

**PEMANFAATAN ALGORITMA C4.5 UNTUK MENENTUKAN
BONUS TAHUNAN PADA TOKO UMANO VERO**

SKRIPSI



Oleh:
Faul Johan
161510062

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

**PEMANFAATAN ALGORITMA C4.5 UNTUK MENENTUKAN
BONUS TAHUNAN PADA TOKO UMANO VERO**

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana**



**Oleh
Faul Johan
161510062**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Faul Johan
NPM : 161510062
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa "Skripsi" yang saya buat dengan judul:

"Pemanfaatan Algoritma C4.5 Untuk Menentukan Bonus Tahunan Pada Toko Umamo Vero"

Adalah hasil karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan saya, didalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Batam, 20 Juli 2020



Faul Johan
161510062

**PEMANFAATAN ALGORITMA C4.5 UNTUK MENENTUKAN
BONUS TAHUNAN PADA TOKO UMANO VERO**

HALAMAN PENGESAHAN

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana**

**Oleh
Faul Johan
161510062**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, 20 Juli 2020



**Sasa Ani Arnomo, S.Kom., M.SI.
Pembimbing**

ABSTRAK

Toko Umano Vero merupakan salah satu toko yang menjual penjualan tas, sepatu, sandal, tali pinggang dan dompet yang berlokasi di Batam, Indonesia. Pada tahun 2019, diketahui pada Toko Umano Vero dalam menentukan bonus tahunan kepada karyawan masih belum merata, karyawan pada Toko Umano Vero tidak puas dengan bonus tahunan yang mereka terima sehingga masalah ini menimbulkan kurangnya kinerja karyawan, kedisiplinan karyawan, dan loyalitas karyawan terhadap Toko Umano Vero. Sehingga penulis melakukan analisis dalam pemanfaatan algoritma C4.5 untuk menentukan bonus tahunan pada Toko Umano Vero. Dalam pelaksanaan penelitian, penulis melakukan metode atau teknik dari *data mining* menggunakan klasifikasi *decision tree* atau pohon keputusan serta algoritma jenis C4.5 yang dilakukan dengan perhitungan secara manual dan juga menggunakan aplikasi WEKA versi 3.9.2. Hasil dari klasifikasi data yang diperoleh dengan aturan pada pohon keputusan dengan jenis algoritma dari C4.5 secara manual dengan aturan *if-then* menggunakan aplikasi WEKA versi 3.9.2 sama-sama benar dan akurat sehingga memudahkan penulis saat melakukan penelitian berdasarkan data yang diberikan oleh pemilik Toko Umano Vero. Dan terakhir sesuai dengan hasil perhitungan dan pengujian, penulis juga berpendapat bahwa dari proses pemanfaatan algoritma C4.5 dalam menentukan bonus tahunan yang dilakukan ternyata secara cukup besar dipengaruhi oleh beberapa atribut dari variabel bonus tahunan, yaitu atribut target, atribut kehadiran dan atribut masa kerja.

Kata kunci: Algoritma C4.5, Analisis, *Data Mining*, Bonus Tahunan.

ABSTRACT

Umano Vero Store is one of the stores that sells sales of bags, shoes, sandals, belts and wallets located in Batam, Indonesia. In the year 2019, known in the store Umano vero in determining the annual bonus to employees still not even evenly, employees at the store Umano Vero not satisfied with the annual bonus they received so that this problem raises less employee performance, employee discipline, and employee loyalty to the store Umano Vero. So the author performs analysis in the utilization of algorithm C4.5 to determine the annual bonus on the store Umano Vero. In the implementation of research, the authors perform methods or techniques of data mining using decision tree Classification or decree tree as well as algorithm type C4.5 done by manually calculation and also use Weka application version 3.9.2. Result of classification of data obtained by rules on decision tree with algorithm type from C4.5 manually with If-then rule using Weka application version 3.9.2 is equally true and accurate so as to facilitate the author while conducting research based on data provided by the store owner umano Vero. And finally in accordance with the results of calculations and tests, the author also argues that from the process of utilization of algorithm C4.5 in determining the annual bonus that is done turns out to be sufficiently large influenced by some attributes of the annual bonus variable, namely the target attribute, the presence attribute and the work period attributes.

Keywords: C4.5 Algorithm, Analysis, Data Mining, Annual Bonus.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa, yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan skripsi yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada program studi sistem informasi di Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

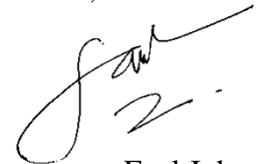
Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. selaku Rektor Universitas Putera Batam.
2. Bapak Muhammat Rasid Ridho, S.Kom., M.SI. selaku Ketua Program Studi Sistem Informasi Universitas Putera Batam.
3. Bapak Sasa Ani Arnomo, S.Kom., M.SI. selaku pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam.
4. Bapak Nopriadi, S.Kom., M.Kom. selaku pembimbing Akademik pada Program Studi Sistem Informasi Universitas Putera Batam.
5. Dosen dan Staff Universitas Putera Batam, yang telah memberikan ilmu dan pengetahuan kepada penulis.
6. Direksi dan karyawan Toko Umano Vero yang telah membantu dan bersedia memberikan data dalam penelitian ini.

7. Kedua orang tua penulis yang selalu setia memberikan dorongan semangat, perhatian dan doa.
8. Seluruh keluarga yang telah memberi semangat dan dukungan selama menempuh pendidikan di Universitas Putera Batam.
9. Teman-teman program studi Sistem Informasi angkatan tahun 2016, yang telah berjuang bersama menempuh pendidikan di Universitas Putera Batam.
10. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu memberikan penyertaanNya, Amin.

Batam, 20 Juli 2020



Faul Johan

DAFTAR ISI

	Halaman
SURAT PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR RUMUS	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah.....	4
1.3. Batasan Masalah.....	4
1.4. Rumusan Masalah	5
1.5. Tujuan Penelitian	5
1.6. Manfaat Penelitian	6
1.6.1. Manfaat Teoritis	6
1.6.2. Manfaat Praktis	7
BAB II TINJAUAN PUSTAKA.....	8
2.1. Teori Dasar.....	8
2.1.1. Sistem Pendukung Keputusan (SPK).....	8
2.1.2. KDD (<i>Knowledge Discovery In Database</i>).....	8
2.1.3. <i>Data Mining</i>	10
2.1.4. <i>Decision Tree</i>	14
2.1.5. <i>Waikato Environment for Knowledge Analysis (WEKA)</i>	17
2.1.6. Algoritma C4.5.....	18
2.1.7. Kinerja Karyawan	20
2.1.8. <i>Reward</i>	21
2.1.9. Bonus	21
2.1.10. Bonus Tahunan	22
2.1.11. Karyawan	22
2.1.12. Insentif.....	22
2.2. Penelitian Terdahulu	23
2.3. Kerangka Pemikiran.....	26
BAB III METODE PENELITIAN	27
3.1. Desain Penelitian.....	27
3.2. Objek Penelitian	29
3.2.1. Lokasi Penelitian.....	29
3.2.2. Jadwal Penelitian.....	29
3.3. Populasi dan Sampel	30
3.3.1. Populasi	30
3.3.2. Sampel Jenuh	30

3.4. Variabel Penelitian	31
3.5. Teknik Pengumpulan Data	31
3.6. Metode Penelitian.....	32
BAB IV HASIL DAN PEMBAHASAN	34
4.1. Hasil	34
4.2. Pembahasan.....	34
4.2.1. Melakukan Pra Proses Data Penelitian	35
4.2.2. Analisa Data	36
4.2.2.1. Pembersihan Data	37
4.2.2.2. Seleksi Data	37
4.2.2.3. Transformasi Data.....	37
4.2.3. Pohon Keputusan	41
4.2.4. Pengujian Dengan Aplikasi WEKA.....	60
BAB V SIMPULAN DAN SARAN	68
5.1. Simpulan	68
5.2. Saran.....	70
DAFTAR PUSTAKA	71
LAMPIRAN	
Lampiran 1. Pendukung Penelitian	
Lampiran 2. Daftar Riwayat Hidup	
Lampiran 3. Surat Keterangan Penelitian	

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Proses Dalam KDD atau <i>Knowledge Discovery in Database</i>	9
Gambar 2.2 Pengelompokan Teknik <i>Data Mining</i>	14
Gambar 2.3 Model <i>Decision Tree</i>	17
Gambar 2.4 Tampilan Utama WEKA	18
Gambar 2.5 Kerangka Berpikir	26
Gambar 3.1 Desain Penelitian	27
Gambar 4.1 Pohon Keputusan Pada <i>Node</i> Pertama.....	48
Gambar 4.2 Pohon Keputusan Pada <i>Node</i> Kedua	54
Gambar 4.3 Pohon Keputusan Pada <i>Node</i> Ketiga	59
Gambar 4.4 Data Pengujian Pada Aplikasi Microsoft Excel	61
Gambar 4.5 Data Pengujian Pada Aplikasi <i>Notepad</i>	62
Gambar 4.6 Tampilan Awal Aplikasi WEKA.....	63
Gambar 4.7 Tampilan Lokasi Penyimpanan Data Pengujian.....	63
Gambar 4.8 Tampilan Pemilihan Atribut Penelitian	64
Gambar 4.9 Tampilan Tab Menu <i>Classify</i>	64
Gambar 4.10 Tampilan Saat Memilih Menu J48	65
Gambar 4.11 Tampilan <i>Classify Output</i>	66
Gambar 4.12 Tampilan <i>Visualize Tree</i>	66
Gambar 4.13 Tampilan Pohon Keputusan Menggunakan Aplikasi Weka.....	67

DAFTAR TABEL

Halaman

Tabel 3.1 Jadwal Penelitian	30
Tabel 4.1 Format Data Karyawan Toko Umano Vero	35
Tabel 4.2 Format Data Karyawan Toko Umano Vero (lanjutan).....	36
Tabel 4.3 Klasifikasi Atribut Tanggung Jawab.....	37
Tabel 4.4 Klasifikasi Atribut Target.....	38
Tabel 4.5 Klasifikasi Atribut Kehadiran	38
Tabel 4.6 Klasifikasi Atribut Masa Kerja.....	38
Tabel 4.7 Klasifikasi Atribut Loyalitas	38
Tabel 4.8 Klasifikasi Atribut Inisiatif.....	39
Tabel 4.9 Klasifikasi Atribut Bonus Tahunan.....	39
Tabel 4.10 Pra Proses Data Karyawan Setelah Klasifikasi	39
Tabel 4.11 Pra Proses Data Karyawan Setelah Klasifikasi (lanjutan).....	40
Tabel 4.12 Hasil Perhitungan Pada Node Pertama.....	47
Tabel 4.13 Hasil Perhitungan Pada <i>Node</i> Kedua	53

DAFTAR RUMUS

	Halaman
Rumus 2.1 Perhitungan <i>Gain</i>	19
Rumus 2.2 Perhitungan <i>Entropy</i>	20

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi saat ini semakin berkembang dan canggih. Perkembangan tersebut memenuhi dalam segala aspek bidang kehidupan, perusahaan harus dapat mengikuti perkembangan teknologi terutama dengan banyaknya masalah dan kesulitan yang dihadapi. Hal tersebut menjadi wajib karena dengan mengikuti perkembangan teknologi dapat mengatasi masalah-masalah yang ada dengan lebih cepat dan efisien.

Toko Umamo Vero merupakan sebuah toko bergerak dibidang penjualan tas, sepatu, sandal, tali pinggang dan dompet yang berlokasi di Batam, Indonesia. Dalam menentukan bonus tahunan karyawan di perusahaan banyak terdapat kendala-kendala atau masalah yang harus dihadapi perusahaan antara lain efisiensi waktu, banyak perbandingan variabel yang diuji, dan pengambilan keputusan untuk menentukan bonus tahunan karyawan. Seiring berlewatnya tahun sering kali karyawan pada Toko Umamo Vero tidak puas dengan bonus tahunan yang mereka terima sehingga masalah ini menimbulkan kurangnya kinerja karyawan, kedisiplinan karyawan, dan loyalitas karyawan terhadap Toko Umamo Vero.

Salah satunya dalam meningkatkan kinerja karyawan, kedisiplinan karyawan, dan loyalitas karyawan terhadap Toko Umamo Vero, maka setiap tahunnya pimpinan toko umamo vero memberikan bonus pada karyawan dengan

kriteria-kriteria yang telah ditetapkan. Karyawan Toko Umamo Vero sendiri berjumlah 50 orang, sehingga karena banyaknya kriteria dan data karyawan yang dianalisa maka seorang pimpinan perusahaan harus bekerja keras dalam menganalisa karyawan mana yang layak dalam menerima bonus tahunan ini. Maka dari itu Toko Umamo Vero juga harus memperhatikan pula faktor pemberian bonus tahunan sebagai salah satu motif bagi karyawan untuk bekerja. Pemberian bonus tahunan bagi karyawan merupakan kewajiban bagi perusahaan yang harus dipenuhi tepat waktu, dilakukan secara adil dan berdasarkan hasil kerja. Bonus tahunan adalah sebuah pembayaran kompensasi dari perusahaan kepada karyawan yang biasanya berbentuk tunai, yang diberikan apabila kinerja tahunan perusahaan melebihi target yang telah ditentukan. Berdasarkan Surat Edaran Menteri Tenaga Kerja Republik Indonesia No.SE-07/MEN/1990 Tahun 1990 tentang pengelompokan komponen upah dan pendapatan non upah, bonus dikategorikan sebagai komponen non-upah. Bonus tahunan diberikan kepada karyawan swasta di akhir tahun ketika perusahaan mengalami keuntungan atau karena pekerja mengalami peningkatan produktivitas dalam bekerja sehingga perusahaan mencapai target laba. Berdasarkan permasalahan yang ada penulis membuat suatu sistem pendukung untuk menentukan bonus tahunan karyawan Toko Umamo Vero dengan menggunakan metode Algoritma C4.5.

Algoritma C4.5 merupakan salah satu solusi yang dapat digunakan untuk memecahkan masalah terkait kasus pada teknik klasifikasi. Keluaran yang diberikan algoritma C4.5 berupa suatu pohon keputusan yang terstruktur, yang

dapat digunakan dalam mengubah kumpulan data menjadi suatu pohon keputusan yang terdiri dari suatu aturan – aturan keputusan.

Alasan Penulis memilih Metode Algoritma C4.5 karena algoritma klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Misalnya dapat mengolah data numerik (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain. Hasil yang diberikan aplikasi ini adalah laporan bonus karyawan yang telah dilakukan proses perhitungan dengan hasil keputusannya apakah karyawan tersebut layak mendapatkan bonus tahunan atau tidak.

Dalam memberikan bonus tahunan Toko Umamo Vero sendiri memberikan bonus tersebut kepada karyawan dari lama berkerja karyawan dan penjualan karyawan saja sehingga pembagian bonus tahunan sering terjadi ketidak rataan dalam membagi bonus tahunan.

Pemanfaatan algoritma ini diharapkan dapat mempercepat proses perhitungan atau lebih efisien dalam menentukan bonus tahunan. Hal ini sangatlah penting karena Algoritma dapat menyampaikan informasi lebih tepat dan efisien kepada pemilik toko. Maka Toko Umamo Vero dijadikan bahan penelitian penulis usulan penelitian dengan judul: **“PEMANFAATAN ALGORITMA C4.5 UNTUK MENENTUKAN BONUS TAHUNAN PADA TOKO UMAMO VERO”**.

1.2. Identifikasi Masalah

Sesuai dengan penjelasan penulis terkait latar belakang pelaksanaan penelitian yang sudah dijelaskan sebelumnya, maka dapat dibuat beberapa penjelasan mengenai identifikasi terkait masalah pada penelitian, sebagai berikut:

1. Belum adanya sistem untuk menentukan bonus tahunan pada Toko Umamo Vero.
2. Kurangnya kinerja karyawan dikarenakan pembagian bonus tahunan yang dinilai kurang efektif.
3. Pemilik Toko Umamo Vero tidak bisa menentukan bonus tahunan yang jelas.

1.3. Batasan Masalah

Demikian penjelasan terkait pembatasan mengenai permasalahan yang akan dibahas pada penelitian agar isi dari penelitian menjadi lebih fokus dan dapat mencapai tujuan dari pelaksanaan penelitian, sebagai berikut:

1. Data yang diambil pada penelitian ini berkaitan dengan kinerja karyawan untuk menentukan bonus tahunan pada Toko Umamo Vero.
2. Pelaksanaan kegiatan terkait menentukan bonus tahunan dari sistem informasi dilakukan dengan menggunakan metode dari *data mining* dan proses klasifikasi *decision tree* serta perhitungan algoritma dari C4.5 dengan pengujian dari *data mining* dilakukan menggunakan *software* atau perangkat lunak WEKA 3.9.2.
3. Hasil dari pelaksanaan kegiatan penelitian yang dibuat adalah dokumen yang berisi penjelasan temuan hasil dari kegiatan analisa mengenai sistem

informasi setelah melalui proses analisis dan evaluasi dan rekomendasi terkait dengan hasil dari analisa mengenai sistem informasi yang didapat.

1.4. Rumusan Masalah

Sesuai dengan isi dari latar belakang, maka penulis dapat memberi penjelasan mengenai rumusan terhadap permasalahan yang ada dalam penelitian, yaitu:

1. Bagaimana memanfaatkan *data mining* dengan menggunakan proses klasifikasi *decision tree* dengan perhitungan algoritma dari C4.5 dalam menganalisa dan menentukan bonus tahunan karyawan pada Toko Umano Vero?
2. Bagaimana menerapkan algoritma C4.5 dalam menentukan bonus tahunan karyawan pada Toko Umano Vero?
3. Bagaimana cara memudahkan pemilik Toko Umano Vero untuk menentukan bonus tahunan dengan memanfaatkan data mining dengan menggunakan *data mining* dan proses klasifikasi *decision tree* dengan perhitungan algoritma dari C4.5?

1.5. Tujuan Penelitian

Untuk menjawab isi dari permasalahan pada penelitian maka pelaksanaan penelitian dilakukan dengan beberapa tujuan, sebagai berikut:

1. Mengetahui cara memanfaatkan *data mining* dengan menggunakan proses klasifikasi *decision tree* dengan perhitungan algoritma dari C4.5 dalam

menganalisa dan menentukan bonus tahunan karyawan pada Toko Umamo Vero.

2. Mengetahui cara menerapkan algoritma C4.5 dalam menentukan bonus tahunan karyawan pada Toko Umamo Vero.
3. Mengetahui cara memudahkan pemilik Toko Umamo Vero untuk menentukan bonus tahunan dengan memanfaatkan data mining dengan menggunakan *data mining* dan proses klasifikasi *decision tree* dengan perhitungan algoritma dari C4.5.

1.6. Manfaat Penelitian

Demikian penjelasan yang menjadi harapan dari penulis terkait beberapa manfaat yang dapat diberikan baik secara teoritis dan praktis dan beberapa penjelasan tersebut adalah sebagai berikut.

1.6.1. Manfaat Teoritis

Demikian penjelasan terkait manfaat secara teoritis yang berasal dari hasil pelaksanaan penelitian yang dilakukan, yaitu:

1. Penelitian ini dapat memberikan informasi terkait wawasan khususnya mengenai cara memanfaatkan algoritma C4.5 mengenai menentukan bonus tahunan.
2. Menjadi kesempatan penulis dalam mempraktikkan hasil pengetahuan yang didapat selama masa perkuliahan dan menambah pengalaman analisa menggunakan C4.5 mengenai sistem informasi.

3. Dapat dijadikan sebagai acuan atau pedoman serta tambahan informasi bagi peneliti lain yang akan melakukan penelitian selanjutnya.

1.6.2. Manfaat Praktis

Demikian penjelasan terkait manfaat secara praktis yang berasal dari hasil pelaksanaan penelitian yang dilakukan, yaitu:

1. Hasil dari analisa menentukan bonus tahunan tersebut bisa digunakan sebagai bahan untuk memperbaiki proses pengelolaan dan sistem untuk menentukan bonus tahunan pada Toko Umano Vero ditahun-tahun berikutnya.
2. Sebagai referensi untuk penelitian-penelitian selanjutnya yang berhubungan dengan teknik *data mining* dengan metode algoritma C4.5.

BAB II

TINJAUAN PUSTAKA

2.1. Teori Dasar

Penulis akan membahas beberapa teori yang menjadi pedoman dasar dalam melaksanakan penelitian, yaitu dengan menjelaskan mengenai pengertian dan istilah yang berkaitan dengan judul penelitian.

2.1.1. Sistem Pendukung Keputusan (SPK)

Menurut (Sri Lestari & Saefudin, 2015), Sistem pendukung keputusan adalah bagian dari sistem informasi berbasis komputer termasuk sistem berbasis pengetahuan atau manajemen pengetahuan yang dipakai untuk mendukung pengambilan keputusan dalam suatu organisasi atau perusahaan.

2.1.2. KDD (*Knowledge Discovery In Database*)

KDD memiliki kepanjangan *Knowledge Discovery In Database* dan KDD adalah salah satu cara dalam mendapatkan dan pengetahuan dengan memanfaatkan data yang berasal dari *database* atau data hasil penyimpanan. Setelah menemukan pengetahuan tersebut, maka terakhir digunakan sebagai dasar pengetahuan dalam melakukan pengambilan keputusan. Demikian tahapan dalam KDD, yaitu:

1. *Data Selection*

Tahap ini dilakukan pada awal proses KDD yaitu dengan melakukan pengumpulan informasi dengan melalui proses seleksi data untuk digunakan sebagai sumber data yang pada akhirnya akan diolah pada tahap *data mining*.

2. *Pre-processing* atau *Cleaning*

Tahap ini dilakukan dengan tujuan menghilangkan beberapa duplikasi yang terdapat pada data, dengan memeriksa yang tidak sesuai dengan kebutuhan karena proses penambahan data harus sesuai dengan kebutuhan penelitian.

3. *Transformation*

Proses perubahan dan penyesuaian terhadap pola penyimpanan basis data.

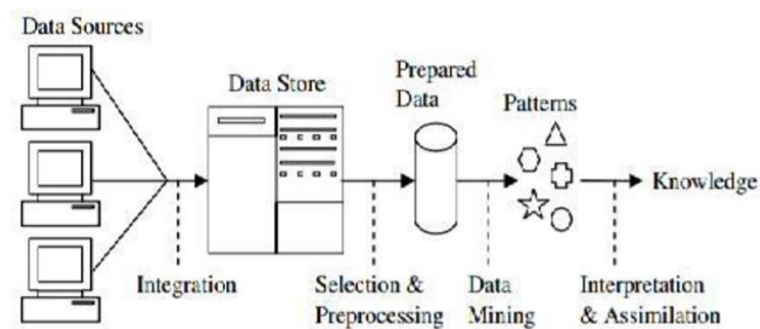
4. *Data Mining*

Proses yang menjadi cara dalam mencari dan membentuk pola aturan untuk menghasilkan informasi berupa keputusan berdasarkan tujuan penelitian.

5. *Interpretation*

Menampilkan hasil pola aturan dari *data mining* agar dapat dipahami, khususnya informasi yang bertentangan dengan hipotesis penelitian (Mardi, 2016).

Sementara itu menurut (Mardi, 2016) proses KDD atau *Knowledge Discovery in Database*, yaitu:



Gambar 2.1 Proses Dalam KDD atau *Knowledge Discovery in Database*

Sumber: (Mardi, 2016)

2.1.3. Data Mining

Apabila dilakukan pembagian berdasarkan tahapan berdasarkan tugas dan proses yang dilakukan dalam data mining dapat dibagi menjadi beberapa tahapan dan bagian, yaitu:

1. *Description* atau Deskripsi

Dalam tahap ini, bermula ketika orang yang sedang melakukan kegiatan analisa berusaha untuk mengelompokkan jenis pola yang terdapat dalam sumber data dan informasi. Hal ini dapat digambarkan seperti pada saat seorang petugas yang bertanggung jawab dalam melakukan kegiatan penghitungan pengumpulan suara dalam kegiatan pemilihan tidak mengetahui atau mendapatkan keterangan yang jelas mengenai baik atau tidaknya kemampuan serta keahlian yang dimiliki oleh calon yang terdaftar dalam pemilihan. Sehingga deskripsi yang dimaksud adalah upaya dalam memberikan beberapa penjelasan akan kemungkinan mengenai kemampuan dari calon yang terdaftar dalam pemilihan tersebut dengan suatu pola.

2. *Estimation* atau Estimasi

Selanjutnya, dalam tahapan atau proses ini kegiatan yang dilakukan dari estimasi memiliki cara kerja yang mirip dengan proses atau tahap dari klasifikasi, namun dapat berbeda apabila variabel dari target yang terdapat dalam proses atau tahapan dari estimasi berisi data yang bersifat numerik. Pola yang dibuat berasal dari rekaman data dengan berisi nilai atau numerik dari variabel yang terdapat dalam untuk dijadikan bahan dalam memberi nilai untuk proses dari prediksi. Kemudian dilanjutkan dengan pemberian nilai dari estimasi yang berasal dari nilai yang terdapat dalam variabel yang bersumber dari nilai dalam prediksi. Hal tersebut

dapat digambarkan dengan nilai yang terdapat dalam IPK atau Indeks Prestasi Kumulatif mahasiswa yang sedang menempuh program pendidikan strata satu (S.1) dapat diketahui nilai estimasinya berdasarkan hasil nilai dari IPK atau Indeks Prestasi Mahasiswa pada saat proses pendidikan setiap semesternya.

3. *Prediction* atau Prediksi

Dalam proses atau tahapan dari prediksi juga hampir mirip dengan proses dan tahapan pada saat melakukan klasifikasi dan juga pada saat melakukan proses estimasi, namun tidak termasuk apabila prediksi dari nilai yang ada diperuntukkan untuk masa yang akan datang. Prediksi yang dimaksud khususnya berkaitan dalam kegiatan bisnis dan juga penelitian, yaitu:

- a. Pada saat melakukan prediksi untuk mengetahui kemungkinan terhadap harga dari beras selama beberapa bulan yang akan datang.
- b. Pada saat memberikan prediksi mengenai kemungkinan yang terjadi mengenai pengangguran selama beberapa tahun kedepan.
- c. Pada saat membuat prediksi dalam bentuk persentase yang membahas mengenai kemungkinan akan kecelakaan yang akan terjadi apabila pengendara mengemudi kendaraannya melawati batas kecepatan yang disarankan selama beberapa tahun kedepan.

4. *Classification* atau Klasifikasi

Dalam proses atau kegiatan saat klasifikasi, digunakan beberapa target dalam variabel berdasarkan kategori yang ada. Hal tersebut dapat digambarkan seperti pada saat melakukan pengelompokkan mengenai penghasilan yang dimiliki oleh pekerja dalam sebuah perusahaan, pengelompokkan tersebut dilakukan dengan

menjadikan penghasilan pekerja memiliki tiga kategori, yaitu pekerja dari kategori penghasilan yang sangat tinggi, pekerja dari kategori tingkat penghasilan yang sedang atau standar sedangkan pekerja dari kategori tingkat penghasilan yang sangat rendah. Kemudian apabila klasifikasi dijelaskan dengan kaitannya dengan kegiatan penelitian dan dalam bisnis, yaitu:

- a. Pada saat melakukan proses klasifikasi terdapat transaksi kredit yang dilakukan dengan curang atau tidak.
- b. Pada saat melakukan prakiraan terhadap rencana kredit dilakukan oleh nasabah dengan tindakan yang tepat atau tidak.
- c. Pada saat dokter melakukan diagnosa terhadap penyakit atau kelainan yang dialami dan dimiliki pasien berdasarkan gejala yang ada.

5. *Clustering* atau Pengklusteran

Proses dari tahapan yang terdapat dalam pengklusteran adalah kegiatan atau proses melakukan pengelompokan hasil perekaman data yang kemudian dikeompokkan lagi berdasarkan data yang memiliki kesamaan atau kemiripan pola. Kluster yang dimaksud memiliki hasil perekaman data yang memiliki suatu kemiripan atau kesamaan pola dalam satu kelompok dan kelompok lainnya tidak memiliki kesamaan atau kemiripan dengan hasil perekaman yang ada. Dalam proses atau tahapan yang terdapat dalam pengklusteran berbeda dengan proses atau tahapan dari klasifikasi karena dalam pengklusteran tidak terdapat variabel dari target. Cara kerja yang dilakukan dalam tahapan dan proses dalam pengklusteran tidak melakukan proses yang terdapat pada saat melakukan estimasi, prediksi dan juga klasifikasi karena perhitungan dari algoritma yang digunakan dalam proses

pengklusteran fokus melakukan pengelompokkan data yang ada berdasarkan kesamaan pola yang dimiliki oleh data dan untuk pemberian nilai yang maksimal berasal atau sesuai dengan tingkat kemiripan atau kesamaan data yang dimiliki dalam kelompok begitu juga berlaku dengan sebaliknya apabila tingkat kesamaan atau kemiripan data yang ada sangat rendah maka hasil dari nilai tersebut menjadi minimal. Hal tersebut dapat digambarkan, sebagai berikut:

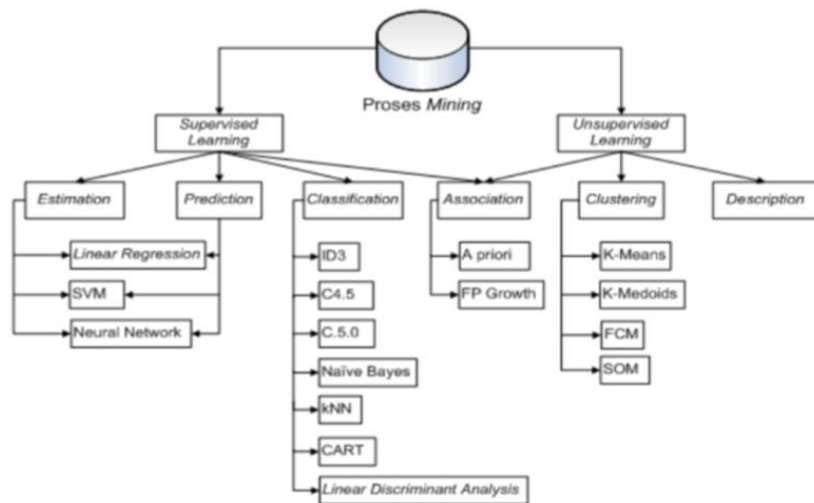
- a. Pada saat melakukan pengelompokkan terhadap jenis-jenis dari konsumen yang dijadikan target dari kegiatan pemasaran dari sebuah produk yang dilakukan suatu perusahaan.
- b. Pada saat melakukan kegiatan audit, dilakukan pemisahan berdasarkan transaksi yang diduga mencurigakan dan transaksi yang diduga aman.

6. *Association* atau Asosiasi

Asosiasi melakukan prosesnya untuk mendapatkan kelengkapan atau keterangan pendukung yang terdapat dalam sebuah data. Hal tersebut dapat digambarkan melalui bentuk kegiatan yang dilakukan, yaitu:

- a. Pada saat melakukan penelitian atau analisa mengenai jumlah pelanggan dari sebuah perusahaan yang memberikan respon secara positif terhadap pelayanan yang diberikan perusahaan.
- b. Untuk mengetahui keseluruhan jenis barang yang ada dipasar namun belum pernah dibeli atau dikonsumsi sebelumnya (Mardi, 2016).

Penjelasan dalam bentuk gambar mengenai hasil pengelompokkan beberapa proses atau tahapan berdasarkan tugas saat seseorang peneliti melakukan penelitian menggunakan teknik dengan *data mining*, sebagai berikut.



Gambar 2.2 Pengelompokan Teknik *Data Mining*
Sumber: (Hermawan, Sukma, & Halfis, 2019)

2.1.4. *Decision Tree*

Akibat dari berbagai jenis permasalahan dan tingkat kesulitan dalam permasalahan yang berbeda dan sering terjadi dalam kehidupan manusia pada akhirnya membuat manusia mulai melakukan berbagai upaya dalam mengembangkan sebuah pola berbentuk sistem agar dapat membantu mereka dalam menyelesaikan berbagai jenis dan bentuk permasalahan yang terjadi dan salah satu hasilnya adalah *decision tree* atau biasa dikenal dengan istilah pohon keputusan. *Decision tree* merupakan salah satu bentuk implementasi pola dengan menerapkan berbagai tahapan dan proses yang terdapat dalam kegiatan klasifikasi dan juga kegiatan prediksi. *Decision tree* menghasilkan alternatif dalam memecahkan masalah dengan hasil yang kuat dan berguna sehingga *decision tree* terkenal dan banyak digunakan oleh peneliti saat melakukan penelitian dengan teknik *data mining*. *Decision tree* sangat bermanfaat apabila digunakan untuk keperluan dalam mengubah data yang ada menjadi sebuah pola berbentuk aturan sehingga dapat

dipahami dengan mudah oleh manusia. Dalam tahapannya, *decision tree* melakukan proses yang dimulai dari keseluruhan bagian-bagian yang terdapat dalam data yang berisi berbagai kondisi dan keterangan yang lengkap dan kemudian diakhiri dengan sebuah keputusan. Arsitektur dari *decision tree* memang dibuat mirip dengan struktur yang terdapat pada sebuah pohon dan penjelasannya, yaitu:

1. Simpul bagian akar

Simpul dari bagian akar terletak pada bagian paling atas dan pertama dari struktur yang terdapat pada *decision tree*.

2. Simpul internal

Simpul dari bagian internal ini adalah percabangan lanjut dari simpul bagian akar. Dalam sebuah simpul ini apabila terdapat satu masukan maka dapat mengeluarkan keluaran dengan jumlah maksimal sebanyak dua.

3. Simpul daun

Simpul daun merupakan simpul yang berada pada ujung bagian dari pohon. Dalam simpul daun terdapat sebuah masukan namun tidak memiliki keluaran.

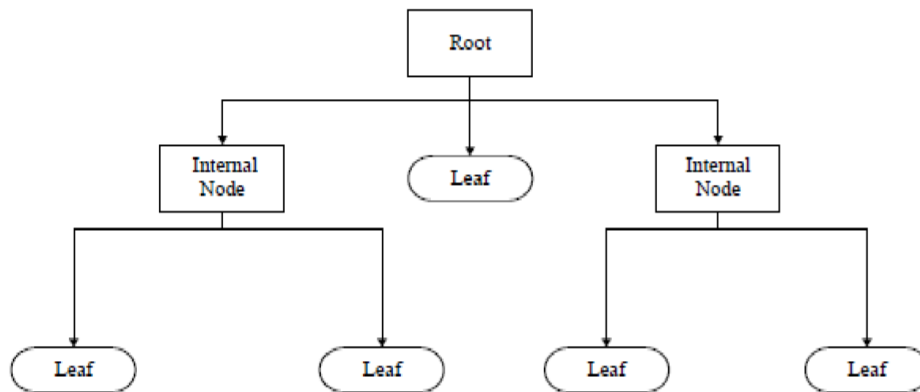
Selanjutnya, demikian penjelasan dari kelebihan dan kekurangan yang dimiliki oleh *decision tree*, yaitu:

1. Kelebihan dari *decision tree*

- a. Dalam penggunaan *decision tree* apabila terdapat cakupan data yang rumit maka dirubah menjadi lebih sederhana untuk kemudian dijadikan dasar pengambilan keputusan.

- b. Kelebihan dari penggunaan *decision tree* yaitu dapat menghapus cara perhitungan yang tidak perlu sehingga dalam melakukan proses dari pengujian hanya menggunakan kriteria sesuai dengan keperluan untuk pengujian saja.
 - c. Dalam proses pemilihan akan kelengkapan atau keterangan pada data, bila menggunakan *decision tree* maka hasil pemilihan terhadap kelengkapan atau keterangan data tersebut akan menjadi pembeda dalam mengelompokkan data berdasarkan kriteria lainnya.
 - d. *Decision tree* bertindak untuk menghindari permasalahan lain dengan memanfaatkan menggunakan jenis kelompok atau kategori dengan jumlah sedikit tanpa mengurangi kualitas hasil keputusan akhir.
2. Kekurangan dari *decision tree*
- a. Dalam menggunakan *decision tree* apabila jumlah dari hasil keputusan yang harus dihasilkan dan kriteria yang digunakan memiliki jumlah yang sangat banyak, maka waktu yang dibutuhkan dalam menghasilkan keputusan akan semakin bertambah serta dibutuhkan memori sistem pendukung yang tinggi.
 - b. Dalam *decision tree* terdapat akumulasi dari jumlah *error* yang terjadi dalam setiap bagian dari struktur yang terdapat pada *decision tree*.
 - c. Membuat struktur dari *decision tree* dengan hasil yang maksimal ternyata sulit untuk dilakukan.

- d. Apabila menggunakan *decision tree*, maka hasil atau kualitas dari keputusan yang dihasilkan sangat ditentukan dari model atau struktur *decision tree* yang dibuat (Harryanto & Hansun, 2017).



Gambar 2.3 Model *Decision Tree*
Sumber: (Takalapeta, 2018)

2.1.5. *Waikato Environment for Knowledge Analysis (WEKA)*

Waikato Environment for Knowledge Analysis atau biasanya disingkat WEKA adalah paket tools machine learning praktis, *Waikato Environment for Knowledge Analysis (WEKA)* ini dikembangkan di Universitas Waikato yang terletak di Selandia Baru untuk penelitian, pendidikan dan berbagai aplikasi (Pujiono, Amborowati, & Suyanto, 2013). WEKA merupakan software yang bersifat open source, jadi siapapun dapat ikut menambahkan perangkat lunak ini (Haryati, Sudarsono, & Suryana, 2015).

WEKA dapat digunakan untuk menyelesaikan masalah-masalah datamining di dunia nyata terutama klasifikasi yang menggunakan machine learning sebagai pendekatannya. WEKA mengandung *tools pre-processing data*, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. Pengujian dengan menggunakan WEKA

lebih fleksibel dan efektif karena pohon keputusan yang terbentuk hampir sama dengan perhitungan secara manual (Azwanti, 2018).



Gambar 2.4 Tampilan Utama WEKA

2.1.6. Algoritma C4.5

Algoritma dari C4.5 bermanfaat dalam melakukan proses klasifikasi data yang menggunakan angka atau yang menggunakan kategori. Setelah tahapan dari proses klasifikasi dilakukan maka kemudian hasilnya adalah beberapa aturan dari pola dan dapat dipergunakan untuk melakukan tahapan dari proses prediksi dengan nilai-nilai prediksi yang berasal dari perekaman data yang lama hingga terbaru. Jenis algoritma dari C4.5 berasal dari algoritma ID3, maksudnya adalah algoritma dari C4.5 adalah hasil dari perbaikan lebih lanjut yang dilakukan terhadap algoritma dari ID3 sehingga dalam algoritma C4.5 memiliki kemampuan dalam mengatasi permasalahan saat terjadi ketiadaan data dan mengatasi permasalahan saat terjadi data yang bersifat pecahan. Pada umumnya algoritma dari C4.5 digunakan dalam

proses pembangunan struktur dari decision tree, dan demikian penjelasan proses atau tahapannya yaitu:

1. Pertama adalah memilih kelengkapan atau keterangan data sebagai simpul akar dari *decision tree*.
2. Kemudian dilanjutkan dengan membuat struktur cabang dari *decision tree* yang sudah berisi nilai-nilai.
3. Pada tahap ketiga, dilakukan pembagian atau pengelompokkan dari permasalahan yang terdapat pada struktur cabang dari *decision tree*.
4. Dan tahap akhir adalah melakukan proses pengulangan yang ada pada tahapan pertama hingga ketiga sampai seluruh cabang pada struktur *decision tree* memiliki kelas dan kategori yang sama.

Dalam melakukan tahap awal dari *decision tree* yaitu pada saat melakukan pemilihan terhadap kelengkapan atau keterangan data dilakukan dengan cara mencari nilai dari gain dengan hasil yang tertinggi dari keluruhan nilai yang dimiliki oleh kelengkapan atau keterangan yang ada. Demikian rumus untuk melakukan pencarian nilai gain tertinggi, yaitu:

$$Gain (A) = Entropi (S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi (S_i)$$

Rumus 2.1 Perhitungan *Gain*

Keterangan simbolnya, yaitu:

S adalah himpunan

A adalah keterangan atau kelengkapan yang digunakan.

n adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian *A*

$|S_i|$ adalah jumlah dari kasus yang terdapat dalam partisi yang ke-*i*

$|S|$ adalah jumlah dari kasus yang terdapat dalam S

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

Rumus 2.2 Perhitungan *Entropy*

Keterangan simbolnya, yaitu:

S adalah himpunan

A adalah keterangan atau kelengkapan yang digunakan.

n adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian S

p_i adalah proporsi dari S_i kepada S (Elisa, 2017).

2.1.7. Kinerja Karyawan

Kinerja karyawan adalah kuantitas dan atau kualitas pekerjaan yang dihasilkan oleh karyawan Toko Umamo Vero seperti standar hasil kerja, target yang ditentukan selama periode tertentu yang berpedoman pada norma, standar operasional prosedur, kriteria dan fungsi yang telah ditetapkan atau yang berlaku dalam perusahaan. Indikator yang digunakan untuk mengukur kinerja karyawan (Fransisca Andreani & Tanto Wijaya, 2015) adalah:

1. Jumlah pekerjaan yang di hasilkan
2. Kualitas pekerjaan yang di hasilkan
3. Ketepatan waktu menyelesaikan pekerjaan
4. Kehadiran pada hari dan jam kerja
5. Kemampuan kerja sama

2.1.8. Reward

Reward adalah usaha menumbuhkan perasaan diterima (diakui) di lingkungan kerja, yang menyentuh aspek kompensasi dan aspek hubungan antara para pekerja yang satu dengan yang lainnya. Manajer mengevaluasi hasil kinerja individu baik secara formal maupun informal (Dhamayanti, 2018).

2.1.9. Bonus

Menurut (Agusli, Dzulhaq, & Khasanah, 2017), Bonus adalah pendapatan tambahan yang berfungsi sebagai pendorong semangat agar karyawan semakin bergairah dalam meningkatkan prestasi kerja dan loyalitas pada perusahaan. Adapun tujuan program bonus yang diberikan perusahaan, sebagai berikut:

1. Bagi Perusahaan

Tujuan pemberian bonus dimaksudkan untuk meningkatkan kinerja dengan cara mendorong karyawan agar bekerja disiplin dan memiliki semangat yang lebih tinggi. Diharapkan dengan pemberian bonus ini dapat mencapai tujuan dalam menghasilkan kualitas kinerja karyawan yang lebih baik, meningkatkan efektivitas dan efisiensi dalam penggunaan faktor kinerja, serta mencegah terjadinya penurunan kinerja karyawan.

2. Bagi Karyawan

Dengan pemberian bonus dari perusahaan, diharapkan karyawan memperoleh banyak keuntungan. Keuntungan yang merupakan komponen non-upah ini, misalnya mendapatkan gaji yang lebih besar di akhir tahun, mendapat

dorongan untuk mengembangkan dirinya, berusaha bekerja dengan sebaik-baiknya, dan loyalitas terhadap perusahaan.

2.1.10. Bonus Tahunan

Bonus merupakan segala sesuatu yang diterima dapat berupa fisik maupun non fisik dan harus dihitung dan diberikan kepada karyawan yang berhak menerimanya, sistem bonus yang baik akan mampu memberikan kepuasan bagi karyawan dan memungkinkan toko/perusahaan memperoleh penghasilan yang banyak tiap bulannya (Afriany & Aisyah, 2019).

2.1.11. Karyawan

Karyawan merupakan faktor pendukung dalam sebuah perusahaan atau instansi, karena dengan adanya karyawan yang memiliki standar kualifikasi perusahaan maka produktivitas perusahaan pasti akan tetap terjaga dan semakin meningkat(Afriany & Aisyah, 2019).

2.1.12. Insentif

Insentif adalah suatu penghargaan dalam bentuk uang yang diberikan oleh pihak pemimpin organisasi kepada karyawan agar mereka bekerja dengan motivasi yang tinggi dan berprestasi dalam mencapai tujuan-tujuan organisasi atau dengan kata lain, insentif kerja merupakan pemberian uang diluar gaji yang dilakukan oleh pihak pemimpin organisasi sebagai pengakuan terhadap prestasi kerja dan kontribusi karyawan kepada organisasi menurut (Sahir & Panjaitan, 2019).

2.2. Penelitian Terdahulu

Peneliti juga menemukan beberapa hasil penelitian terdahulu yang berkaitan dengan menganalisa menggunakan algoritma C4.5 yang dirangkum yaitu:

1. Dari penelitian “Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5” yang dilakukan oleh (Sri Lestari & Saefudin, 2015) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membangun sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5 dengan kriteria-kriteria yang telah ditentukan oleh perusahaan.
2. Dari penelitian “Komparasi Algoritma C4.5 Dan *Naive Bayes* Yang diKembangkan Menjadi *Web Intelligence* Pada Perhitungan Bonus Tahunan Karyawan di PT.ABC” yang dilakukan oleh (Ibnu Alfarobi & Taransa Agasya Tutupoly, 2019) didapatkan informasi bahwa bonus karyawan dari kedua algoritma *Decision Tree* C4.5 dan *Naive Bayes* yang dimana memiliki nilai *Area Under Curve (AUC)* sebesar 0.991 dan *Accuracy* atau *Confusion Matrix* sebesar 99%.
3. Dari penelitian “Penerapan Algoritma C4.5 Dalam Analisa Kelayakan Penerima Bonus Tahunan Pegawai (Studi Kasus : PT. Multi Pratama Nauli Medan)” yang dilakukan oleh (Syahputra, M.Kom, 2015) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membantu seorang pimpinan perusahaan dalam menganalisa pegawai yang layak untuk menerima bonus tahunan.

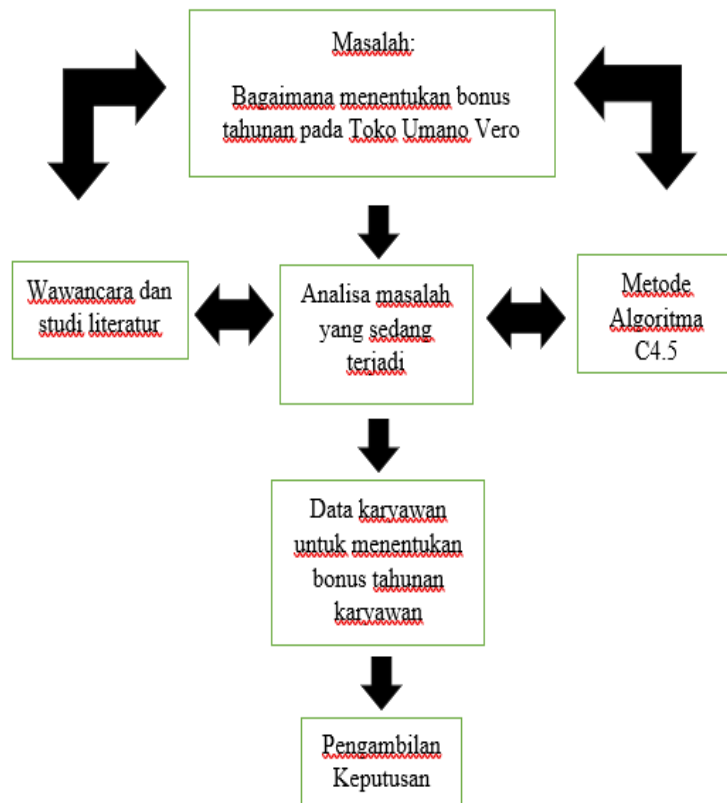
4. Dari penelitian “Penerapan Algoritma C4.5 Pada Analisis Penentuan Kelayakan Penerima Bonus Tahunan Pegawai Bank Sinarmas Multifinance” yang dilakukan oleh (Intan Komala Dewi Patwari & Baiq Andriskha Candra Permana, 2019) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membantu seorang pemimpin dalam mengambil suatu keputusan apakah seorang pegawai layak atau tidak untuk diberi bonus.
5. Dari penelitian “Sistem Pendukung Keputusan Penentuan Insentif Bulanan Pegawai Dengan Menggunakan Metode *Naïve Bayes*” yang dilakukan oleh (Victor Marudut Mulai Siregar, 2018) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat mempermudah pihak manager Edene Sayangku *Cafe & Bakery* dalam menghitung insentif pegawai dengan tepat dan akurat.
6. Dari penelitian “*Taiga: Performance Optimization of the C4.5 Decision Tree Construction Algorithm*” yang dilakukan oleh (Yang & Chen, 2016) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 yang terdapat pada decision tree dapat dimaksimalkan dengan melakukan penyelasaran data dengan kategori atau kriteria dalam melakukan *data mining* sehingga waktu yang dibutuhkan untuk mendapatkan keputusan akhir menjadi lebih singkat.
7. Dari penelitian “*Performance improvement of datamining ini WEKA through GPU acceleration*” yang dilakukan oleh (Augusto & Char, 2014) didapatkan informasi bahwa proses yang memakan waktu banyak dapat

mengadaptasi terhadap *GPU*. *show that set of operations that are time consuming can easily adapted to GPUs*. Ketika memparalelkan metode penggandaan matriks, metode ini dapat beradaptasi untuk memanfaatkan *CPU* dan *GPU* sesuai dengan ukuran matriks, metode ini mempercepat proses hingga 49%.

8. Dari penelitian “Penerapan Algoritma C.45 Untuk Kelayakan Penerima Uang Dan Beras” yang dilakukan oleh (Harman, 2018) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 diharapkan dapat memperbaiki kebijakan-kebijakan yang selama ini dianggap tidak adil atau faktor kedekatan lebih berperan besar untuk penerima bantuan untuk uang dan beras kedepannya setelah penelitian ini selesai dilakukan.
9. Dari penelitian “Analisis Dengan Metode Klasifikasi Menggunakan *Decision Tree* Untuk Prediksi Penentuan Resiko Kredit Pada Bank Bukopin Batam” yang dilakukan oleh (Intan Utma Sari, 2015) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 diharapkan dapat membantu pihak bank agar lebih cepat dan mudah dalam memprediksi data dan menarik suatu kesimpulan dari data yang ada. Salah satu cara memprediksi data tersebut dengan menggunakan *software Dtree* dan *Decision Tree*. *Dtree* dapat menghasilkan pohon keputusan, salah satu nya yaitu hasil keputusan resiko dari jumlah kredit pemilikan rumah berdasarkan jumlah nasabah sehingga dan dengan teknik *Decision Tree* diharapkan dapat membantu pihak bank agar lebih cepat dan mudah dalam memprediksi data dan menarik suatu kesimpulan dari data yang ada.

2.3. Kerangka Pemikiran

Berikut adalah kerangka pemikiran dalam penelitian ini:



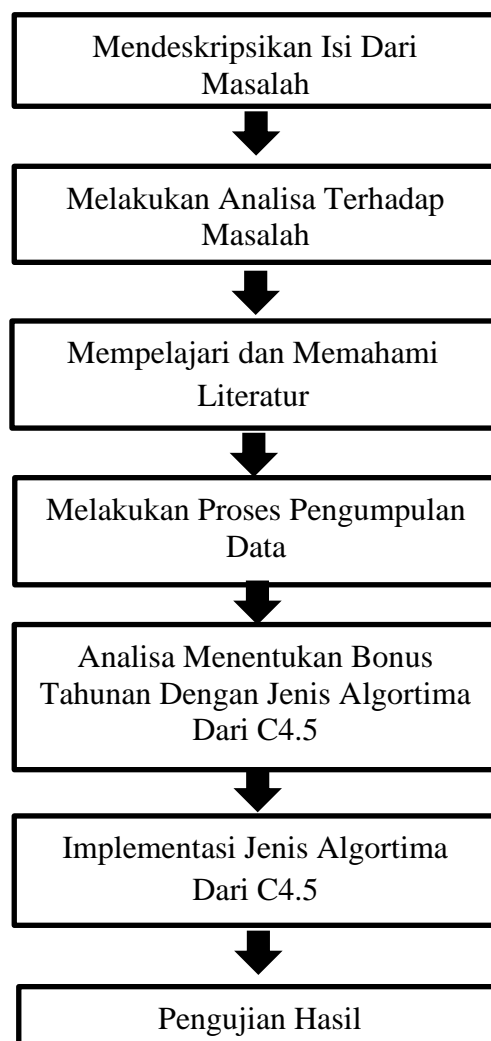
Gambar 2.5 Kerangka Berpikir
(Sumber: Peneliti, 2020)

BAB III

METODE PENELITIAN

3.1. Desain Penelitian

Berikut gambaran desain penelitian:



Gambar 3.1 Desain Penelitian
(Sumber: Penulis,2020)

Demikian penjelasan berdasarkan desain dari penelitian sesuai dengan model atau gambar diatas, yaitu:

1. Mendeskripsikan isi dari masalah

Penulis berupaya memberikan penjelasan permasalahan yang akan diteliti, berkaitan dengan kepuasan pelanggan dengan jenis algoritma dari C4.5.

2. Melakukan analisa terhadap masalah

Tahap ini merupakan kegiatan lanjutan setelah mendeskripsikan masalah. Penulis berupaya dalam melakukan analisa berdasarkan beberapa permasalahan yang ingin diangkat dan diteliti oleh penulis.

3. Mempelajari dan memahami literatur

Penulis berupaya mempelajari dan memahami beberapa sumber teori yang mendukung penelitian dari jurnal untuk dijadikan pedoman bagi penulis dalam melakukan penelitian.

4. Melakukan proses pengumpulan data

Pada tahap ini, peneliti melakukan wawancara dengan mendatangi objek penelitian dan melakukan tanya jawab kepada pemilik dan pekerja di Toko Umamo Vero untuk mendapatkan lebih banyak data dan informasi yang dibutuhkan.

5. Analisa menentukan bonus tahunan dengan jenis algoritma dari C4.5

Penggunaan jenis algoritma dari C4.5 mulai dari pada saat proses pembuatan struktur *decision tree* hingga menghasilkan pola berupa aturan dari menentukan bonus tahunan.

6. Implementasi jenis algoritma dari C4.5

Penulis menentukan kelengkapan data yang dijadikan sebagai akar dalam *decision tree* dan kemudian dilakukan perhitungan nilai gain tertinggi dari keseluruhan kelengkapan yang terdapat pada data yang selanjutnya digunakan juga dalam proses membentuk struktur *decision tree*.

7. Pengujian Hasil

Pada tahapan terakhir ini, peneliti melakukan pengujian setelah selesai melakukan perhitungan terhadap nilai gain yang tertinggi dan membentuk *decision tree* dengan menggunakan bantuan aplikasi WEKA versi 3.9.2.

3.2. Objek Penelitian

Pada sub bab objek penelitian, penulis menjelaskan keterangan lokasi penelitian dan penjelasan jadwal selama kegiatan penelitian, sebagai berikut.

3.2.1.Lokasi Penelitian

Penelitian ini berlokasi di Toko Umamo Vero dengan alamat Jl. Komp. Bumi Indah No.45 & 47, Lubuk Baja Kota, Kec. Lubuk Baja, Kota Batam, Kepulauan Riau.

3.2.2.Jadwal Penelitian

Demikian penjelasan penulis mengenai rencana waktu yang akan dilakukan dalam melaksanakan kegiatan analisa mengenai menentukan bonus tahunan pada Toko Umamo Vero, sebagai berikut.

Tabel 3.1 Jadwal Penelitian

No	Kegiatan	Mar-20	Apr-20	Mei-20	Jun-20	Jul-20
1	Bimbingan					
2	Identifikasi Masalah					
3	Studi Literatur					
4	Penentuan Judul					
5	Pengumpulan Data					
6	Pengolahan Data					
7	Pengujian Hasil					
8	Penyusunan Laporan					

Sumber: (Penulis,2020)

3.3. Populasi dan Sampel

Dalam sub bab ini penulis akan menjelaskan mengenai keterangan yang dimiliki oleh populasi dan sampel yang dijadikan bahan dalam pelaksanaan kegiatan penelitian sebagai berikut.

3.3.1. Populasi

Populasi pada penelitian ini adalah data kinerja karyawan di Toko Umamo Vero periode Januari 2019 sampai dengan Nopember 2019.

3.3.2. Sampel Jenuh

Menurut (Fitria & Ariva, 2018) teknik sampling jenuh adalah teknik penentuan sampel dimana semua anggota populasi digunakan menjadi sampel.

Data yang akan dijadikan sampel pada penelitian ini, peneliti akan mengambil data kinerja karyawan sebanyak 50 data sampel dari populasi data yaitu data periode Januari 2019 sampai dengan Nopember 2019.

3.4. Variabel Penelitian

Untuk menganalisis data dalam sebuah penelitian memerlukan variable-variabel, data yang akan diolah adalah data kinerja karyawan periode Januari 2019 sampai Nopember 2019 dari Toko Umano Vero. Beberapa indikator yang digunakan dalam penelitian khususnya yang berkaitan dengan variabel untuk menentukan bonus tahunan, sebagai berikut:

1. Tangung Jawab
2. Target
3. Kehadiran
4. Masa Kerja
5. Loyalitas
6. Inisiatif

Untuk variabel keputusannya adalah layak dan tidak layak.

3.5. Teknik Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan untuk mencapai tujuan penelitian, atau mengungkap informasi atau kondisi lokasi penelitian sesuai dengan kebutuhan penelitian. Teknik pengumpulan data adalah sebagai berikut:

1. Teknik studi literatur adalah pembahasan literatur pada bidang tertentu dari sebuah penelitian. Studi ini adalah gambaran singkat dari apa yang telah dipelajari, argumentasi, dan ditetapkan tentang suatu topik, dan biasanya diorganisasikan secara kronologis atau tematis. Dalam hal ini, peneliti

mempelajari berbagai jurnal yang memiliki keterkaitan terhadap permasalahan yang dihadapi.

2. Teknik observasi adalah teknik pengumpulan data dengan cara peneliti melakukan pengamatan secara langsung dilapangan, dalam hal ini peneliti terjun langsung ke Toko Umamo Vero untuk mengamati bagaimana proses kinerja karyawan pada Toko Umamo Vero.
3. Teknik dokumentasi adalah peneliti mempelajari sumber-sumber tertulis yang ada pada lokasi penelitian yang menyangkut masalah data yang telah ada sebelumnya, atau melakukan proses penggalian terhadap data yang sudah ada tentang dan kemudian mengelompokkan data tersebut sesuai dengan kebutuhan dari penelitian yang akan dilakukan. Dalam hal ini, peneliti mengamati dan mempelajari data-data kinerja karyawan yang ada di Toko Umamo Vero.
4. Teknik wawancara adalah peneliti melakukan proses tanya jawab dalam penelitian yang berlangsung secara lisan dalam mana dua orang atau lebih bertatap muka mendengarkan secara langsung informasi-informasi atau keterangan-keterangan. Dalam hal ini, peneliti melakukan tanya jawab terhadap pemilik dan pekerja di Toko Umamo Vero.

3.6. Metode Penelitian

Penulis melakukan jenis model penelitian menggunakan data mining dan jenis algoritma dari C4.5. Tahapan awal dilakukan klasifikasi sesuai dengan 50 kategori atau kriteria data kinerja karywan, selanjutnya dilakukan perhitungan

algoritma dari C4.5 untuk mendapatkan nilai *entropy* dan *gain* tertinggi pada saat membuat struktur dari *decision tree* dilakukan secara berulang hingga simpul pada *decision* tidak memiliki turunan lagi sampai didapatkan pola berupa aturan dasar penilaian mengenai menentukan bonus tahunan pada Toko Umano Vero.