifikasi dengan cara menganalisis pelatihan data. Tahap pembelajaran bisa juga dilihat menjadi tahap pembentukan fungsi atau pemetaan $y=f(x)$ dimana y artinya kelas hasil prediksi dan x ialah yang ingin diprediksi kelasnya (Ardiansyah & Walim, 2018).

2.3. Algoritma C.45

Algoritma c.45 artinya salah satu solusi pemecahan perkara yang acapkali digunakan dalam pemecahan persoalan pada teknik klasifikasi. Keluaran dari algoritma c.45 itu berupa sebuah *decision tree* layaknya teknik klasifikasi lain. Sebuah pohon keputusan artinya sebuah struktur yang dapat digunakan buat membagi perpaduan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan pembagian, anggota himpunan hasil sebagai mirip satu dengan yang lain (Nofriansyah & Nurcahyo, 2019).

2.3.1. Perhitungan Algoritma C.45

Tahapan dari algoritma c.45 adalah sebagai berikut :

- a) Memilih atribut sebagai akar,
- b) Buat cabang untuk setiap nilai,
- c) Membagi kasus dari tiap cabang,
- d) Ulangi langkah yang sama untuk setiap cabang hingga semua kasus memiliki kelas yang sama. Untuk menghitung nilai *entropy* dapat dihitung dengan persamaan :

$$Entropy(S) = \sum \sum_{i=1}^n -p_i * \log_2 p_i \quad (2.1)$$

dimana :

S = himpunan kasus

A = fitur

n = jumlah partisi S

p = proporsi dari S_i terhadap S

Sementara itu nilai *information gain* (*Gain*) dapat dihitung menggunakan persamaan :

$$Gain(S, A) = entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.2)$$

dimana :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

$|S_i|$ = jumlah kasus pada partisi ke-i

$|S|$ = jumlah kasus dalam S

Hasil perhitungan *entropy* untuk masing-masing variabel kemudian digunakan untuk mencari *gain* tertinggi dan *gain* tertinggi digunakan sebagai akar dari pohon keputusan yang akan dibentuk, kemudian proses akan berulang semua kasus pada cabang memiliki kelas yang sama. Pohon keputusan yang terbentuk kemudian digariskan dalam suatu aturan (Tanjung D.Y.,2021).

2.4. Penduduk

Penduduk itu orang yang berdomisili ataupun yang sedang bertempat tinggal di suatu negara. Definisi penduduk dipandang dari segi pandang penulis ialah suatu serikat sejumlah orang yang menempati suatu negara serta yang terikat dengan sebuah hukum-aturan yang ada pada suatu negara. Dikatakan penduduk Indonesia bila orang tersebut memiliki tanda pengenal misalnya KTP, Kartu Keluarga, Akte Kelahiran yang berdomisili serta berkewarganegaraan Indonesia (Watik, Novitasari, & Trisiana, 2022).

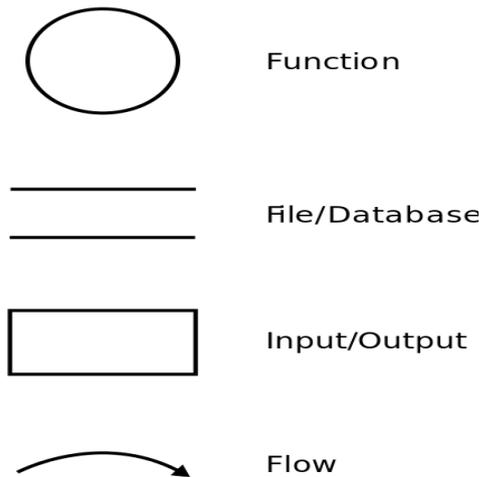
Dalam sebuah keputusan Menteri Dalam Negeri No.54 tentang kependudukan pedoman penyelenggaraan pendaftaran penduduk yang dijelaskan diantaranya : Penduduk adalah Warga Negara Indonesia (WNI) dan Warga Negara Asing (WNA) pemegang ijin tinggal tetap di wilayah Negara Kesatuan Republik Indonesia atau semua orang yang berdomisili di desa tersebut selama enam bulan lebih atau mereka yang berdomisili kurang dari enam bulan tapi bertujuan menetap (Huda, 2018).

Pengangguran artinya angkatan kerja yang sama sekali tidak memiliki pekerjaan. pengangguran ini terjadi sebab banyak berasal angkatan kerja yang belum menerima pekerjaan padahal sudah berusaha secara aporisma atau dikarenakan faktor malas untuk mencari pekerjaan atau malas bekerja (Wahyudi , Pujiastuti, & Solikhun, 2020).

Tingkat pengangguran merupakan indikator dalam hal pertumbuhan ekonomi di suatu negara atau di suatu wilayah. berdasarkan tingkat pengangguran, mampu memilih apakah negara atau pada suatu daerah tersebut tumbuh, stagnan atau menurun. Pengangguran dapat terjadi ketika jumlah orang mencari suatu pekerjaan semakin tinggi dengan cepat namun lapangan pekerjaan yang tersedia tidak sinkron (Fajri & Iriani, 2022).

2.5. DFD (*Data Flow Diagram*)

DFD adalah salah satu komponen dalam serangkaian pembuatan perancangan sebuah sistem komputerisasi. DFD menggambarkan sirkulasi data berasal dari sumber pemberi data (*input*) ke penerima data (*output*). Peredaran data itu perlu diketahui agar si penghasil sistem memahami persis kapan sebuah data harus disimpan, kapan harus ditanggapi (*proses*), dan kapan wajib didistribusikan ke bagian lain.



Gambar 2.2. Data Flow Diagram (Utami & Asnawati, 2015)

Komponen-komponen DFD Komponen-komponen DFD terdiri atas :

1) Terminator

Terminator bisa diklaim juga “kesatuan luar”, yaitu suatu unit kerja/jabatan, atau sejenisnya yang berada pada luar sistem tetapi memberi andil atas pemberian atau penerimaan data dari sistem secara eksklusif. Terminator bisa juga diklaim menggunakan “sumber pemberi data (*input*)”, juga “tujuan pemberian data (*output*)”. Pemberi data serta penerima data yang dimaksud merupakan pihak yang sangat dekat dan memiliki hubungan langsung dengan sistem.

2) Proses

Proses ialah suatu tindakan yang akan diambil terhadap data yang masuk. Karena proses merupakan tindakan, maka proses berisi istilah kerja. Proses diberikan identifikasi (angka) supaya mempermudah sekuen buat diagram detailnya.

3) Alur Data

Alur data mendeskripsikan data yang mengalir dari terminator ke proses lainnya. Data yang dibawa oleh alur data wajib disebutkan dan diletakkan pada lambang alur data. Dan jika alur data digambar panjang, panjang sebaiknya penulisan data mendekati lambang anak panahnya. Data yang

menempati alur data dapat berupa elemen data tunggal, maupun deretan elemen data.

4) Penyimpanan Data (*Data Store*)

Data yang akan disimpan perlu ditempatkan ke satu tempat penyimpanan data. Data yang disimpan dapat berupa data manual maupun data digital. Untuk data digital, penyimpan data tersebut kelak akan dijadikan *file* data di komputer. Alur data yang anak panahnya menuju penyimpan data, kegiatannya adalah “menulis/merekam” data, sebagai akibatnya isi arsip data akan berubah karenanya. Sedangkan alur data yang anak panahnya menuju ke proses dari penyimpanan data, kegiatannya adalah “membaca” data sehingga isi arsip data tidak akan berubah karenanya.

2.5.1. Levelisasi DFD

DFD digambarkan secara bertingkat, dari tingkat yang global berturut-turut hingga taraf yang sangat detail. Taraf yang global (awam) disebut dengan “Diagram Konteks” atau “*Context Diagram*”. Ini termasuk level 0.

Selanjutnya, dari diagram konteks, prosesnya dijabarkan lebih rinci lagi di “Diagram Nol” atau “*Zero Diagram*”. Ini disebut level 1. Pada diagram nol ini yang berkembang hanya proses dan alur data yang menghubungkan proses-prosesnya, sedangkan jumlah terminatornya dan alur data yang masuk atau keluar dari terminator, tetap.

Bila, masih dirasakan perlu memerinci proses berikutnya, maka diagram selanjutnya disebut sebagai “Diagram Detail” atau “Diagram Primitif”. Ini disebut dengan level 2. Pada diagram detail, yang diagram relatif proses (nomor berapa) yang perlu didetailkan saja, selain itu (proses lainnya, atau terminatornya) tidak perlu digambarkan. Bila masih dapat lebih di detailkan lagi, maka level 3, dan seterusnya mampu dirancang (Utami & Asnawati, 2015).

2.6. Data Model

Data model artinya cara formal buat mendeskripsikan data yang dipergunakan serta diciptakan pada suatu bisnis. Model ini memberikan orang, tempat atau benda dimana data diambil dan korelasi antar data tersebut.

Pemodelan data pula dibedakan sebagai dua bagian, yaitu model data logis (*logical data model*) dan model fisik (*physical data model*). Model data logis memberikan pengaturan data tanpa menandakan bagaimana data tadi disimpan, dibuat, dan dimanipulasi. Contoh data fisik membagikan bagaimana data akan disimpan sebenarnya dalam *database* atau arsip. Penyusunan pemodelan data wajib seimbang dengan pemodelan proses. Salah satu cara pemodelan data ialah menggunakan ERD (*Entity Relationship Diagram*).

2.6.1 *Entity Relationship Diagram (ERD)*

ERD merupakan gambar atau diagram yang memberikan informasi yang didesain, disimpan, dan digunakan dalam sistem bisnis, entitas umumnya menggambarkan jenis informasi yang sama. Pada entitas digunakan buat menghubungkan antar entitas yang sekaligus menunjukkan korelasi antar data. Pada akhirnya ERD bisa pula dipergunakan untuk membagikan aturan-aturan bisnis yang ada pada sistem informasi yang akan dibangun (Trisyanto, 2017).

Entity Relationship Diagram (ERD) merupakan pemodelan data primer yang membantu mengorganisasikan data dalam suatu proyek ke dalam entitas-entitas serta menghubungkan antar entitas bersama atributnya. ERD memiliki entitas, relasi, serta atribut. (Setyaningrum, 2013).



Gambar 2.3. *Entity Relationship Diagram* (Setyaningrum, 2013)

a. Entitas (*Entity*)

Entitas artinya sesuatu apa saja yang terdapat pada sistem, nyata ataupun abstrak pada data tersimpan atau di mana ada data. Entitas tersaji pada bentuk persegi panjang. Entitas diberi nama menggunakan kata benda serta dapat dikelompokkan dalam (empat kelas, yaitu : *role* (peran), *events* (peristiwa), *location* (lokasi), *concepts* (sesuatu yang tidak nyata/konsep).

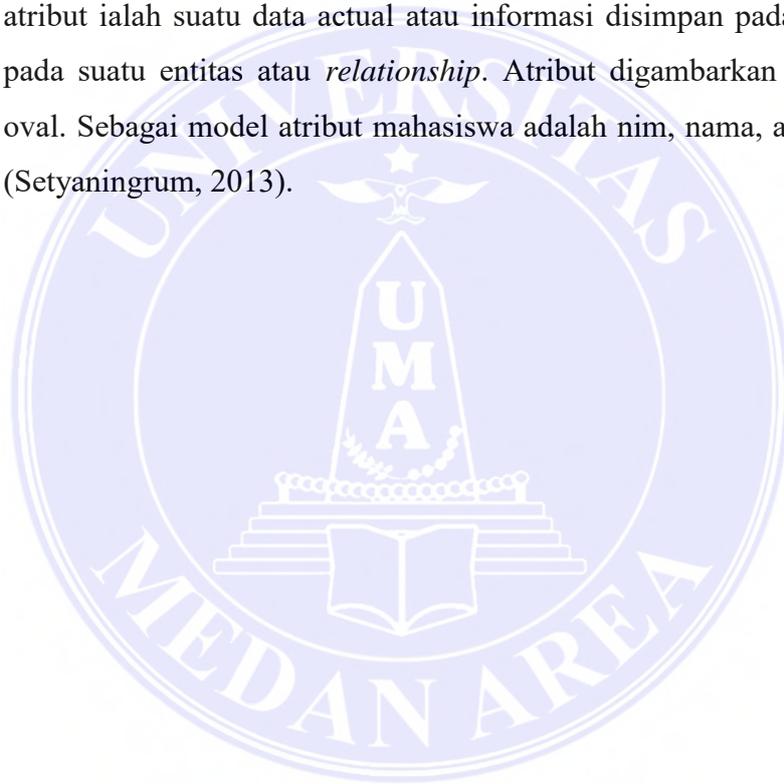
Contoh : pegawai, mahasiswa, perusahaan, kampus, buku, dan lain-lain. Model lebih jelasnya dari suatu entitas disebut *instance*, contohnya mahasiswa bernama Alfa, Beta, Gama.

b. Relasi (*Relationship*)

Relasi adalah penghubung antara suatu entitas dengan entitas yang lain serta merupakan bagian yang sangat krusial dalam mendesain basis data.

c. Atribut

Atribut adalah karakteristik dari entitas atau *relationship*, yang menyediakan penerangan detail tentang entitas atau *relationship* itu. Nilai atribut ialah suatu data actual atau informasi disimpan pada suatu atribut pada suatu entitas atau *relationship*. Atribut digambarkan dalam bentuk oval. Sebagai model atribut mahasiswa adalah nim, nama, alamat, jurusan (Setyaningrum, 2013).



2.7. Penelitian Terdahulu

Berikut ini adalah beberapa penelitian terdahulu tentang analisa algoritma c.45 yang berkaitan dengan metode algoritma c.45 dan pengangguran :

Tabel 2.1 Penelitian Terdahulu

Penulis (Tahun)	Judul	Metode	Hasil
Mochamad Wahyudi, Lise Pujiastuti, Solikhun (2020)	Penerapan Data Mining Dalam Mengelompokkan Data Pengangguran Terbuka Menurut Provinsi Menggunakan Algoritma <i>K-Means</i>	Algoritma <i>K-Means</i>	Hasil akhir dari penelitian yang menggunakan data sebanyak 34 provinsi ini, dapat disimpulkan bahwa telah didapatkan masing-masing nilai <i>cluster</i> yakni : 1. <i>Cluster</i> tinggi (C1) dengan jumlah sebanyak 12 Provinsi yaitu: Aceh, Sumatera Utara, Sumatera Barat, Riau, Kep. Riau, DKI Jakarta, Jawa Barat, Banten, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Maluku, dan Papua Barat. 2. <i>Cluster</i> rendah (C2) dengan jumlah sebanyak 22 Provinsi selain dari <i>cluster</i> tinggi. 3. Proses pemberentian iterasi pada pengujian yang dilakukan pada penelitian ini yaitu terjadi pada iterasi ke 3. 4. Nilai hasil akurasi yang dilakukan dengan perhitungan

			manual dan dengan aplikasi <i>RapidMiner</i> bernilai sama.
Febti Eka Pratiwi, Ismaini Zain (2014)	Klasifikasi Pengangguran Terbuka Menggunakan CART (<i>CLASSIFICATION AND REGRESSION TREE</i>) Di Provinsi SULAWESI UTARA	CART (<i>CLASSIFICATION AND REGRESSION TREE</i>)	Faktor yang mempengaruhi pengangguran terbuka di provinsi Sulawesi Utara dengan menggunakan CART yaitu jenis kelamin, pendidikan terakhir, usia, status dalam rumah tangga, dan status perkawinan. Metode CART pada penelitian ini memiliki ketepatan 78,90%.
Abdi Rahim Damanik , Sarjon Defit, Dedi Hartama,P.P.P.A.N.W Fikrul Ilmi R.H Zer (2020)	Implementasi Metode <i>K-Means</i> Dalam Perbandingan Tingkat Pengangguran Dengan Tenaga Kerja Berdasarkan Provinsi	Algoritma <i>K- Means</i>	Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa metode data mining dengan metode k-means dapat diterapkan dalam perbandingan pengelompokkan data tingkat pengangguran dan tenaga kerja. Dari hasil yang diperoleh dapat dilihat bahwa tingkat pengangguran lebih tinggi dibandingkan dengan tenaga kerja untuk provinsi yang ada di Indonesia.
Tri Herdiawan Apani, Roby Bayu Maulana, Rian Piarna, Dwi Vernanda (2018)	Menganalisis Kemungkinan Keterlambatan Pembayaran SPP	Algoritma C.45	Berdasarkan hasil penelitian yang telah dilakukan maka dapat diambil beberapa kesimpulan, antara lain dari hasil uji coba partisi data set ditemukan tingkat akurasi tertinggi pada data partisi 2 sebanyak 80% data training dan

	<p>Dengan Aalgoritma C.45 (Studi Kasus Politenik TEDC Bandung)</p>		<p>menghasilkan akurasi sebesar 75,00%. Dengan demikian model yang didapatkan dari partisi 2 ini dapat digunakan sebagai data training untuk memprediksi data yang baru. Atribut pelayanan akademik memiliki pengaruh yang cukup besar dalam semua partisi yang diuji, setelah dilakukan percobaan dengan menggunakan <i>RapidMiner</i> bahwa atribut pelayanan akademik berada pada node paling atas. Dari penelitian ini penulis juga memperoleh pola aturan yang dapat digunakan untuk memprediksi mahasiswa yang terindikasi tepat waktu atau terlambat ketika melakukan pembayaran SPP.</p>
<p>Pandu Pratama Putra, Andi Supriadi Chan (2018)</p>	<p>Pengembangan Aplikasi Perhitungan Prediksi <i>Stock Motor</i> Menggunakan Algoritma C.45 Sebagai Bagian Dari Sistem Pengambilan Keputusan (Studi Kasus di Saudara</p>	<p>Algoritma C.45</p>	<p>Berdasarkan penelitian dan pembahasan yang dilakukan, maka dapat disimpulkan beberapa hal. Hasil dari perhitungan nilai <i>entropy</i> telah memberikan nilai. Berdasarkan proses perhitungan yang telah dilakukan melalui beberapa langkah pada proses penerapan data mining menggunakan algoritma c.45 maka diperoleh pohon keputusan tersebut diperoleh 12 aturan-aturan (<i>rule</i>) dalam menentukan prediksi jumlah persediaan stok motor pada</p>

	Motor)		<i>dealer</i> Saudara Motor.
Ardi Ramdani, Christian Dwi Sofyan, Fauzi Ramdani, Muhammad Fauzi Arya Tama, Muhammad Angga Rachmadsyah (2022)	Algoritma Klasifikasi Data Mining Untuk Memprediksi Masyarakat Dalam Menerima Bantuan Sosial	Algoritma C.45	Dari hasil pengujian digunakan tes algoritma c.45 menghasilkan nilai <i>accuracy</i> sebesar 100.00, nilai untuk <i>preccission</i> sebesar 100.00, dan nilai untuk <i>sensitivity</i> atau <i>recall</i> 100.00. Dengan demikian algoritma c.45 merupakan algoritma dan teknik terbaik untuk memprediksi masyarakat dalam menerima bantuan sosial.

BAB III

DESAIN DAN PERANCANGAN SISTEM

3.1. Analisis Sistem

Pada pembahasan bab ini, akan dilakukan penganalisaan mengenai analisa dan perancangan pembuatan sistem untuk pengklasifikasian “Analisa Algoritma C.45 Untuk Penyediaan Lapangan Pekerjaan Di Kelurahan Sukaramai II”. Dimana analisis sistem ini yaitu penggambaran suatu masalah yang ada pada sistem yang sedang berjalan pada kelurahan tersebut dan sistem yang akan diusulkan pada kelurahan tersebut. Analisis sistem ini bertujuan agar pada pembuatan sistem terwujud klasifikasi pada penduduk dan dapat terwujud sesuai dengan variabel-variabel yang telah ditentukan.

3.1.1. Deskripsi Sistem

3.1.1.1. Sistem Yang Sedang Berjalan

Pada sistem yang sedang berjalan sekarang di kelurahan Sukaramai II masih sulitnya menentukan tersedianya lapangan pekerjaan bagi para penduduk. Yakni dikarenakan tidak ada sebuah standar keputusan yang benar di pemilihan lapangan pekerjaan untuk penduduk. Maka kebiasaan dalam pengambilan keputusan yang diambil masih belum sesuai dimana akan dipilih ialah hanya beberapa saja tidak secara merata, sehingga kesempatan bagi penduduk yang kurang dalam hal pendapatan menjadi hilang. Sehingga hal tersebut menjadi tidak efektif dan efisien dalam segi waktu maupun data.

3.1.1.2. Sistem Yang Akan Diusulkan

Adapun sistem yang diusulkan dalam pengklasifikasian untuk penyediaan lapangan pekerjaan agar penentuan lapangan pekerjaan bagi penduduk kelurahan Sukaramai II secara akurat. Pada sistem ini nanti akan mencari ketepatan data dengan menggunakan algoritma c.45 serta keefektifan dalam penentuan lapangan pekerjaan bagi penduduk di kelurahan Sukaramai II. Adapun variabel-variabel yang digunakan dalam melakukan pengklasifikasian yaitu status perkawinan, luas tanah yang dimiliki, penghasilan, dan umur. Berikut ini akan dijelaskan dari beberapa variabel untuk pengklasifikasian data penduduk, yaitu sebagai berikut :

- a. Status perkawinan, yaitu dapat dilihat penduduk tersebut sudah menikah atau masih lajang.
- b. Luas tanah yang dimiliki, yaitu dapat dilihat penduduk tersebut memiliki luas tanah yang besar, sedang atau kecil.
- c. Penghasilan, yaitu dapat dilihat dari berapa besar jumlah penghasilan yang dimiliki penduduk.
- d. Umur, yaitu dapat dilihat dari berapa umur yang di miliki oleh penduduk tersebut.

3.1.2. Analisa Sistem

Pada penelitian ini selain digunakan sebagai bahan penelitian juga dibutuhkan unsur pendukung yaitu *software* dan *hardware* sebagai bahan pendukung untuk berjalannya penelitian tersebut, unsur pendukung tersebut ialah sebagai berikut :

1. Analisis Perangkat Keras
 - a. *Processor* minimal *Dual Core*
 - b. *RAM* minimal 4 Gb
 - c. *Hardisk* minimal 1024 Gb
2. Analisis Kebutuhan Perangkat Lunak
 - a. Sistem Operasi *Windows 7,8,10,11*
 - b. *XAMPP*
 - c. *Web Browser (google chrome, mozilla firefox, opera, dll)*.
 - d. *Database MySQL*
 - e. Bahasa pemograman PHP

3.1.3. Analisis User

Pengguna adalah orang yang menggunakan sistem aplikasi. Dalam hal ini pengguna harus mengerti dalam setiap prosedur dari penggunaan sistem yang akan diimplementasikan ke dalam perusahaan yang dimana aplikasi tersebut berbasis web. Pengguna yang bertugas untuk mengoperasikan aplikasi berbasis web adalah operator.

3.1.4. Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional menggambarkan proses kegiatan yang akan dipergunakan dalam sebuah sistem dan menjabarkan kebutuhan yang diperlukan oleh sistem agar dapat berjalan dengan baik sesuai yang diperlukan. Fungsionalitas yang akan digunakan dalam sistem ini ialah :

1. Operator

a. Fungsionalitas Unggah Olah Data

Berisikan tempat untuk menambahkan data pada *form input* olah data dengan mengunggah data dalam bentuk *file excel*.

b. Fungsionalitas Hapus Olah Data

Berisikan untuk menghapus seluruh data yang ada di olah data.

c. Fungsionalitas Proses Mining

Berisikan untuk melakukan perhitungan *entropy*, *gain* dan *rasio gain* untuk menentukan pohon keputusan (*rule*) untuk sebagai aturan dalam melakukan prediksi data.

d. Fungsionalitas Uji *Rule*

Berisikan pengujian *rule* dengan mengunggah data penduduk untuk melakukan prediksi yang benar sesuai dengan pohon keputusan (*rule*).

e. Fungsionalitas Hitung Akurasi

Berisikan perhitungan akurasi data yang diunggah di sistem untuk membandingkan kelas asli dengan hasil prediksi dari sistem.

f. Fungsionalitas Prediksi

Berisikan penginputan data penduduk yang akan diprediksi dengan melakukan penginputan nama, status perkawinan, luas tanah yang dimiliki, penghasilan dan umur kemudian akan di submit. Sehingga nantinya akan keluar hasil prediksinya.

g. Fungsionalitas Unggah Data Uji

Berisikan untuk melakukan penambahan data pada *form input* data uji *rule* dengan mengunggah data dalam bentuk *file excel*.

h. Fungsionalitas Hapus Data Uji

Berisikan untuk menghapus seluruh data yang ada di uji *rule*.

- i. Fungsionalitas Hasil Prediksi
Berisikan hasil data yang telah diprediksi di *form* prediksi.
- j. Fungsionalitas Hapus Semua Data Prediksi
Berisikan untuk menghapus seluruh data yang telah diprediksi sebelumnya di *form* prediksi.

3.1.5. Batasan – Batasan Sistem

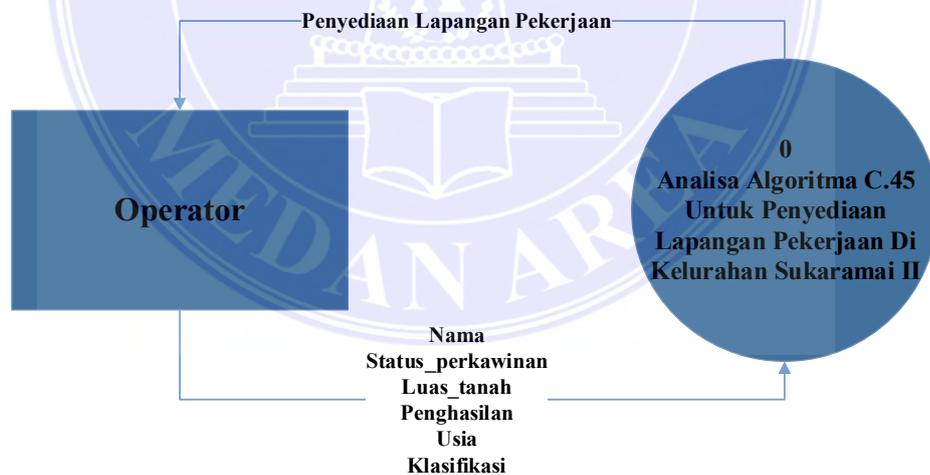
Adapun batasan-batasan dari sistem yang dibangun ini yaitu:

1. Sistem ini digunakan untuk mengklasifikasi data penduduk berdasarkan variabel-variabel yaitu status pernikahan, luas tanah yang dimiliki, penghasilan dan umur menggunakan algoritma c.45
2. Sistem yang akan dibangun nantinya berbasis web.

3.2. Desain Sistem

3.2.1. Diagram Konteks

Diagram konteks yang akan dibuat ialah tingkatan tertinggi dalam diagram aliran data dan hanya ada satu proses, yang dimana ini memperlihatkan sistem secara keseluruhan. Diagram konteks dapat dilihat pada Gambar 3.1.



Gambar 3.1. Diagram Konteks

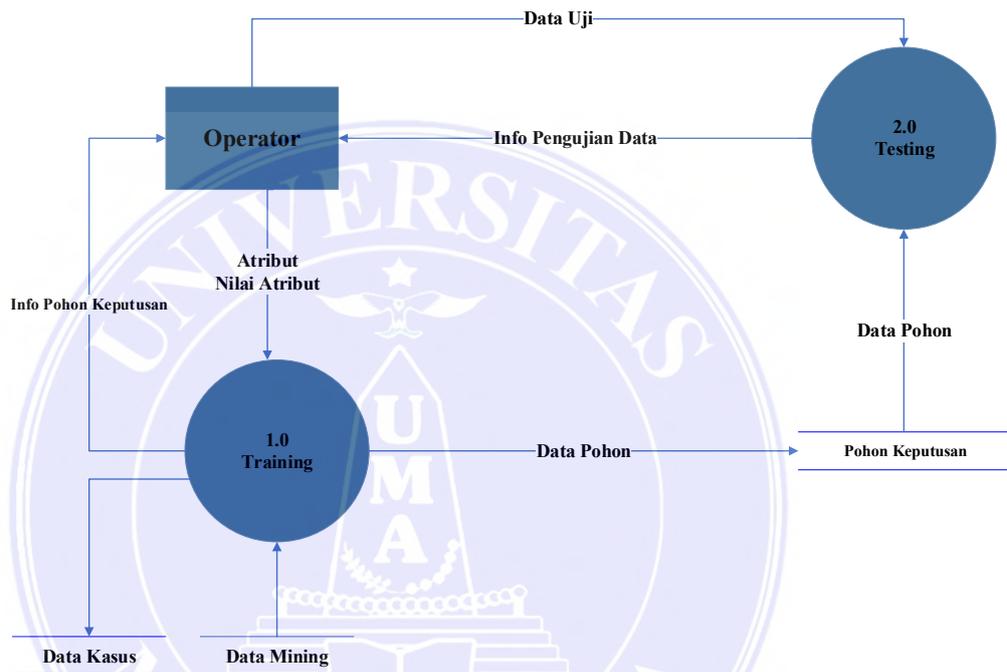
Aliran data berasal dari masukan oleh operator ke dalam sistem yaitu atribut dan nilai atribut. Atribut yang dibutuhkan berupa data penelitian yang diambil dari data sekunder kantor kelurahan Sukaramai II yaitu data penduduk, dan hasil

keputusan. Setelahnya masuk ke proses sistem dan mendapatkan klasifikasi data penduduk.

3.2.2. Data Flow Diagram (DFD)

a. Data Flow Diagram (DFD) level 1

Proses diagram konteks atau bisa dikatakan dengan diagram level 0 akan dipecahkan lagi ke dalam DFD level 1, seperti yang ada pada Gambar 3.2.

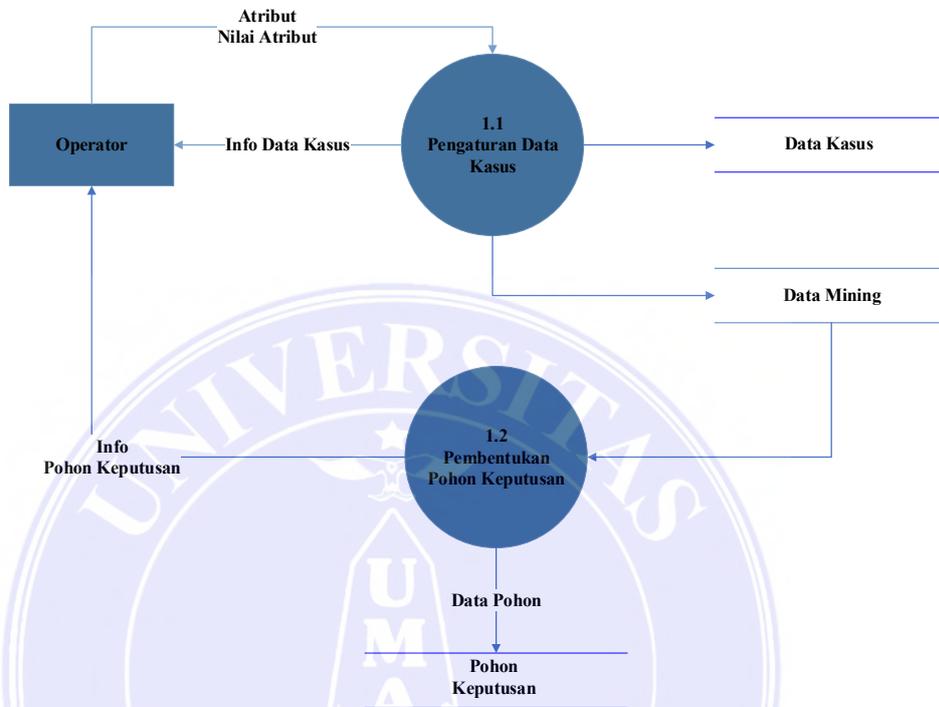


Gambar 3.2. Data Flow Diagram (DFD) Level 1

Aliran data bersumber dari atribut dan nilai atribut (seperti yang sudah dijelaskan di diagram konteks isi dari atribut dan nilai atributnya) yang diperlukan ke dalam proses level 1.0 yang kemudian akan dilakukan *training* data dan menghasilkan *output* info pohon keputusan. Dan data pohon dari proses level 1.0 akan dimasukkan ke *database* yang nantinya data tersebut dimasukkan ke level 2.0. Kemudian masukkan data atribut dan nilai yang sama tadi ke proses *testing* data menggunakan aturan yang sudah dibentuk tadi di pohon keputusan. Hasil keluaran nantinya berupa info pengujian data.

b. Data Flow Diagram (DFD) Level 2

Proses DFD Level 1 dipecah lagi dan akan dibagi dua menjadi proses *Training* dan proses *Testing*. Untuk proses *Training* dapat dilihat pada gambar di bawah ini.



Gambar 3.3. Data Flow Diagram Level 2 Proses

Proses 1.1 merupakan data kasus yang akan diproses di data mining untuk mencari pola klasifikasinya sesuai dengan algoritma C.45. Proses 1.2 yaitu pembentukan pohon keputusan yang nantinya akan menghasilkan keluaran akhir proses yang berupa pohon keputusan dan aturan/*rule*.

3.2.3. Kamus Data (Data Dictionary)

Pada bab ini akan dijelaskan Kamus Data dari DFD yang telah dirancang

1. Tempat Penyimpanan Data

- a. $data_mining = gain + entropy + rasio_gain$
- b. $pohon\ keputusan = parent + akar + keputusan$
- c. $data_penduduk = @id + nama + status_pernikahan + luas_tanah + penghasilan + umur + kelas\ asli$

2. Arus Data (*Data Flow*)

- a. atribut = @id + nama + status_pernikahan + luas_tanah + penghasilan + umur + kelas asli
- b. info_pohon_keputusan = @id + parent + akar + keputusan
- c. info_pengujian_data = @id + nama + status_pernikahan + luas_tanah + penghasilan + umur + kelas asli + kelas hasil + id_rule

Keterangan

id = [0-9]

nama = [A-Z | a-z | 0-9 | . | ']

status_pernikahan = [A-Z | a-z | 0-9 | . | ']

luas_tanah = [A-Z | a-z | 0-9 | . | ']

penghasilan = [0-9]

umur = [A-Z | a-z | 0-9 | . | ']

kelas_asli = [A-Z | a-z | 0-9 | . | ']

gain = [0-9]

entropy = [0-9]

rasio_gain = [0-9]

parent = [A-Z | a-z | 0-9 | . | ']

akar = [A-Z | a-z | 0-9 | . | ']

keputusan = [A-Z | a-z | 0-9 | . | ']

kelas_hasil = [A-Z | a-z | 0-9 | . | ']

id_rule = [A-Z | a-z | 0-9 | . | ']

3.2.4. Spesifikasi Proses (*Process Specification*)

1. Proses 1.1 Pengaturan Data Kasus

Mulai

Baca data kasus

ketika dimasukkan *file* atribut dan nilai atribut lakukan

Baca atribut dan nilai atribut

Verifikasi

jika tidak benar maka tulis pesan kalau tidak rekam ke tabel data mining

berhenti

2. Proses 1.2 Pembetulan Pohon Keputusan

Mulai

Baca data kasus

Pilih proses

jika pilihan = proses pembetulan pohon keputusan maka pohon keputusan;*exit*

berhenti

3. Proses 2.0 Uji Data

Mulai

ketika masukkan data penduduk lakukan

Baca atribut dan nilai atribut

Baca aturan pohon keputusan

Hitung nilai atribut

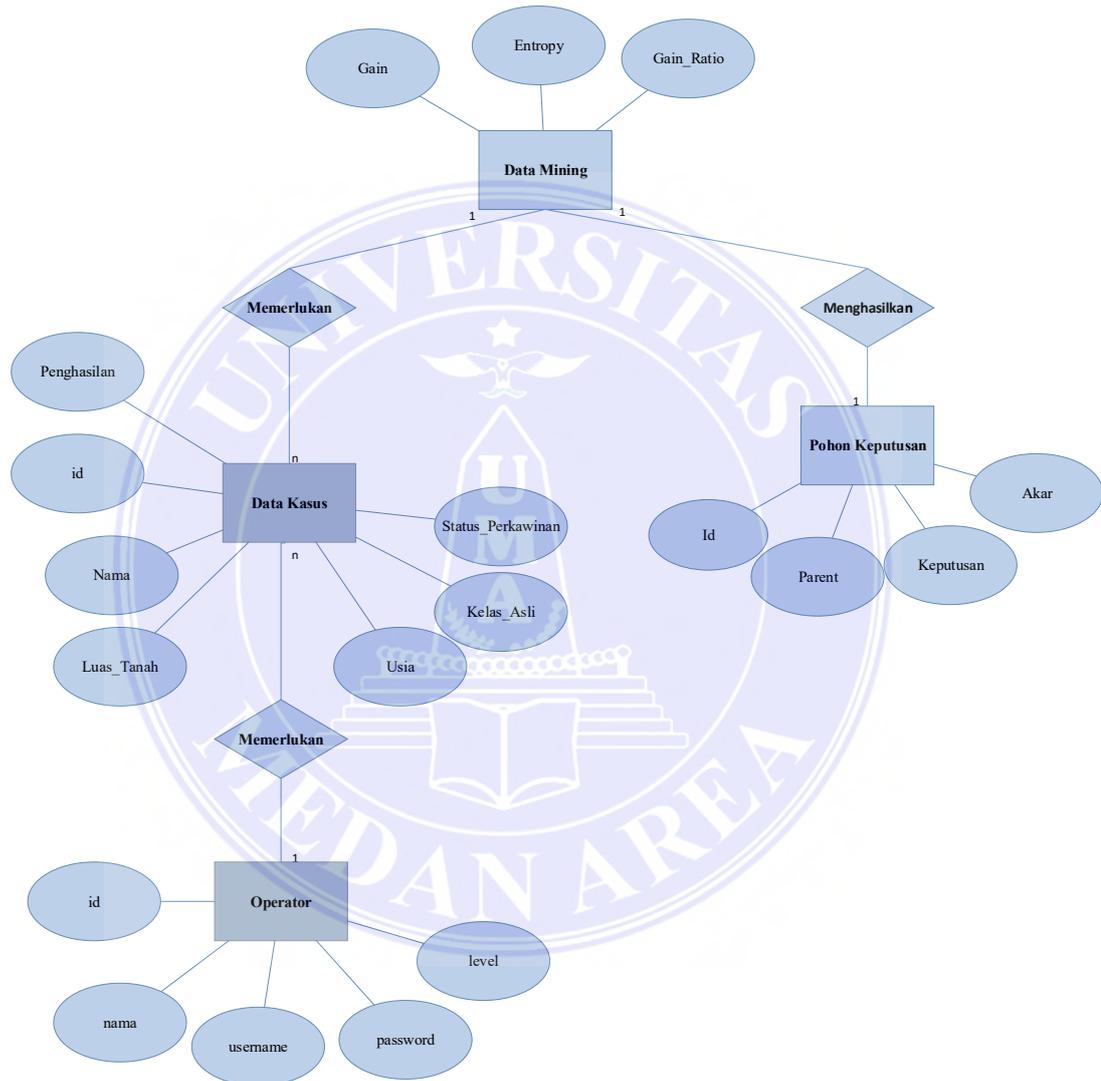
Rekam ke tabel hasil prediksi

selesai

3.3. Desain Basis Data

3.3.1. Diagram Entity Relationship (ER-Diagram)

Adapun ERD (*Entity Relationship Diagram*) yang digunakan dalam analisa algoritma c.45 untuk penyediaan lapangan pekerjaan sebagai berikut:



Gambar 3.6. Entity Relationship Diagram

Hubungan yang dibutuhkan ialah proses data mining yang memerlukan data kasus yang kemudian akan menghasilkan pohon keputusan. Pada entitas pohon keputusan terdapat *key* atribut untuk proses klasifikasi yaitu atribut *rule*.

3.3.2. Struktur Tabel

Table 3.1 Tabel Data Kasus

Nama Field	Jenis Field	Ukuran Field	Keterangan
Id	Integer	11	Primary Key
Nama	Varchar	200	NULL
Status_perkawinan	Varchar	100	NULL
Luas_tanah	Varchar	100	NULL
Umur	Integer	11	NULL
Penghasilan	Double	-	NULL
Kelas_asli	Varchar	200	NULL

Tabel 3.2 Tabel Data Mining

Nama Field	Jenis Field	Ukuran Field	Keterangan
Gain	Double	-	NULL
Entropy	Double	-	NULL
Rasio_gain	Double	-	NULL

Tabel 3.3 Tabel Pohon Keputusan

Nama Field	Jenis Field	Ukuran Field	Keterangan
Id	Integer	11	Primary Key
Parent	Text	100	NULL
Akar	Text	100	NULL
Keputusan	Varchar	100	NULL

Tabel 3.4 Tabel Operator

Nama Field	Jenis Field	Ukuran Field	Keterangan
Id	Integer	11	Primary Key
Nama	Varchar	200	NULL
Username	Varchar	50	NULL
Password	Text	100	NULL
Level	Char	1	NULL

3.4. Contoh Perhitungan Algoritma C.45

Berikut ialah tahapan penyelesaian untuk pembentukan sebuah pohon keputusan dari data penduduk yang akan dipilih sebagai dasar pada penentuan lapangan pekerjaan bagi penduduk di kelurahan Sukaramai II.

Tabel 3.5 Tabel Data Penduduk

No	Nama	Status Pernikahan	Luas Tanah Yang Dimiliki	Penghasilan	Umur	Kelas Asli
1	Yulaini SE	menikah	besar	3800000	55	tidak_tersedia
2	Williem	lajang	besar	2000000	31	tidak_tersedia
3	IR Muliady Kosman	menikah	besar	3500000	59	tidak_tersedia
4	Michael	lajang	besar	2800000	27	tidak_tersedia
5	Mhd Eko Harapan SE	menikah	sedang	3000000	55	tidak_tersedia
6	Gustina	menikah	sedang	1000000	52	tersedia
7	Suwandi Kasim	lajang	sedang	1800000	53	tidak_tersedia
8	Nasrah	menikah	kecil	1200000	66	tidak_tersedia
9	Farhan Arfan	menikah	kecil	1200000	35	tersedia
10	Iksan Arfan	menikah	kecil	1500000	31	tersedia
11	Hendry Halim Lam	menikah	besar	1800000	60	tidak_tersedia
12	Supardi	lajang	sedang	2000000	38	tidak_tersedia
13	Edi Wijaya	lajang	sedang	1900000	30	tidak_tersedia
14	Merlyn Florencia, S.Fam	lajang	sedang	3000000	31	tidak_tersedia
15	Ani Kasim	menikah	sedang	1000000	92	tidak_tersedia
16	Ana Kosman	menikah	kecil	1000000	84	tidak_tersedia
17	Julia	menikah	kecil	800000	91	tidak_tersedia
18	Bran Hart	menikah	besar	1500000	69	tidak_tersedia
19	Furudin Tanwin	menikah	besar	1500000	72	tidak_tersedia
20	Tiffany Karmelia Suwandy	lajang	besar	2500000	29	tidak_tersedia

3.4.1. Pilih Atribut Akar

Dalam penentuan sebuah atribut sebagai akar, dilandaskan dengan sebuah nilai *gain* yang tertinggi dan atribut-atribut yang sudah ada. Untuk menghitung nilai *gain* dan nilai *entropy* dengan menggunakan rumus yang dapat dilihat di bab II.

Dengan penggunaan Rumus 2.1 di bab II maka bisa dihasilkan nilai dari *entropy* dan dengan penggunaan Rumus 2.2 maka bisa dihasilkan nilai *gain* yang nantinya dipakai sebagai akar dalam pembuatan sebuah pohon keputusan.

1. Entropy Total

$$\begin{aligned} \text{Entropy Total} &= \left(-\frac{3}{20} * \log_2 \left(\frac{3}{20}\right)\right) + \left(-\frac{17}{20} * \log_2 \left(\frac{17}{20}\right)\right) \\ &= 0,6098403 \end{aligned}$$

Entropy Total ialah untuk menghitung nilai hasil tersedia (3) dan tidak tersedia (17) dengan nilai 20 ialah total kasus semuanya.

2. Entropy Atribut Status Pernikahan

Untuk atribut status pernikahan terdapat 2 nilai yaitu: menikah dan lajang dengan nilai *entropy* keduanya ialah sebagai berikut :

Status Pernikahan = Menikah

(Total kasus = 13, hasil tersedia = 3, hasil tidak tersedia = 10)

$$\begin{aligned} \text{Entropy (menikah)} &= \left(-\frac{3}{13} * \log_2 \left(\frac{3}{13}\right)\right) + \left(-\frac{10}{13} * \log_2 \left(\frac{10}{13}\right)\right) \\ &= 0,779349837 \end{aligned}$$

Status Pernikahan = Lajang

(Jumlah Kasus = 7, hasil tersedia = 0, hasil tidak tersedia = 7)

$$\text{Entropy (lajang)} = \left(-\frac{0}{7} * \log_2 \left(\frac{0}{7}\right)\right) + \left(-\frac{7}{7} * \log_2 \left(\frac{7}{7}\right)\right)$$

$$= 0$$

3. Entropy Atribut Luas Tanah Yang Dimiliki

Untuk atribut luas tanah terdiri dari 3 nilai yaitu : besar, sedang dan kecil dengan nilai *entropy* masing-masing ialah sebagai berikut:

Luas Tanah = Besar

(Jumlah kasus = 8, hasil tersedia = 0, hasil tidak tersedia = 8)

$$\begin{aligned} Entropy (1) &= \left(-\frac{0}{8} * \log_2 \left(\frac{0}{8}\right)\right) + \left(-\frac{8}{8} * \log_2 \left(\frac{8}{8}\right)\right) \\ &= 0 \end{aligned}$$

Luas Tanah = Sedang

(Total kasus = 7, hasil tersedia = 1, hasil tidak tersedia = 6)

$$\begin{aligned} Entropy (2) &= \left(-\frac{1}{7} * \log_2 \left(\frac{1}{7}\right)\right) + \left(-\frac{6}{7} * \log_2 \left(\frac{6}{7}\right)\right) \\ &= 0,591672779 \end{aligned}$$

Luas Tanah = Kecil

(Total kasus = 5, hasil tersedia = 2, hasil tidak tersedia = 3)

$$\begin{aligned} Entropy (3) &= \left(-\frac{2}{5} * \log_2 \left(\frac{2}{5}\right)\right) + \left(-\frac{3}{5} * \log_2 \left(\frac{3}{5}\right)\right) \\ &= 0,970950594 \end{aligned}$$

4. Entropy Atribut Penghasilan

Untuk atribut status penghasilan terdapat 2 nilai yaitu : (≤ 2000000 & > 2000000) dengan nilai *entropy* masing-masing ialah sebagai berikut:

Penghasilan = (≤ 2000000 & > 2000000)

(Total kasus ≤ 2000000 = 14, hasil tersedia = 3, hasil tidak tersedia = 11)

$$\begin{aligned} Entropy (1) &= \left(-\frac{3}{14} * \log_2 \left(\frac{3}{14}\right)\right) + \left(-\frac{11}{14} * \log_2 \left(\frac{11}{14}\right)\right) \\ &= 0,7495595257 \end{aligned}$$

(Total kasus >2000000 = 6, hasil tersedia = 0, hasil tidak tersedia = 6)

$$\begin{aligned} Entropy (2) &= \left(-\frac{0}{6} * \log_2 \left(\frac{0}{6}\right)\right) + \left(-\frac{6}{6} * \log_2 \left(\frac{6}{6}\right)\right) \\ &= 0 \end{aligned}$$

5. Entropy Atribut Umur

Untuk atribut umur terdiri dari 2 nilai yaitu : (<=64 & >64) dengan nilai *entropy* masing-masing ialah sebagai berikut:

Penghasilan = (<=64 & >64)

(Total kasus <=64 = 14, hasil tersedia = 3, hasil tidak tersedia = 11)

$$\begin{aligned} Entropy (1) &= \left(-\frac{3}{14} * \log_2 \left(\frac{3}{14}\right)\right) + \left(-\frac{11}{14} * \log_2 \left(\frac{11}{14}\right)\right) \\ &= 0,7495595257 \end{aligned}$$

(Total kasus >64 = 6, hasil tersedia = 0, hasil tidak tersedia = 6)

$$\begin{aligned} Entropy (2) &= \left(-\frac{0}{6} * \log_2 \left(\frac{0}{6}\right)\right) + \left(-\frac{6}{6} * \log_2 \left(\frac{6}{6}\right)\right) \\ &= 0 \end{aligned}$$

Untuk selanjutnya ialah melakukan perhitungan nilai *gain* di setiap atribut.

6. Gain (Total, Status Pernikahan)

Gain untuk status pernikahan adalah :

$$= 0,609840305 - \left(\left(\frac{13}{20} * 0,77934498\right) + \left(\frac{7}{20} * 0\right)\right) = 0,10326291$$

Dengan total = 20, jumlah kasus (status pernikahan menikah = 13, status pernikahan lajang = 7).

7. *Gain* (Total, Luas Tanah)

Gain untuk luas tanah adalah :

$$= 0,609840305 - \left(\left(\frac{8}{20} * 0 \right) + \left(\frac{7}{20} * 0,591672779 \right) + \left(\frac{5}{20} * 0,970950594 \right) \right)$$

$$= 0,160017184$$

Dengan total = 20, jumlah kasus (luas tanah besar = 8, sedang = 7, kecil = 5).

8. *Gain* (Total, Penghasilan)

Gain untuk penghasilan adalah :

$$= 0,609840305 - \left(\left(\frac{14}{20} * 0,749595257 \right) + \left(\frac{6}{20} * 0 \right) \right) = 0,085123625$$

Dengan total = 20, jumlah kasus (penghasilan ≤ 2000000 = 14, dan penghasilan > 2000000 = 6).

9. *Gain* (Total, Umur)

Gain untuk umur adalah :

$$= 0,609840305 - \left(\left(\frac{14}{20} * 0,749595257 \right) + \left(\frac{6}{20} * 0 \right) \right) = 0,085123625$$

Dengan total = 20, jumlah kasus (umur ≤ 64 = 14, dan umur < 64 = 6).

Setelah keseluruhan nilai *entropy* dan *gain* didapatkan, maka kemudian hasil dari perhitungan tersebut dibuat ke dalam Tabel 3.6.

Tabel 3.6 Tabel Perhitungan Node 1

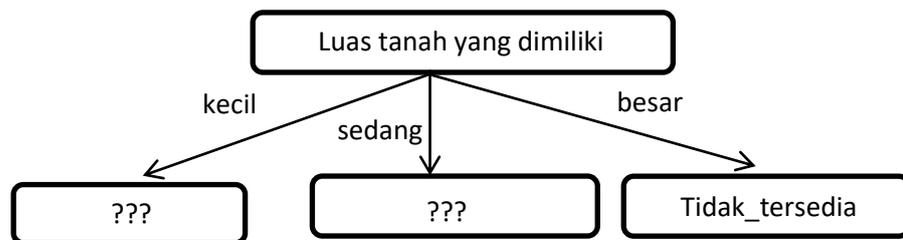
Node		Jumlah kasus	Tersedia	Tidak tersedia	Entropy	Gain
	Total	20	3	17	0,609840305	
						0,160017184
	Luas Tanah Yang Dimiliki	Besar	8	0	8	0
		Sedang	7	1	6	0,591672779
		Kecil	5	2	3	0,970950594
						0,10326291
	Status Pernikahan	Menikah	13	3	10	0,779349837
		Lajang	7	0	7	0
						0,085123625
	Penghasilan	<=2000000	14	3	11	0,749595257
		>2000000	6	0	6	0
						0,085123625
	Umur					
		<=64	14	3	11	0,749595257
		>64	6	0	6	0

Dengan hasil yang ada di Tabel 3.6 dapat dilihat maka dengan *gain* tertinggi ialah luas tanah yang dimiliki, maka atribut status pernikahan dijadikan sebagai *node* akar.

3.4.2. Cabang Untuk Setiap Nilai

Pada atribut luas tanah yang dimiliki yang dibuat menjadi akar terdapat 3 nilai ialah “luas tanah yang dimiliki = besar”, “luas tanah yang dimiliki = sedang” dan “luas tanah yang dimiliki = kecil”. Dari nilai atribut ini, maka nilai “luas tanah yang dimiliki = besar” memiliki hasil = “tidak_tersedia” sehingga tidak dibutuhkannya perhitungan kembali.

Sedangkan nilai “luas tanah yang dimiliki = sedang” dan “luas tanah yang dimiliki = kecil” memiliki hasil = “tersedia” dan hasil “tidak_tersedia” sehingga nilai atribut “luas tanah yang dimiliki = sedang” dan “luas tanah yang dimiliki = kecil” perlu dibutuhkannya perhitungannya lagi. Gambar pohon keputusan sementara dapat dilihat pada Gambar 3.7.



Gambar 3.7 Pohon Keputusan Hasil Perhitungan Node 1

Dari Gambar 3.7, bisa dilihat yang dijadikan sebagai *node* akar ialah luas tanah yang dimiliki dimana di bagian ini dibagi atas dalam 3 nilai ialah “luas tanah yang dimiliki = besar”, “luas tanah yang dimiliki = sedang” dan “luas tanah yang dimiliki = kecil”. Untuk nilai dari kelas “luas tanah yang dimiliki = besar” tidak dibutuhkannya perhitungan lagi disebabkan hasilnya ialah “tidak_tersedia”, sedangkan untuk nilai “luas tanah yang dimiliki = sedang” dan “luas tanah yang dimiliki = kecil” sebaiknya kembali dihitung karena hasilnya bernilai antara “tersedia” atau “tidak_tersedia”.

3.4.3. Pembagian Kasus Setiap Cabang

Berikutnya untuk *node* 1 sebagai akar, sesuai dengan cara yang ada diatas maka menghitung kembali nilai *entropy* dari atribut yang masih ada yaitu status pernikahan, penghasilan, dan umur. Setelah menghitung *entropy*, selanjutnya dihitung kembali *gain* untuk di setiap atribut.

1. Entropy Atribut Penghasilan

Untuk atribut penghasilan terdapat 2 nilai yaitu : (≤ 2000000 & > 2000000) dengan nilai *entropy* nya ialah sebagai berikut :

Penghasilan ≤ 2000000

(Total kasus $\leq 2000000 = 5$, hasil tersedia = 1, hasil tidak tersedia = 4)

$$\begin{aligned} Entropy(1) &= \left(-\frac{1}{5} * \log_2 \left(\frac{1}{5} \right) \right) + \left(-\frac{4}{5} * \log_2 \left(\frac{4}{5} \right) \right) \\ &= 0,721928095 \end{aligned}$$

Penghasilan > 2000000

(Total kasus $> 2000000 = 2$, hasil tersedia = 0, hasil tidak tersedia = 2)

$$\begin{aligned} Entropy(2) &= \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2} \right) \right) + \left(-\frac{2}{2} * \log_2 \left(\frac{2}{2} \right) \right) \\ &= 0 \end{aligned}$$

2. Entropy Atribut Status Pernikahan

Untuk atribut status pernikahan terdiri dari 2 nilai yaitu : menikah dan lajang dengan nilai *entropy* masing-masing ialah sebagai berikut :

Status pernikahan = menikah

(Total kasus = 3, hasil tersedia = 1, hasil tidak tersedia = 2)

$$\begin{aligned} \text{Entropy (menikah)} &= \left(-\frac{1}{3} * \log_2 \left(\frac{1}{3}\right)\right) + \left(-\frac{2}{3} * \log_2 \left(\frac{2}{3}\right)\right) \\ &= 0,918295834 \end{aligned}$$

Status pernikahan = lajang

(Total Kasus = 4, hasil tersedia = 0, hasil tidak tersedia = 4)

$$\begin{aligned} \text{Entropy (lajang)} &= \left(-\frac{0}{4} * \log_2 \left(\frac{0}{4}\right)\right) + \left(-\frac{4}{4} * \log_2 \left(\frac{4}{4}\right)\right) \\ &= 0,9182958 \end{aligned}$$

3. Entropy Atribut Umur

Untuk atribut umur terdiri dari 2 nilai yaitu : (<=64 & >64) dengan nilai *entropy* masing-masing ialah sebagai berikut :

Umur <=64

(Total kasus <=64 = 6, hasil tersedia = 1, hasil tidak tersedia = 5)

$$\begin{aligned} \text{Entropy (1)} &= \left(-\frac{1}{6} * \log_2 \left(\frac{1}{6}\right)\right) + \left(-\frac{5}{6} * \log_2 \left(\frac{5}{6}\right)\right) \\ &= 0,650022422 \end{aligned}$$

Umur >64

(Total kasus >64 = 1, hasil tersedia = 0, hasil tidak tersedia = 1)

$$\text{Entropy (2)} = \left(-\frac{0}{1} * \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * \log_2 \left(\frac{1}{1}\right)\right) = 0$$

Untuk selanjutnya ialah melakukan perhitungan nilai *gain* di setiap atribut

4. *Gain* (Total, Status Pernikahan)

Gain untuk status pernikahan adalah :

$$\begin{aligned} &= 0,591672779 - \left(\left(\frac{3}{7} * 0,918295834 \right) + \left(\frac{4}{7} * 0 \right) \right) \\ &= 0,198117421 \end{aligned}$$

Dengan jumlah total = 7, jumlah kasus (status pernikahan menikah = 3, lajang = 4).

5. *Gain* (Total, Penghasilan)

Gain untuk penghasilan adalah :

$$= 0,591672779 - \left(\left(\frac{5}{7} * 0,721928095 \right) + \left(\frac{2}{7} * 0 \right) \right) = 0,076009854$$

Dengan jumlah total = 7, jumlah kasus (penghasilan $\leq 2000000 = 5$, penghasilan $> 2000000 = 2$).

6. *Gain* (Total, Umur)

Gain pada umur ialah :

$$= 0,591672779 - \left(\left(\frac{6}{7} * 0,650022422 \right) + \left(\frac{1}{7} * 0 \right) \right) = 0,034510703$$

Dengan total = 7, total kasus (Umur $\leq 64 = 6$, dan Umur $> 64 = 1$).

Setelah keseluruhan nilai *entropy* dan *gain* yang didapatkan, maka hasil dari perhitungan itu dibuat ke dalam Tabel 3.7.

Tabel 3.7 Tabel Perhitungan Node 1.1

Node			Jumlah kasus	Tersedia	Tidak tersedia	Entropy	Gain
1.1	Total		7	1	6	0,591672779	
	Status Pernikahan	Menikah	3	1	2	0,918295834	0,198117421
		Lajang	4	0	4	0	
	Penghasilan	<=2000000	5	1	4	0,721928095	0,076009854
		>2000000	2	0	2	0	
	Umur	<=64	6	1	5	0,650022422	0,034510703
		>64	1	0	1	0	

Dari Tabel 3.7 maka dilihat bahwa atribut status pernikahan dengan nilai *gain* tertinggi, yang berarti penghasilan ialah *node* selanjutnya. Ada 2 atribut dari status pernikahan yaitu “menikah” dan nilai “lajang”. Nilai “lajang” mempunyai hasil = “tidak_tersedia” sehingga tidak dibutuhkannya perhitungan kembali.

Sedangkan nilai “menikah” memiliki hasil = “tersedia” dan hasil “tidak_tersedia” sehingga nilai atribut “menikah” perlu dibutuhkannya perhitungannya lagi. Setelah mendapatkan cabang kedua maka akan disambung dengan pohon keputusan sebelumnya yang masih belum terlihat keseluruhannya. Untuk melihat cabang pohon keputusan berikutnya dapat dilihat pada gambar pohon keputusan sementara dapat dilihat pada Gambar 3.8.



Gambar 3.8 Pohon Keputusan Node 1.1

3.4.4. Pembagian Kasus Pada Node 1.1.1

Berikutnya untuk *node* 1.1.1, sesuai dengan cara yang ada diatas maka menghitung kembali nilai *entropy* dari atribut yang masih ada yaitu umur dan penghasilan. Setelah menghitung *entropy*, selanjutnya dihitung kembali *gain* untuk di setiap atribut.

1. Entropy Atribut Umur

Untuk atribut umur terdapat 2 nilai yaitu : (≤ 64 & > 64) dengan nilai *entropy* nya ialah sebagai berikut :

Umur ≤ 64

(Total kasus $\leq 64 = 2$, hasil tersedia = 1, hasil tidak tersedia = 1)

$$\begin{aligned} \text{Entropy (1)} &= \left(-\frac{1}{2} * \log_2 \left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} * \log_2 \left(\frac{1}{2}\right)\right) \\ &= 1 \end{aligned}$$

Umur > 64

(Total kasus $> 64 = 1$, hasil tersedia = 0, hasil tidak tersedia = 1)

$$\begin{aligned} \text{Entropy (2)} &= \left(-\frac{0}{1} * \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * \log_2 \left(\frac{1}{1}\right)\right) \\ &= 0 \end{aligned}$$

2. Entropy Atribut Penghasilan

Untuk atribut penghasilan terdiri dari 2 nilai yaitu : ≤ 2000000 dan > 2000000 dengan nilai *entropy* masing-masing ialah sebagai berikut :

Penghasilan ≤ 2000000

(Total kasus = 2, hasil tersedia = 1, hasil tidak tersedia = 1)

$$\begin{aligned} \text{Entropy (1)} &= \left(-\frac{1}{2} * \log_2 \left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} * \log_2 \left(\frac{1}{2}\right)\right) \\ &= 1 \end{aligned}$$

Penghasilan >2000000

(Total Kasus = 1, hasil tersedia = 0, hasil tidak tersedia = 1)

$$Entropy(2) = \left(-\frac{0}{1} * \log_2\left(\frac{0}{2}\right)\right) + \left(-\frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right)$$

$$= 0$$

Untuk selanjutnya ialah melakukan perhitungan nilai *gain* di setiap atribut

3. *Gain* (Total, Umur)

Gain untuk umur adalah :

$$= 0,918295834 - \left(\left(\frac{2}{3} * 1\right) + \left(\frac{1}{3} * 0\right)\right) = 0,251629167$$

Dengan jumlah total = 3, jumlah kasus (umur <=64 = 2, umur >64 = 1).

4. *Gain* (Total, Penghasilan)

Gain pada penghasilan ialah :

$$= 0,918295834 - \left(\left(\frac{2}{3} * 1\right) + \left(\frac{1}{3} * 0\right)\right) = 0,251629167$$

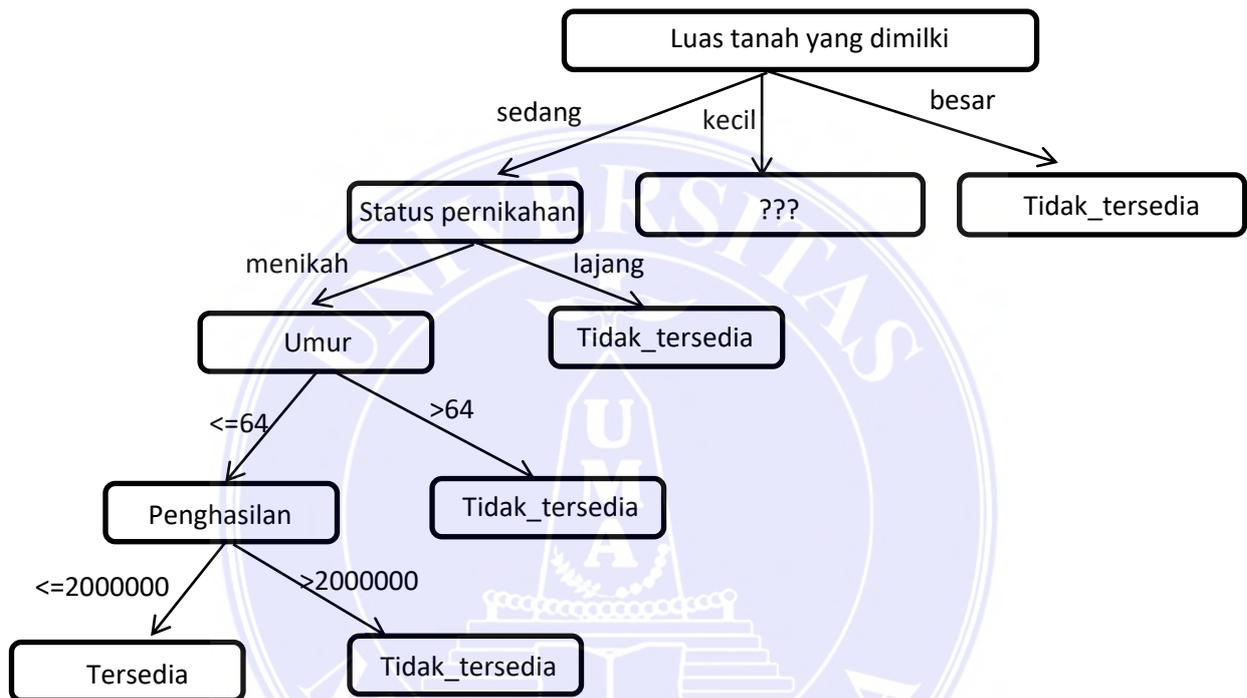
Dengan total = 3, total kasus (penghasilan <=2000000 = 2, dan penghasilan >2000000 = 1).

Setelah keseluruhan nilai *entropy* dan *gain* yang didapatkan, maka hasil dari perhitungan itu dibuat ke dalam Tabel 3.8.

Tabel 3.8 Tabel Perhitungan Node 1.1.1

Node			Jumlah Kasus	Tersedia	Tidak Tersedia	Entropy	Gain
1.1.1	Total		3	1	2	0,918295834	
	Umur	<=64	2	1	1	1	0,251629167
		>64	1	0	1	0	
	Penghasilan	<=2000000	2	1	1	1	0,251629167
		>2000000	1	0	1	0	

Dari Tabel 3.8 maka dilihat bahwa atribut umur dan penghasilan memiliki nilai *gain* yang sama, yang berarti umur ialah node selanjutnya dan penghasilan sebagai node terakhir. Untuk umur yang bernilai ≤ 64 memiliki hasil “tersedia” dan >64 “tidak_tersedia” sehingga tidak dibutuhkannya perhitungan kembali. Untuk penghasilan yang bernilai ≤ 2000000 memiliki hasil “tersedia” dan >2000000 “tidak_tersedia” sehingga tidak dibutuhkannya perhitungan kembali. Untuk melihat cabang pohon keputusan berikutnya dapat dilihat pada gambar pohon keputusan sementara dapat dilihat pada Gambar 3.9.



Gambar 3.9 Pohon Keputusan Node 1.1.1

3.4.5. Pembagian Kasus Pada Node 1.2

Berikutnya untuk *node* 1.2, sesuai dengan cara yang ada diatas maka menghitung kembali nilai *entropy* dari atribut yang masih ada yaitu penghasilan dan luas tanah yang dimiliki. Setelah menghitung *entropy*, selanjutnya dihitung kembali *gain* untuk di setiap atribut.

1. Entropy Atribut Penghasilan

Untuk atribut penghasilan terdapat 2 nilai yaitu : (≤ 2000000 & >2000000) dengan nilai *entropy* nya ialah sebagai berikut :

Penghasilan ≤ 2000000

(Total kasus $\leq 2000000 = 5$, hasil tersedia = 5, hasil tidak tersedia = 0)

$$\begin{aligned} Entropy (1) &= \left(-\frac{5}{5} * \log_2 \left(\frac{5}{5}\right)\right) + \left(-\frac{0}{5} * \log_2 \left(\frac{0}{5}\right)\right) \\ &= 0,970950594 \end{aligned}$$

Penghasilan > 2000000

(Total kasus $> 2000000 = 0$, hasil tersedia = 0, hasil tidak tersedia = 0)

$$\begin{aligned} Entropy (2) &= \left(-\frac{0}{0} * \log_2 \left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2 \left(\frac{0}{0}\right)\right) \\ &= 0 \end{aligned}$$

2. Entropy Atribut Status Pernikahan

Untuk atribut status pernikahan terdapat 2 nilai yaitu : (menikah & lajang)
dengan nilai *entropy* nya ialah sebagai berikut :

Status pernikahan = menikah

(Total kasus menikah = 5, hasil tersedia = 5, hasil tidak tersedia = 0)

$$\begin{aligned} Entropy (menikah) &= \left(-\frac{5}{5} * \log_2 \left(\frac{5}{5}\right)\right) + \left(-\frac{0}{5} * \log_2 \left(\frac{0}{5}\right)\right) \\ &= 0,970950594 \end{aligned}$$

Status pernikahan = lajang

(Total kasus lajang = 0, hasil tersedia = 0, hasil tidak tersedia = 0)

$$\begin{aligned} Entropy (lajang) &= \left(-\frac{0}{0} * \log_2 \left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2 \left(\frac{0}{0}\right)\right) \\ &= 0 \end{aligned}$$

3. Entropy Atribut Umur

Untuk atribut umur terdapat 2 nilai yaitu : (≤ 64 & > 64) dengan nilai *entropy* nya ialah sebagai berikut :

Umur ≤ 64

(Total kasus $\leq 64 = 2$, hasil tersedia = 2, hasil tidak tersedia = 0)

$$\begin{aligned} \text{Entropy (1)} &= \left(-\frac{2}{2} * \log_2 \left(\frac{2}{2}\right)\right) + \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2}\right)\right) \\ &= 0 \end{aligned}$$

Umur > 64

(Total kasus $> 64 = 3$, hasil tersedia = 0, hasil tidak tersedia = 3)

$$\begin{aligned} \text{Entropy (2)} &= \left(-\frac{0}{3} * \log_2 \left(\frac{0}{3}\right)\right) + \left(-\frac{3}{3} * \log_2 \left(\frac{3}{3}\right)\right) \\ &= 0 \end{aligned}$$

3. Gain (Total, Penghasilan)

Gain pada penghasilan ialah :

$$= 0,970950594 - \left(\left(\frac{5}{5} * 0,970950594\right) + \left(\frac{0}{5} * 0\right)\right) = 0$$

Dengan total = 5, total kasus (penghasilan $\leq 2000000 = 5$, dan penghasilan $> 2000000 = 0$).

4. Gain (Total, Status Pernikahan)

Gain pada status pernikahan ialah :

$$= 0,970950594 - \left(\left(\frac{5}{5} * 0,970950594\right) + \left(\frac{0}{5} * 0\right)\right) = 0$$

Dengan total = 5, total kasus (status pernikahan menikah = 5, status pernikahan lajang = 0).

5. *Gain* (Total, Umur)

Gain pada umur ialah :

$$= 0,970950594 - \left(\left(\frac{2}{5} * 0 \right) + \left(\frac{3}{5} * 0 \right) \right) = 0$$

Dengan total = 5, total kasus (status pernikahan menikah = 2, status pernikahan lajang = 3).

Setelah keseluruhan nilai *entropy* dan *gain* yang didapatkan, maka hasil dari perhitungan itu dibuat ke dalam Tabel 3.9.

Tabel 3.9 Tabel Perhitungan Node 1.2

Node		Jumlah kasus	Tersedia	Tidak tersedia	Entropy	Gain
1.2	Total	5	2	3	0,970950594	0
	Status Pernikahan					
	Menikah	5	2	3	0,970950594	
	Lajang	0	0	0	0	
	Penghasilan					0
	<=2000000	5	2	3	0,970950594	
	>2000000	0	0	0	0	
	Umur					0,970950594
	<=64	2	2	0	0	
	>64	3	0	3	0	

Dari Tabel 3.9 maka dilihat bahwa atribut umur dengan nilai *gain* tertinggi, yang berarti umur ialah node terakhir.



Gambar 3.10 Pohon Keputusan Node 1.2

3.4.6. Pohon Keputusan

Pada perhitungan *entropy* dan *gain* yang telah dilakukan di atas, maka diperoleh sebuah pohon keputusan akhir seperti pada Gambar 3.11.



Gambar 3.11 Pohon Keputusan Cabang Akhir

3.4.7. General Rule

Berdasarkan hasil dari pohon keputusan yang telah terbentuk pada Gambar 3.11, maka dapat disimpulkan sebuah aturan (*rule*) sebagai berikut :

1. *If* luas tanah yang dimiliki = besar *then* tidak tersedia
2. *If* luas tanah yang dimiliki = kecil *and* umur > 64 *then* tidak_tersedia
3. *If* luas tanah yang dimiliki = kecil *and* umur <= 64 *then* tersedia
4. *If* luas tanah yang dimiliki = sedang *and* status pernikahan = menikah *and* umur <= 64 *and* penghasilan <= 2000000 *then* tersedia
5. *If* luas tanah yang dimiliki = sedang *and* status pernikahan = menikah *and* umur <= 64 *and* penghasilan > 2000000 *then* tidak_tersedia
6. *If* luas tanah yang dimiliki = sedang *and* status pernikahan = menikah *and* umur > 64 *then* tidak_tersedia

Keterangan :

1. Jika penduduk memiliki sebuah nilai dengan luas tanah yang dimiliki = besar, maka akan dihasilkan nilai tidak_tersedia. Bahwasanya jika penduduk memiliki nilai dengan luas tanah yang dimiliki = besar, maka akan dilihat atribut selanjutnya. Karena ada penduduk yang memiliki nilai dengan status pernikahan menikah.
2. Jika penduduk mempunyai nilai luas tanah yang dimiliki = kecil, kemudian dapat dilihat di nilai atribut berikutnya yakni umur (beralaskan perhitungan *gain* dengan nilai tertinggi setelah luas tanah yang dimiliki). Apabila dengan kasus ini penduduk mempunyai nilai dengan umur > 64 , maka hasilnya tidak_tersedia.
3. Jika penduduk mempunyai nilai dengan luas tanah yang dimiliki = kecil, maka akan dilihat di nilai atribut berikutnya, yaitu umur (berdasarkan pada perhitungan *gain* dengan nilai tertinggi setelah luas tanah yang dimiliki). Jika pada kasus ini, penduduk mempunyai nilai dengan umur ≤ 64 , maka hasilnya akan dilihat atribut berikutnya.
4. Jika penduduk mempunyai nilai dengan luas tanah yang dimiliki = sedang, maka akan dilihat di nilai atribut berikutnya, yaitu status pernikahan (berdasarkan pada perhitungan *gain* dengan nilai tertinggi setelah luas tanah yang dimiliki). Jika pada kasus ini, penduduk mempunyai nilai dengan sedang, maka hasilnya akan dilihat atribut berikutnya. Yaitu status pernikahan = menikah. Kemudian akan lihat kembali nilai atribut berikutnya, yaitu umur ≤ 64 , dan atribut terakhir yaitu penghasilan ≤ 2000000 , maka hasilnya tersedia.
5. Jika penduduk mempunyai nilai dengan luas tanah yang dimiliki = sedang, maka akan dilihat di nilai atribut berikutnya, yaitu status pernikahan (berdasarkan pada perhitungan *gain* dengan nilai tertinggi setelah luas tanah yang dimiliki). Jika pada kasus ini, penduduk mempunyai nilai dengan sedang, maka hasilnya akan dilihat atribut berikutnya. Yaitu status pernikahan = menikah. Kemudian akan lihat kembali nilai atribut

berikutnya, yaitu umur ≤ 64 , dan atribut terakhir yaitu penghasilan > 2000000 , maka hasilnya tidak tersedia.

6. Jika penduduk mempunyai nilai dengan luas tanah yang dimiliki = sedang, maka akan dilihat di nilai atribut berikutnya, yaitu status pernikahan (berdasarkan pada perhitungan *gain* dengan nilai tertinggi setelah luas tanah yang dimiliki). Jika pada kasus ini, penduduk mempunyai nilai dengan sedang, maka hasilnya akan dilihat atribut berikutnya. Yaitu status pernikahan = menikah. Kemudian akan lihat kembali nilai atribut berikutnya, yaitu umur > 64 , maka hasilnya tidak tersedia.
7. Tiap atribut harus dihitung *entropy* dan *gain* secara satu persatu. Ini nantinya guna dengan memeriksa sebuah nilai *gain* tertinggi, sehingga akan dihasilkannya node akar (lihat pada Gambar 3.7). Pohon keputusan tersebut sudah sesuai dengan data yang telah diperoleh, agar pohon keputusan tersebut dapat sebagai dasar untuk memproses pengklasifikasian penduduk untuk penyediaan lapangan pekerjaan di kelurahan Sukaramai II.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil pembahasan dan uji coba yang telah dilakukan, dapat disimpulkan:

- 1 Dengan menggunakan sistem ini dalam melakukan penentuan lapangan pekerjaan bagi penduduk dapat dilakukan dengan akurat dan data pun dapat tersimpan dengan baik di *database*.
- 2 Dari analisa yang telah dilakukan, dalam proses penerapan algoritma c.45 maka dapat terlihat keakuratan data nya dalam penentuan lapangan pekerjaan bagi penduduk. Meski dengan data yang banyak tetapi kekarutan data nya sangat baik.
- 3 Dari hasil pengujian yang dilakukan sebanyak 3 kali pengujian dengan sampel data sebanyak 378 data, diperoleh nilai setiap pengujian memiliki tingkat akurasi yang berbeda. Pada pengujian pertama didapatkan nilai akurasi 100%, pada pengujian kedua memiliki tingkat akurasi 100%, dan pada pengujian ketiga memiliki akurasi sebesar 100%.

5.2. Saran

Adapun saran untuk menyempurnakan sistem yang telah dibuat adalah sebagai berikut:

1. Dalam penelitian selanjutnya dapat menggunakan algoritma lainnya yang lebih *update* atau dibandingkan untuk melihat tingkat akurasinya.
2. Selanjutnya sistem yang dibangun nantinya bisa dapat bekerja secara daring.
3. Dalam penelitian selanjutnya bisa menambahkan enkripsi pada *database* dengan berbagai metode enkripsi yang tersedia agar tidak terjadi kebocoran data.

DAFTAR PUSTAKA

- Ardiansyah, D., & Walim, W. (2019). Algoritma C.45 Untuk Klasifikasi Calon Peserta Lomba Cerdas Cermat Siswa SMP Dengan Menggunakan Aplikasi Rapid Miner. *Infokar*, 5-12.
- Fajri, A. A., & Iriani, R. (2022). Pengaruh Kemiskinan Dan Pengangguran Terhadap Pertumbuhan Ekomomi Di Perovinsi Bali. *Ekonomi Pembangunan*, 53-66.
- Franita , R. (2016). Analisa Pengangguran Di Indonesia. *Ilmu Pengetahuan Sosial*, 88-93.
- Huda, B. (2018). Sistem Informasi Data Penduduk Berbasis Android Dan Web Monitoring Studi Kasus Pemerintah Kota Karawang. 62-69.
- Jollyta, D., Ramdhan , W., & Zarlis, M. (2020). *Konsep Data Mining Dan Penerapan*. Yogyakarta: CV Budi Utama.
- Kalsum, U. (2017). Pengaruh Pengangguran Dan Inflasi Terhadap Pertumbuhan Ekonomi Di Sumatera Utara. *Jurnal Ekonomikawan*, 87-94.
- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C.45. *Edik Informatika*, 213-219.
- Mulyadi, M. (2016). Peran Pemerintah Dalam Mengatasi Pengangguran Dan Kemiskinan Dalam Masyarakat. 221-236.
- Nofriansyah, D., & Nurcahyo, G. W. (2019). *Algoritma Data Mining Dan Pengujian*. Yogyakarta: CV Budi Utama.
- Novianti, B., Rismawan, T., & Bahri, S. (2016). Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus : SMA Negeri 1 Pontianak). *Jurnal Coding*, 75-84.
- Seran, K. J., & Naiheli, V. N. (2021). Development Of Promotional Media For Oepuah Village Potentiatility Waterfall Method. *Journal Of Information And Technology Unimor (JITU)*, 31-36.
- Setyaningrum, S. (2013). *Konsep dan Perancangan Basis Data*. Yogyakarta: Skripta.
- Sholeh, A. (2017). Masalah Ketenagakerjaan Dan Pengangguran Di Indonesia. *Jurnal Ilmiah Cano Ekonomos*, 83-92.
- Tanjung, D. Y. (2021). Optimalisasi Algoritma C.45 Untuk Prediksi Kerusakan Mesin ATM. 12-21.

- Tanjung, F. A., Windarto, A. P., & Fauzan, M. (2020). Penerapan Metode K-Means Pada Pengelompokan Pengangguran Di Indonesia. *Jurnal Riset Sistem Informasi Dan Teknik Informatika*, 61-74.
- Trisyanto. (2017). *Analisis & Perancangan Sistem Basis Data*. Surabaya: CV. Garuda Mas Sejahtera.
- Utami, H. F., & Asnawati. (2015). *Rekayasa Perangkat Lunak*. Yogyakarta: Deepublish.
- Wahyudi, M., Pujiastuti, L., & Solikhun. (2020). Penerapan Data Mining Dalam Mengelompokkan Data Pengangguran Terbuka Menurut Provinsi Dengan Menggunakan Algoritma K-Means. 433-440.
- Watik, D., Novitasari, F., & Trisiana, A. (2022). Analisis Peran Pemerintahan Dalam Pengendalian Pertumbuhan Penduduk. 45-56.
- Apandi, T. H., Maulana, R. B., Piarna, R., & Vernanda, D. (2019). MENGANALISIS KEMUNGKINAN KETERLAMBATAN PEMBAYARAN SPP Dengan Algoritma C.45 (Studi Kasus Politeknik TEDC Bandung). *Techno Nusa Mandiri*, 93-98.
- Putra, P. P., & Chan, A. S. (2018). Pengembangan Aplikasi Perhitungan Prediksi Stock Motor Menggunakan Algoritma C.45 Sebagai Bagian Dari Sistem Pengambilan Keputusan. *Inovtek Polbeng-Seri Informatika*, 24-33.
- Ramdani, A., Sofyan, C. D., Ramdani, F., Tama, M. F., & Rachmatsyah, M. A. (2022). Algoritma Klasifikasi Data Mining Untuk Memprediksi Masyarakat Dalam Menerima Bantuan Sosial. *Jurnal Ilmiah Sistem Informasi*, 39-47.

LAMPIRAN

Lampiran Kode Program

Halaman Proses Mining (mencari Entropy dan Gain)

```

<?php
function format_decimal($value){
    return round($value, 3);
}
function proses_DT($db_object, $parent, $kasus_cabang1, $kasus_cabang2) {
    echo "cabang 1<br>";
    pembentukan_tree($db_object, $parent, $kasus_cabang1);
    echo "cabang 2<br>";
    pembentukan_tree($db_object, $parent, $kasus_cabang2);
}
function pembentukan_tree($db_object, $N_parent, $kasus) {
    if ($N_parent != '') {
        $kondisi = $N_parent . " AND " . $kasus;
    } else {
        $kondisi = $kasus;
    }
    echo $kondisi . "<br>";

    $cek = cek_heterohomogen($db_object, 'kelas_asli', $kondisi);
    if ($cek == 'homogen') {
        echo "<br>LEAF ||";
        $sql_keputusan = $db_object->
        db_query("SELECT DISTINCT(kelas_asli) FROM "
            . "data_latih WHERE $kondisi");
        $row_keputusan = $db_object->db_fetch_array($sql_keputusan);
        $keputusan = $row_keputusan['0'];
        pangkas($db_object, $N_parent, $kasus, $keputusan);
    }
    else if ($cek == 'heterogen') {
        $kondisi_kelas_asli = '';
        if ($kondisi != '') {
            $kondisi_kelas_asli = $kondisi . " AND ";
        }
        $jml_tersedia = jumlah_data($db_object, "$kondisi_kelas_asli kelas_asli='tersedia'");
        $jml_tidak_tersedia = jumlah_data($db_object, "$kondisi_kelas_asli kelas_asli='tidak_tersedia'");
        $jml_total = $jml_tersedia + $jml_tidak_tersedia ;
    }
}

```



```

    }
    hitung_gain($db_object, $kondisi, "Penghasilan=1000000", $entropy_all, "penghasilan<=1000000", "penghasilan>1000000", "", "", "")
;
    hitung_gain($db_object, $kondisi, "Penghasilan=2000000", $entropy_all, "penghasilan<=2000000", "penghasilan>2000000", "", "", "")
;
    hitung_gain($db_object, $kondisi, "Penghasilan=3000000", $entropy_all, "penghasilan<=3000000", "penghasilan>3000000", "", "", "")
;
    hitung_gain($db_object, $kondisi, "Umur=35", $entropy_all, "umur<=35", "umur>35", "", "", "");
    hitung_gain($db_object, $kondisi, "Umur=40", $entropy_all, "umur<=40", "umur>40", "", "", "");
    hitung_gain($db_object, $kondisi, "Umur=45", $entropy_all, "umur<=45", "umur>45", "", "", "");
    hitung_gain($db_object, $kondisi, "Umur=64", $entropy_all, "umur<=60", "umur>60", "", "", "");
    echo "</tbody>";
    echo "</table>";
    $sql_max = $db_object->db_query("SELECT MAX(gain) FROM gain");
    $row_max = $db_object->db_fetch_array($sql_max);
    $max_gain = $row_max[0];
    $sql = $db_object->db_query("SELECT * FROM gain WHERE gain=$max_gain");
    $row = $db_object->db_fetch_array($sql);
    $atribut = $row[2];
    echo "Atribut terpilih = " . $atribut . ", dengan nilai gain = " . $max_gain . "<br>";
    echo "<br>=====<br>";
    if ($max_gain == 0) {
        echo "<br>LEAF ";
        $Ntersedia = $kondisi . " AND kelas_asli='tersedia'";
        $Ntidak_tersedia = $kondisi . " AND kelas_asli='tidak_tersedia'";
        $jumlahtersedia = jumlah_data($db_object, "$Ntersedia");
        $jumlahtidak_tersedia = jumlah_data($db_object, "$Ntidak_tersedia");
        if ($jumlahtersedia >= $jumlahtidak_tersedia) {
            $keputusan = 'tersedia';
        }
        else {
            $keputusan = 'tidak_tersedia';
        }
    }
}

```

```

    }
    pangkas($db_object, $N_parent, $kasus, $keputusan);
  }
  else {
    if ($atribut == "luas_tanah") {
      proses_DT($db_object, $kondisi, "($atribut='besar')", "($atribut='sedang')", "($atribut='kecil')");
    }
    if ($atribut == "status_pernikahan") {
      if($jmlStatusPernikahan==3){
        $cabang = array();
        $cabang = hitung_rasio($db_object, $kondisi , 'status_pernikahan', $max_gain, $nilai_status_pernikahan[0], $nilai_status_pernikahan[1], $nilai_status_pernikahan[2], ',');
        $exp_cabang = explode(" , ", $cabang[1]);
        proses_DT($db_object, $kondisi , "($atribut='$cabang[0]')", "($atribut='$exp_cabang[0]' OR $atribut='$exp_cabang[1]')");
      }
      else if($jmlStatusPernikahan==2){
        proses_DT($db_object, $kondisi , "($atribut='$nilai_status_pernikahan[0]')", "($atribut='$nilai_status_pernikahan[1]')");
      }
    }
    if ($atribut == "Penghasilan=1000000") {
      proses_DT($db_object, $kondisi, "(penghasilan<=1000000)", "(penghasilan>1000000)");
    } else if ($atribut == "Penghasilan=2000000") {
      proses_DT($db_object, $kondisi, "(penghasilan<=2000000)", "(penghasilan>2000000)");
    } else if ($atribut == "Penghasilan=3000000") {
      proses_DT($db_object, $kondisi, "(penghasilan<=3000000)", "(penghasilan>3000000)");
    }
    if ($atribut == "Umur=35") {
      proses_DT($db_object, $kondisi, "(umur<=35)", "(umur>35)");
    } else if ($atribut == "Umur=40") {
      proses_DT($db_object, $kondisi, "(umur<=40)", "(umur>40)");
    } else if ($atribut == "Umur=45") {
      proses_DT($db_object, $kondisi, "(umur<=45)", "(umur>45)");
    } else if ($atribut == "Umur=64") {
      proses_DT($db_object, $kondisi, "(umur<=64)", "(umur>64)");
    }
  }
}
}

```

```
function cek_nilaiAtribut($db_object, $field , $kondisi){
    $hasil = array();
    if($kondisi==''){
        $sql = $db_object-
>db_query("SELECT DISTINCT($field) FROM data_latih");
    }else{
        $sql = $db_object-
>db_query("SELECT DISTINCT($field) FROM data_latih WHERE $kondisi");
    }
    $a=0;
    while($row = $db_object->db_fetch_array($sql)){
        $hasil[$a] = $row['0'];
        $a++;
    }
    return $hasil;
}

function cek_heterohomogen($db_object, $field, $kondisi) {
    if ($kondisi == '') {
        $sql = $db_object-
>db_query("SELECT DISTINCT($field) FROM data_latih");
    } else {
        $sql = $db_object-
>db_query("SELECT DISTINCT($field) FROM data_latih WHERE $kondisi");
    }
    if ($db_object->db_num_rows($sql) == 1) {
        $nilai = "homogen";
    } else {
        $nilai = "heterogen";
    }
    return $nilai;
}

function jumlah_data($db_object, $kondisi) {
    if ($kondisi == '') {
        $sql = "SELECT COUNT(*) FROM data_latih";
    } else {
        $sql = "SELECT COUNT(*) FROM data_latih WHERE $kondisi";
    }
    $query = $db_object->db_query($sql);
    $row = $db_object->db_fetch_array($query);
    $jml = $row['0'];
    return $jml;
}

function pangkas($db_object, $PARENT, $KASUS, $LEAF) {
    $sql_in = "INSERT INTO t_keputusan "
        . "(parent,akar,keputusan)"
```

```

        . " VALUES (\ "$PARENT\" , \ "$KASUS\" , \ "$LEAF\" );
        $db_object->db_query($sql_in);
        echo "Keputusan = " . $LEAF . "<br>=====  

====<br>";
    }
function hitung_gain($db_object, $kasus, $atribut, $ent_all, $kondisi1, $kondisi2, $kondisi3, $kondisi4, $kondisi5) {
    $data_kasus = '';
    if ($kasus != '') {
        $data_kasus = $kasus . " AND ";
    }
    if ($kondisi3 == '') {
        $j_tersedia1 = jumlah_data($db_object, "$data_kasus kelas_asli='tersedia' AND $kondisi1");
        $j_tidak_tersedia1 = jumlah_data($db_object, "$data_kasus kelas_asli='tidak_tersedia' AND $kondisi1");
        $jml1 = $j_tersedia1 + $j_tidak_tersedia1;
        $j_tersedia2 = jumlah_data($db_object, "$data_kasus kelas_asli='tersedia' AND $kondisi2");
        $j_tidak_tersedia2 = jumlah_data($db_object, "$data_kasus kelas_asli='tidak_tersedia' AND $kondisi2");
        $jml2 = $j_tidak_tersedia2 + $j_tidak_tersedia2 ;
        $jml_total = $jml1 + $jml2;
        $ent1 = hitung_entropy($j_tersedia1, $j_tersedia1);
        $ent2 = hitung_entropy($j_tersedia2, $j_tidak_tersedia2);
        $gain = $ent_all -
        ((( $jml1 / $jml_total ) * $ent1 ) + ( ( $jml2 / $jml_total ) * $ent2 ));
        $gain = format_decimal($gain);
        echo "<tr>";
        echo "<td>" . $kondisi1 . "</td>";
        echo "<td>" . $jml1 . "</td>";
        echo "<td>" . $j_tersedia1 . "</td>";
        echo "<td>" . $j_tidak_tersedia1 . "</td>";
        echo "<td>" . $ent1 . "</td>";
        echo "<td>&nbsp;</td>";
        echo "</tr>";
        echo "<tr>";
        echo "<td>" . $kondisi2 . "</td>";
        echo "<td>" . $jml2 . "</td>";
        echo "<td>" . $j_tersedia2 . "</td>";
        echo "<td>" . $j_tidak_tersedia2 . "</td>";
        echo "<td>" . $ent2 . "</td>";
        echo "<td>" . $gain . "</td>";
        echo "</tr>";
        echo "<tr><td colspan='8'></td></tr>";
    }
}

```

```

}
else if($kondisi4==''){
    $j_tersedia1 = jumlah_data($db_object, "$data_kasus kelas_asli
='tersedia' AND $kondisi1");
    $j_tidak_tersedia1 = jumlah_data($db_object, "$data_kasus kela
s_asli='tidak_tersedia' AND $kondisi1");
    $jml1 = $j_tersedia1 + $j_tersedia1 ;
    $j_tersedia2 = jumlah_data($db_object, "$data_kasus kelas_asli
='tersedia' AND $kondisi2");
    $j_tidak_tersedia2 = jumlah_data($db_object, "$data_kasus kela
s_asli='tidak_tersedia' AND $kondisi2");
    $jml2 = $j_tersedia2 + $j_tidak_tersedia2;
    $j_tersedia3 = jumlah_data($db_object, "$data_kasus kelas_asli
='tersedia' AND $kondisi3");
    $j_tidak_tersedia3 = jumlah_data($db_object, "$data_kasus kela
s_asli='tidak_tersedia' AND $kondisi3");
    $jml3 = $j_tersedia3 + $j_tidak_tersedia3;
    $jml_total = $jml1 + $jml2 + $jml3;
    $sent1 = hitung_entropy($j_tersedia1 , $j_tidak_tersedia1);
    $sent2 = hitung_entropy($j_tersedia2 , $j_tidak_tersedia2);
    $sent3 = hitung_entropy($j_tersedia3 , $j_tidak_tersedia3);
    $gain = $ent_all -
(((($jml1/$jml_total)*$sent1) + (($jml2/$jml_total)*$sent2)
+ (($jml3/$jml_total)*$sent3));
    $gain = format_decimal($gain);
    echo "<tr>";
    echo "<td>".$kondisi1."</td>";
    echo "<td>".$jml1."</td>";
    echo "<td>".$j_tersedia1."</td>";
    echo "<td>".$j_tidak_tersedia1."</td>";
    echo "<td>".$sent1."</td>";
    echo "<td>&nbsp;</td>";
    echo "</tr>";
    echo "<tr>";
    echo "<td>".$kondisi2."</td>";
    echo "<td>".$jml2."</td>";
    echo "<td>".$j_tersedia2."</td>";
    echo "<td>".$j_tidak_tersedia2."</td>";
    echo "<td>".$sent2."</td>";
    echo "<td>&nbsp;</td>";
    echo "</tr>";
    echo "<tr>";
    echo "<td>".$kondisi3."</td>";
    echo "<td>".$jml3."</td>";
    echo "<td>".$j_tersedia3."</td>";

```

```

        echo "<td>".$j_tidak_tersedia3."</td>";
        echo "<td>".$sent3."</td>";
        echo "<td>".$gain."</td>";
        echo "</tr>";
        echo "<tr><td colspan='8'></td></tr>";
    }
    $db_object->
    >db_query("INSERT INTO gain VALUES ('','1','$atribut','$gain')");
}
function hitung_entropy($nilai1, $nilai2) {
    $total = $nilai1 + $nilai2;
    $atribut1 = (-
($nilai1 / $total) * (log(($nilai1 / $total), 2)));
    $atribut2 = (-
($nilai2 / $total) * (log(($nilai2 / $total), 2)));
    $atribut1 = is_nan($atribut1)?0:$atribut1;
    $atribut2 = is_nan($atribut2)?0:$atribut2;
    $entropy = $atribut1 + $atribut2 ;
    $entropy = format_decimal($entropy);
    return $entropy;
}
function hitung_rasio($db_object, $kasus , $atribut , $gain , $nilai
1 , $nilai2 , $nilai3 , $nilai4 , $nilai5){
    $data_kasus = '';
    if($kasus!=''){
        $data_kasus = $kasus." AND ";
    }
    $jmlNilai=5;
    if($nilai5==''){
        $jmlNilai=4;
    }
    if($nilai4==''){
        $jmlNilai=3;
    }
    $db_object->db_query("TRUNCATE rasio_gain");
    if($jmlNilai==3){
        $opsi11 = jumlah_data($db_object, "$data_kasus ($atribut='$nila
i2' OR $atribut='$nilai3')");
        $opsi12 = jumlah_data($db_object, "$data_kasus $atribut='$nilai
1'");
        $tot_opsi1=$opsi11+$opsi12;
        $opsi21 = jumlah_data($db_object, "$data_kasus ($atribut='$nila
i3' OR $atribut='$nilai1')");
        $opsi22 = jumlah_data($db_object, "$data_kasus $atribut='$nilai
2'");

```

```

    $tot_opsi2=$opsi21+$opsi22;
    $opsi31 = jumlah_data($db_object, "$data_kasus ($atribut='$nilai1' OR $atribut='$nilai2')");
    $opsi32 = jumlah_data($db_object, "$data_kasus $atribut='$nilai3'");
    $tot_opsi3=$opsi31+$opsi32;
    $opsi1 = (-
($opsi11/$tot_opsi1)*(log(($opsi11/$tot_opsi1),2))) + (-
($opsi12/$tot_opsi1)*(log(($opsi12/$tot_opsi1),2)));
    $opsi2 = (-
($opsi21/$tot_opsi2)*(log(($opsi21/$tot_opsi2),2))) + (-
($opsi22/$tot_opsi2)*(log(($opsi22/$tot_opsi2),2)));
    $opsi3 = (-
($opsi31/$tot_opsi3)*(log(($opsi31/$tot_opsi3),2))) + (-
($opsi32/$tot_opsi3)*(log(($opsi32/$tot_opsi3),2)));
    $opsi1 = format_decimal($opsi1);
    $opsi2 = format_decimal($opsi2);
    $opsi3 = format_decimal($opsi3);
    $rasio1 = $gain/$opsi1;
    $rasio2 = $gain/$opsi2;
    $rasio3 = $gain/$opsi3;
    $rasio1 = format_decimal($rasio1);
    $rasio2 = format_decimal($rasio2);
    $rasio3 = format_decimal($rasio3);
    echo "Ops1 1 : <br>jumlah ".$nilai2."/ ".$nilai3." = ".$opsi11.
        "<br>jumlah ".$nilai1." = ".$opsi12.
        "<br>Split = ".$opsi1.
        "<br>Rasio = ".$rasio1."<br>";
    echo "Ops1 2 : <br>jumlah ".$nilai3."/ ".$nilai1." = ".$opsi21.
        "<br>jumlah ".$nilai2." = ".$opsi22.
        "<br>Split = ".$opsi2.
        "<br>Rasio = ".$rasio2."<br>";
    echo "Ops1 3 : <br>jumlah ".$nilai1."/ ".$nilai2." = ".$opsi31.
        "<br>jumlah ".$nilai3." = ".$opsi32.
        "<br>Split = ".$opsi3.
        "<br>Rasio = ".$rasio3."<br>";

    $db_object->db_query("INSERT INTO rasio_gain VALUES
(' , 'opsi1' , '$nilai1' , '$nilai2 , $nilai3' , '$rasio1'),
(' , 'opsi2' , '$nilai2' , '$nilai3 , $nilai1' , '$rasio2'),
(' , 'opsi3' , '$nilai3' , '$nilai1 , $nilai2' , '$rasio3')");
}
$sql_max = $db_object-
>db_query("SELECT MAX(rasio_gain) FROM rasio_gain");
$row_max = $db_object->db_fetch_array($sql_max);
$max_rasio = $row_max['0'];

```

```

    $sql = $db_object-
>db_query("SELECT * FROM rasio_gain WHERE rasio_gain=$max_rasio");
    $row = $db_object->db_fetch_array($sql);
    $opsiMax = array();
    $opsiMax[0] = $row[2];
    $opsiMax[1] = $row[3];
    echo "<br>=====<br>";
    return $opsiMax;
}
function klasifikasi($db_object, $n_status_pernikahan, $n_luas_tanah
, $n_penghasilan, $n_umur) {
    $sql = $db_object->db_query("SELECT * FROM t_keputusan");
    $keputusan = $id_rule_keputusan = "";
    while ($row = $db_object->db_fetch_array($sql)) {
        if ($row['parent'] != '') {
            $rule = $row['parent'] . " AND " . $row['akar'];
        } else {
            $rule = $row['akar'];
        }
        $rule = str_replace("<=", " k ", $rule);
        $rule = str_replace("=", " s ", $rule);
        $rule = str_replace(">", " l ", $rule);
        $rule = str_replace("status_pernikahan", "'$n_status_pernikaha
n'", $rule);
        $rule = str_replace("luas_tanah", "'$n_luas_tanah'", $rule);
        $rule = str_replace("penghasilan", "'$n_penghasilan'", $rule);
        $rule = str_replace("umur", "'$n_umur'", $rule);
        $rule = str_replace("'", "", $rule);
        $explodeAND = explode(" AND ", $rule);
        $jmlAND = count($explodeAND);
        $explodeAND = str_replace("(", "", $explodeAND);
        $explodeAND = str_replace(")", "", $explodeAND);
        $bolAND=array();
        $n=0;
        while($n<$jmlAND){
            $explodeOR = explode(" OR ", $explodeAND[$n]);
            $jmlOR = count($explodeOR);
            $bol=array();
            $a=0;
            while($a<$jmlOR){
                $exrule2 = explode(" ", $explodeOR[$a]);
                $parameter = $exrule2[1];
                if($parameter=='s'){
                    $explodeRule = explode(" s ", $explodeOR[$a]);
                    if($explodeRule[0]==$explodeRule[1]){

```

```
    $bol[$a]="Benar";
}else if($explodeRule[0]!=$explodeRule[1]){
    $bol[$a]="Salah";
}
}else if($parameter=='k'){
    $explodeRule = explode(" k ",$explodeOR[$a]);
if($explodeRule[0]<=$explodeRule[1]){
    $bol[$a]="Benar";
}else{
    $bol[$a]="Salah";
}
}else if($parameter=='l'){
    $explodeRule = explode(" l ",$explodeOR[$a]);
if($explodeRule[0]>$explodeRule[1]){
    $bol[$a]="Benar";
}else{
    $bol[$a]="Salah";
}
}
}
$a++;
}
$bolAND[$n]="Salah";
$b=0;
while($b<$jmlOR){
if($bol[$b]=="Benar"){
    $bolAND[$n]="Benar";
}
$b++;
}
$n++;
}
$boolRule="Benar";
$a=0;
while($a<$jmlAND){
if($bolAND[$a]=="Salah"){
    $boolRule="Salah";
break;
}
$a++;
}
if($boolRule=="Benar"){
    $keputusan=$row['keputusan'];
    $id_rule_keputusan=$row['id'];
break;
}
```

```
if ($keputusan == '') {
    $que = $db_object->db_query("SELECT parent FROM t_keputusan");
    $jml = array();
    $exParent = array();
    $i = 0;
    while ($row_baris = $db_object->db_fetch_array($que)) {
        $exParent = explode(" AND ", $row_baris['parent']);
        $jml[$i] = count($exParent);
        $i++;
    }
    $maxParent = max($jml);
    $sql_query = $db_object-
>db_query("SELECT * FROM t_keputusan");
    while ($row_bar = $db_object->db_fetch_array($sql_query)) {
        $explP = explode(" AND ", $row_bar['parent']);
        $jmlT = count($explP);
        if ($jmlT == $maxParent) {
            $keputusan = $row_bar['keputusan'];
            $id_rule[$it] = $row_bar['id'];
            $id_rule_keputusan = $row_bar['id'];
            break;
        }
    }
}
return array('keputusan' => $keputusan, 'id_rule' => $id_rule_ke
putusan);
}
```

Lampiran SK Pembimbing Skripsi

	UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK	
<small>Kampus I : Jalan Negeri Nomor 1 Medan Estetis/Jalan PEGI Nomor 1 (81) 7366070, 7366100, 7364340, 7366781, Fax (81) 7366886 Medan 20223 Kampus II : Jalan Selatua/ Nomor 70 / Jalan Sei Serayu Nomor 70 A, (81) 6229402, Fax. (81) 6226331 Medan 20122 Website: www.umma.ac.id, E-mail: umma.medan@umma.ac.id</small>		
Nomor	: 75/FT.6/01.10/III/2022	9 Maret 2022
Lamp	: -	
Hal	: Perubahan Judul Tugas Akhir	
Yth. Pembimbing Tugas Akhir Juanda Hakim Lubis, ST, M. Kom Andre Hasudungan Lubis, S. Ti, MS. c di Tempat		
Dengan hormat, Sehubungan dengan adanya perubahan judul tugas akhir maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa tersebut :		
Nama	: Widya Rizki Utami	
N P M	: 178160050	
Jurusan	: Informatika	
Maka dengan hormat kami mengharapkan kesediaan saudara :		
1. Juanda Hakim Lubis, ST, M. Kom	(Sebagai Pembimbing I)	
2. Andre Hasudungan Lubis, S. Ti, MS. c	(Sebagai Pembimbing II)	
Adapun Tugas Akhir Skripsi berjudul :		
"Analisis Algoritma C.45 untuk Penyediaan Lapangan Pekerjaan di Kelurahan Sukaramai II".		
SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.		
Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.		
		 Dr. Rahmad Syah, S. Kom, M. Kom

Lampiran Surat Pengantar Riset

	UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax. (061) 7366998 Medan 20223 Kampus II : Jalan Seliabudi Nomor 79 / Jalan Sel Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122 Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id									
Nomor	: 105/FT.6/01.10/IV/2022	23 April 2022								
Lamp	: -									
Hal	: Penelitian Dan Pengambilan Data Tugas Akhir									
Yth. Lurah Kelurahan Sukaramai II Jln. A.R Hakim Gg. Kamboja No.2 Di Medan										
Dengan hormat, Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :										
	<table border="1"><thead><tr><th>NO</th><th>NAMA</th><th>NPM</th><th>PRODI</th></tr></thead><tbody><tr><td>1</td><td>Widya Rizki Utami</td><td>178160050</td><td>Informatika</td></tr></tbody></table>	NO	NAMA	NPM	PRODI	1	Widya Rizki Utami	178160050	Informatika	
NO	NAMA	NPM	PRODI							
1	Widya Rizki Utami	178160050	Informatika							
Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.										
Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana lengkap pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :										
Analisis Algoritma C.45 untuk Penyediaan Lapangan Pekerjaan di Kelurahan Sukaramai II										
Atas perhatian dan kerja sama yang baik diucapkan terima kasih.										
	 Dekan, Dr. Rahmat Syah, S. Kom, M. Kom									
Tembusan : 1. Ka. BAMAI 2. Mahasiswa 3. File										

Lampiran Surat Selesai Riset



PEMERINTAH KOTA MEDAN KECAMATAN MEDAN AREA KELURAHAN SUKARAMAI II

Alamat Kantor : Jl. A.R. Hakim Gg. Kamboja No. 2 Telp. 061 - 7345902 Medan 20216

Medan, 05 Juli 2022

Nomor : 420 / 101 / SK - II / VII / 2022

Lampiran :

Perihal : Selesai Penelitian dan Pengambilan

Data Tugas Akhir

Kepada Yth,

Dekan Fakultas Teknik

Universitas Medan Area

Di -

M E D A N

1. Berdasarkan Surat Dekan Fakultas Teknik Universitas Medan Area Nomor 105/FT.6/01.10/IV/022 tanggal 23 April 2022 perihal Penelitian dan Pengambilan Data Tugas Akhir, Judul : Analisis Algoritma C.45 untuk Penyediaan Lapangan Pekerjaan di Kelurahan Sukaramai II
2. Sehubungan dengan hal tersebut diatas, mahasiswa yang bernama :

No.	Nama	NPM	Prog. Studi
1.	Widya Rizki Utami	178160050	Informatika

Telah menyelesaikan penelitian di Kelurahan Sukaramai II Kecamatan Medan Area, selanjutnya kepada yang bersangkutan agar menyampaikan hasil penelitian tersebut kepada Kepala Kelurahan Sukaramai II Kec. Medan Area.

3. Demikian disampaikan, untuk dimaklumi.

KEPALA KELURAHAN SUKARAMAI-II
KECAMATAN MEDAN AREA

KELURAHAN
SUKARAMAI-II

TAUFIKA KAMBE, SH.
NIP.19860401 201001 1 015.

TERWUJUDNYA MASYARAKAT KOTA MEDAN YANG BERKAH, MAJU DAN KONDUSIF

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 23/12/22

Access From (repository.uma.ac.id)23/12/22

Lampiran Turnitin



Similarity Report ID: oid:29477:25964764

PAPER NAME

Skripsi Full - Widya.pdf

AUTHOR

Dinda Rizky Aprillya

WORD COUNT

13058 Words

CHARACTER COUNT

74829 Characters

PAGE COUNT

83 Pages

FILE SIZE

1.8MB

SUBMISSION DATE

Nov 2, 2022 3:43 PM GMT+7

REPORT DATE

Nov 2, 2022 3:46 PM GMT+7

● 30% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 28% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 15% Submitted Works database

● Excluded from Similarity Report

- Small Matches (Less than 10 words)

Summary