

**PENERAPAN ALGORITMA C4.5 UNTUK
MENENTUKAN REWARD KARYAWAN PADA
PT INDOLAND BATAM**

SKRIPSI



**Oleh:
Vincent Kiu Kok Seng
171510014**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

**PENERAPAN ALGORITMA C4.5 UNTUK
MENENTUKAN REWARD KARYAWAN PADA
PT INDOLAND BATAM**

SKRIPSI

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



**Oleh
Vincent Kiu Kok Seng
171510014**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Vincent Kiu Kok Seng
NPM/NIP : 171510014
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

Penerapan Algoritma C4.5 Untuk Menentukan reward Karyawan Pada PT Indoland Batam

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.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 22 Januari 2021



Vincent Kiu Kok Seng
171510014

**PENERAPAN ALGORITMA C4.5 UNTUK MENENTUKAN
REWARD KARYAWAN PADA PT INDOLAND BATAM**

Oleh
Vincent Kiu Kok Seng
171510014

SKRIPSI

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

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

Batam, 22 Januari 2021



Rika Harman, S.Kom., M.SI.
Pembimbing



Universitas Putera Batam

ABSTRAK

Dalam menghadapi persaingan yang ketat akibat perubahan teknologi dan perubahan lingkungan di berbagai bidang kehidupan, semua perusahaan mencari karyawan yang berpotensi untuk memberikan layanan terbaik dan bermanfaat. Dengan cara ini, perusahaan tidak hanya dapat memberikan layanan yang memuaskan (kepuasan pelanggan), tetapi juga mewujudkan nilai (nilai pelanggan). Perusahaan yang beroperasi tidak hanya mengejar produktivitas tetapi juga kinerja. PT Indoland Batam merupakan sebuah perusahaan yang bergerak di bidang penjualan *online* atau *online shop* yang menyediakan barang-barang seperti peralatan kebutuhan rumah tangga, tas, sepatu, alat kecantikan, pakaian, dan lain sebagainya. Perusahaan ini berdiri sejak tahun 2017 yang beralamat di Citra Buana Park 1, Blok H No 2, Jalan Yos Sudarso, Kampung Seraya Batam. Dalam menentukan *reward* 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 *reward* karyawan. Seiring berlewatnya tahun sering kali karyawan pada PT Indoland Batam tidak puas dengan *reward* yang mereka terima sehingga masalah ini menimbulkan kurangnya kinerja karyawan, kedisiplinan karyawan, dan loyalitas karyawan terhadap PT Indoland Batam.

Kata Kunci: Algoritma C4.5, Analisis, Data Mining, *Reward*.

ABSTRACT

In the face of intense competition due to technological changes and environmental changes in various areas of life, all companies are looking for employees who have the potential to provide the best and useful services. In this way, the company can not only provide satisfactory service (customer satisfaction), but also embody value (customer value). Operating companies not only pursue productivity but also performance. PT Indoland Batam is a company engaged in online sales or an online shop that provides items such as household appliances, bags, shoes, beauty tools, clothes, and so on. This company was founded in 2017, which is located at Citra Buana Park 1, Blok H No 2, Jalan Yos Sudarso, Kampung Seraya Batam. In determining the reward for employees in the company, there are many obstacles or problems that the company must face, including time efficiency, many comparisons of the variables tested, and decision making to determine the employee's reward. As the years passed, employees at PT Indoland Batam were often dissatisfied with the reward they received, so this problem resulted in underperformance of employees, employee discipline, and employee loyalty to PT Indoland Batam.

Keywords: *Algorithm C4.5, Data Mining, Reward.*

KATA PENGANTAR

Dengan mengucapkan puji dan syukur kepada Tuhan yang Maha Esa yang telah melimpahkan segala rahmat dan karunianya, sehingga penulis dapat menyelesaikan skripsi yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

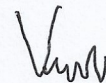
Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

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 Welly Sugianto, S.T., M.M, selaku Dekan Fakultas Teknik Dan Komputer Universitas Putera Batam.
3. Bapak Muhammad Rasyid Ridho, S.Kom., M.SI., selaku Ketua Program Studi Sistem Informasi Universitas Putera Batam.
4. Bapak Rika Harman, S.Kom., M.SI., sebagai pembimbing skripsi yang telah menyediakan waktu dalam memberikan bimbingan dan petunjuk sampai selesainya skripsi ini.
5. Seluruh dosen dan staff Universitas Putera Batam.
6. Dan pihak-pihak yang telah memberikan dukungannya hingga tersusunnya penelitian ini yang tidak dapat disebutkan oleh penulis satu persatu.

Semoga Tuhan YME membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 22 Januari 2021



Penulis



Universitas Putera Batam

DAFTAR ISI

HALAMAN SAMBUNG DEPAN.....	i
HALAMAN JUDUL.....	ii
SURAT PERNYATAAN.....	iii
HALAMAN PENGESAHAN.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
DAFTAR RUMUS	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah	5
1.3. Batasan Masalah.....	6
1.4. Perumusan Masalah.....	6
1.5. Tujuan Penelitian.....	7
1.6. Manfaat Penelitian.....	8
1.6.1. Manfaat Teoritis	8
1.6.2. Manfaat Praktis	8
BAB II TINJAUAN PUSTAKA.....	11
2.1. Teori Dasar	11
2.1.1. Teori Umum	11
2.1.2. Teori Khusus	22
2.2. Penelitian Terdahulu.....	26
2.3. Kerangka Pemikiran	30
BAB III METODE PENELITIAN.....	33
3.1. Desain Penelitian	33
3.2. Objek Penelitian	34
3.2.1. Lokasi Penelitian.....	34

3.2.2.	Jadwal Penelitian.....	34
3.3.	Populasi dan Sampel	35
3.3.1.	Populasi.....	35
3.3.2.	Sampel.....	35
3.4.	Variabel Penelitian	35
3.5.	Teknik Pengumpulan Data	36
3.6.	Metode Penelitian.....	37
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Hasil Penelitian.....	39
4.2	Pembahasan	39
4.2.1.	Melakukan Pra-Proses Data	39
4.2.2.	Analisa Data	41
4.2.3.	Pohon Keputusan	46
4.2.4.	Pengujian Hasil Dengan Manual.....	46
4.2.5.	Pengujian menggunakan WEKA.....	60
4.2.6.	Keputusan Hasil Pengujian	66
BAB V SIMPULAN DAN SARAN		70
5.1	Simpulan.....	70
5.2	Saran	70
DAFTAR PUSTAKA		72
LAMPIRAN.....		74

DAFTAR GAMBAR

	Halaman
Gambar 1.1. Data Bonus Karyawan PT Indoland Batam.....	3
Gambar 2.1. Proses Dalam KDD atau Knowledge Discovery in Database.....	12
Gambar 2.2. Bidang Ilmu Data Mining.....	12
Gambar 2.3. Pengelompokan Teknik Data Mining.....	17
Gambar 2.4. Model <i>Decision Tree</i>	20
Gambar 2.5. Tampilan Utama WEKA.....	27
Gambar 2.6. Kerangka Berpikir.....	31
Gambar 3.1. Desain Penelitian.....	32
Gambar 4.1. Pohon Keputusan Hasil Perhitungan Node 1.....	52
Gambar 4.2. Pohon Keputusan Hasil Perhitungan Node 2.....	56
Gambar 4.3. Pohon Keputusan Hasil Perhitungan Node 3.....	60
Gambar 4.4. Data Uji <i>reward.xls</i>	61
Gambar 4.5. Data Uji <i>weka.csv</i>	62
Gambar 4.6. Menu Awal <i>WEKA 3.8.3</i>	62
Gambar 4.7. Open File pada <i>WEKA Explorer</i>	63
Gambar 4.8. Pemilihan Variabel pada <i>WEKA</i>	63
Gambar 4.9. <i>Tab Classify</i> pada <i>WEKA</i>	64
Gambar 4.10. Pemilihan Model <i>Classify</i>	64
Gambar 4.11. <i>Classify Output</i> pada <i>WEKA</i>	65
Gambar 4.12. <i>Visualize Tree</i> Pada <i>WEKA</i>	66
Gambar 4.13. Hasil <i>Visualize Tree WEKA</i>	66

DAFTAR TABEL

	Halaman
Tabel 3.1. Jadwal Penelitian.....	34
Tabel 4.1. Format Data Kinerja Karyawan.....	39
Tabel 4.2. Klasifikasi Masa Kerja.....	42
Tabel 4.3. Klasifikasi Loyalitas.....	42
Tabel 4.4. Klasifikasi Kehadiran.....	42
Tabel 4.5. Klasifikasi Kerajinan.....	43
Tabel 4.6. Klasifikasi Sikap.....	43
Tabel 4.7. Klasifikasi Inisiatif.....	43
Tabel 4.8. Format Data Akhir Kinerja Karyawan.....	44
Tabel 4.9. Perhitungan Node 1 (manual).....	50
Tabel 4.10. Perhitungan Node 2 (manual).....	54
Tabel 4.11. Perhitungan Node 3 (manual).....	59

DAFTAR RUMUS

	Halaman
Tabel 2.1. Perhitungan Gain.....	21
Tabel 2.2. Perhitungan Entropy.....	22



Universitas Putera Batam

BAB I

PENDAHULUAN

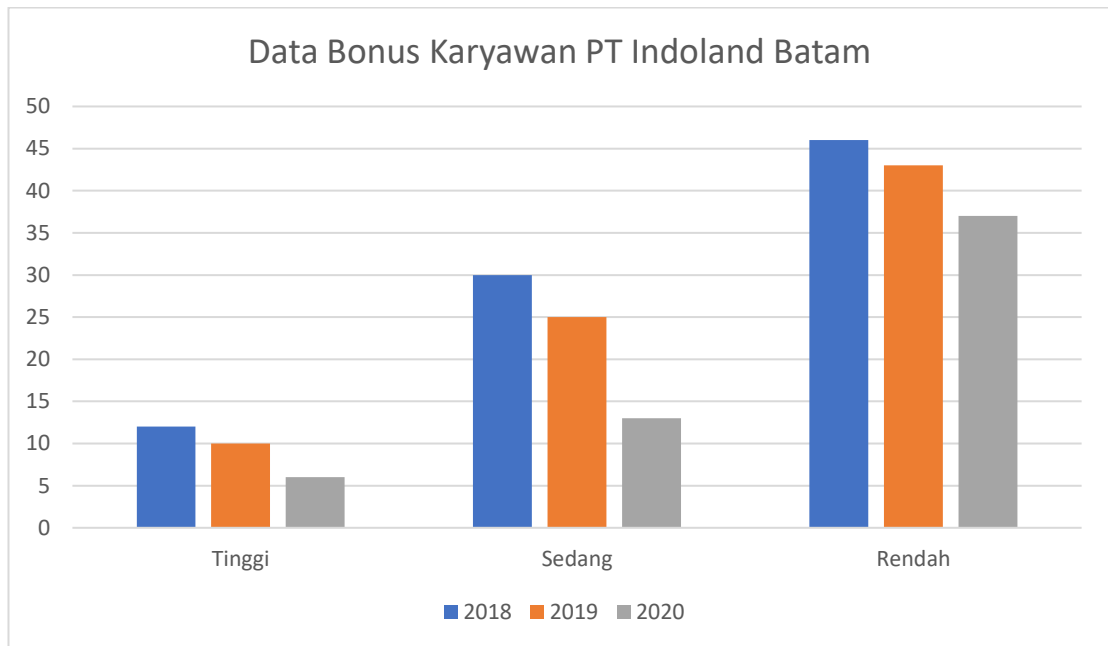
1.1. Latar Belakang

Dalam menghadapi persaingan yang ketat akibat perubahan teknologi dan perubahan lingkungan di berbagai bidang kehidupan, semua perusahaan mencari karyawan yang berpotensi untuk memberikan layanan terbaik dan bermanfaat. Dengan cara ini, perusahaan tidak hanya dapat memberikan layanan yang memuaskan (kepuasan pelanggan), tetapi juga mewujudkan nilai (nilai pelanggan). Perusahaan yang beroperasi tidak hanya mengejar produktivitas tetapi juga kinerja. Menurut (Bangun, 2012) menyatakan kinerja adalah hasil seseorang menyelesaikan pekerjaan sesuai dengan kebutuhan pekerjaan. Kinerja karyawan merupakan faktor penting, karena kemajuan suatu organisasi atau perusahaan bergantung pada sumber daya manusianya. Jika kinerja meningkat maka peluang keberhasilan pencapaian tujuan perusahaan semakin besar, namun jika kinerja menurun dapat menyebabkan perusahaan mengalami kemunduran dan perusahaan tidak dapat mempertahankan usahanya.

Menurut (Bangun, 2012), bagi perusahaan penilaian kinerja memiliki banyak manfaat, antara lain evaluasi antar individu dalam organisasi, pengembangan diri masing-masing individu, pemeliharaan sistem, dan dokumentasi. Output berkaitan erat dengan kuantitas yang dihasilkan dalam waktu singkat, apa saja capaian maksimal, capaian ideal dan capaian minimal serta

faktor-faktor lain yang mempengaruhinya. Kinerja karyawan terkadang meningkat, terkadang menurun, dan penurunan ini pun dapat membuat organisasi atau perusahaan kehilangan reputasinya di mata masyarakat. Ini perlu dipertimbangkan dan dipelajari bersama untuk mencapai tujuan yang diinginkan.

PT Indoland Batam merupakan sebuah perusahaan yang bergerak di bidang penjualan *online* atau *online shop* yang menyediakan barang-barang seperti peralatan kebutuhan rumah tangga, tas, sepatu, alat kecantikan, pakaian, dan lain sebagainya. Perusahaan ini berdiri sejak tahun 2017 yang beralamat di Citra Buana Park 1, Blok H No 2, Jalan Yos Sudarso, Kampung Seraya Batam. Dalam menentukan *reward* 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 *reward* karyawan. Seiring berlewatnya tahun sering kali karyawan pada PT Indoland Batam tidak puas dengan *reward* yang mereka terima sehingga masalah ini menimbulkan kurangnya kinerja karyawan, kedisiplinan karyawan, dan loyalitas karyawan terhadap PT Indoland Batam.



Gambar 1.1 Data Bonus Karyawan PT Indoland Batam

Sumber: (PT Indoland Batam, 2020)

Dalam hal peningkatan kinerja karyawan, disiplin karyawan dan loyalitas karyawan kepada PT Indoland Batam, Pimpinan PT Indoland Batam akan memberikan bonus kepada karyawan sesuai standar yang telah ditetapkan setiap tahun. PT Indoland Batam memiliki karyawan sebanyak 50. Oleh karena itu, karena analisis terhadap sejumlah besar standar dan data karyawan, pimpinan perusahaan harus bekerja keras untuk menganalisis karyawan mana saja yang berhak atas *reward* ini. Oleh karena itu, PT Indoland Batam juga harus memperhatikan faktor pemberian *reward* yang menjadi salah satu motivasi karyawan untuk bekerja. Pemberian *reward* kepada karyawan merupakan kewajiban perusahaan yang harus dibayar tepat waktu, dilakukan secara adil dan berdasarkan hasil kerja. Bonus adalah segala sesuatu yang diterima, baik fisik

maupun non fisik yang harus diperhitungkan dan diberikan kepada karyawan yang berhak atas *reward*. Sistem bonus yang baik akan dapat memuaskan karyawan dan memungkinkan toko / perusahaan mendapatkan banyak penghasilan setiap bulannya (Afriany & Aisyah, 2019). Ketika perusahaan memperoleh keuntungan atau mencapai target keuntungan perusahaan karena peningkatan efisiensi kerja pekerja, *reward* diberikan kepada karyawan swasta di akhir tahun.

Indikator yang selama ini digunakan dalam memberikan *reward* pada PT Indoland Batam sendiri adalah antara lain:

1. Lama bekerja karyawan.
2. Penjualan karyawan saja (jumlah yang terjual).

Hal ini sering menimbulkan kesalahan data serta kecemburuan antar satu dengan yang lain dan pada akhirnya jumlah *reward* sering terjadi ketidak rataan antar satu dengan yang lainnya. Algoritma C4.5 merupakan solusi yang dapat digunakan untuk menyelesaikan masalah yang berkaitan dengan kasus dalam teknik klasifikasi. Keluaran dari algoritma C4.5 berupa pohon keputusan terstruktur yang dapat digunakan untuk mengubah kumpulan data menjadi pohon keputusan yang terdiri dari aturan keputusan. Jadi untuk menentukan indikator pembagian bonus menjadi lebih mudah. Berdasarkan permasalahan yang ada penulis membuat suatu sistem pendukung untuk menentukan *reward* karyawan PT Indoland Batam dengan menggunakan metode Algoritma C4.5.

Teknik datamining dengan algoritma C4.5 pada penelitian (Dhamayanti, 2018), yang meneliti tentang penentuan *reward* bagi karyawan yang berprestasi

dengan algoritma C4.5 menyatakan algoritma C4.5 merupakan suatu sistem algoritma yang digunakan untuk membentuk sebuah pohon keputusan atau sistem pendukung keputusan yang memiliki fungsi untuk menganalisa masing-masing data atau kriteria karyawan, kemudian dibandingkan satu sama dengan yang lainnya dan mendapatkan suatu hasil penilaian untuk menentukan karyawan yang layak mendapatkan *reward* pegawai. Sedangkan menurut (Baiq Andriskha Candra Permana & Intan Komala Dewi Patwari, 2019) Algoritma C4.5 merupakan suatu solusi yang digunakan untuk memecahkan masalah-masalah yang terkait dengan klasifikasi. Algoritma C4.5 berupa sebuah pohon keputusan atau decision tree yang terstruktur yang berfungsi untuk mengumpulkan data-data yang telah diubah dan menjadi suatu pohon keputusan atau decision tree yang terdiri dari suatu aturan keputusan sehingga hasil data tersebut dapat membantu melakukan suatu keputusan.

Pemanfaatan algoritma ini diharapkan dapat mempercepat proses perhitungan atau lebih efisien dalam menentukan *reward*. Hal ini sangatlah penting karena Algoritma dapat menyampaikan informasi lebih tepat dan efisien kepada pemilik perusahaan. Berdasarkan uraian latar belakang diatas, maka peneliti tertarik mengambil sebuah penelitian dengan judul “Penerapan Algoritma C4.5 Untuk Menentukan Reward Karyawan Pada PT Indoland Batam”.

1.2. Identifikasi Masalah

Berdasarkan uraian latar belakang diatas, maka masalah pokok identifikasi masalah yaitu sebagai berikut :

1. Pimpinan atau ketua perusahaan mengalami kesulitan dalam menganalisis data-data kriteria karyawan pada PT Indoland Batam untuk menentukan pemberian *reward* karyawan.
2. Adanya keluhan dari beberapa karyawan dari PT Indoland Batam dalam pemberian *reward* yang tidak merata atau dengan tidak benar.
3. Belum adanya sistem untuk menentukan *reward* karyawan pada PT Indoland Batam.

1.3. Batasan Masalah

Agar pembahasan dalam penelitian ini bisa lebih terarah dan jelas , maka penelitian yang dileliti berfokus pada :

1. Menganalisa data karyawan yang dilakukan pada PT Indoland Batam yang beralamat di Citra Buana Park 1, Blok H No 2, Jalan Yos Sudarso, Kampung Seraya Batam.
2. Mengelola data karyawan yang dilakukan pada PT Indoland Batam selama 3 tahun terakhir yaitu 2018, 2019, dan 2020.
3. Penelitian ini menggunakan algoritma C4.5 dan Software WEKA versi 3.8.4 untuk menentukan pohon keputusan yang dibutuhkan dalam pengambilan keputusan yang kemudian diuji dengan *Microsoft Excel* untuk proses pencarian data secara manual.

1.4. Perumusan Masalah

Berdasarkan uraian identifikasi masalah diatas, maka permasalahan yang akan dibahas yaitu sebagai berikut :

1. Bagaimanakah penerapan Data mining khususnya Teknik klasifikasi menggunakan Algoritma C4.5 dalam menentukan pemberian *reward* pada karyawan PT Indoland Batam?
2. Apakah Teknik klasifikasi menggunakan algoritma C4.5 dapat menentukan pola dalam menentukan pemberian *reward* pada karyawan PT Indoland Batam?
3. Bagaimanakah cara mengimplementasikan Teknik klasifikasi Algoritma C4.5 sebagai alat untuk menemukan pola yang digunakan dalam menentukan *reward* pada karyawan PT Indoland Batam?
4. Bagaimanakah Tindakan kedepannya dengan adanya hasil yang diperoleh dari proses penelitian yang dilakukan mengenai pemberian *reward* pada karyawan PT Indoland Batam?

1.5. Tujuan Penelitian

Berdasarkan uraian masalah diatas, tujuan pelaksanaan penelitian tersebut yaitu sebagai berikut :

1. Menerapkan Data Mining khususnya teknik klasifikasi menggunakan Algoritma C4.5 dalam menentukan pemberian *reward* karyawan pada PT Indoland Batam.
2. Pengujian teknik klasifikasi dan Algoritma C4.5 dalam wujud penerapan dari Data Mining yang digunakan dalam menentukan pemberian *reward* pada PT Indoland Batam.

3. Pengujian Algoritma C4.5 menggunakan software Data Mining WEKA dalam menentukan pemberian *reward* pada PT Indoland Batam.
4. Menguji dan menganalisa penerapan Data Mining teknik klasifikasi menggunakan Algoritma C4.5 dalam menentukan pemberian *reward* pada PT Indoland Batam.

1.6. Manfaat Penelitian

Dalam melakukan penelitian ini Adapun manfaat penelitian ini bagi pembaca atau peneliti lain yaitu sebagai berikut :

1.6.1. Manfaat Teoritis

Berikut adalah manfaat teoritis bagi pembaca dalam penelitian ini adalah sebagai berikut :

1. Dapat dijadikan sebuah tambahan wawasan atau pengetahuan baik bagi pembaca maupun bagi peneliti.
2. Dapat dijadikan referensi bagi para peneliti yang ingin penelitiannya menggunakan metode Algoritma C4.5 dalam menentukan *reward* karyawan.

1.6.2. Manfaat Praktis

Selain manfaat teoritis, manfaat praktis bagi pembaca dalam penelitian ini adalah sebagai berikut :

1. Dengan menggunakan penerapan Algoritma C.45 ini diharapkan agar dapat membantu menjadi praktis dan mudah untuk menentukan pemberian *reward* karyawan pada PT. Indoland Batam.
2. Sebagai penambahan wawasan atau pengetahuan pada perusahaan PT Indoland Batam agar kedepannya tidak terjadi kesalahan lagi.



Universitas Putera Batam

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. Teori Umum

Penulis membahas beberapa teori umum yang berkaitan dengan pelaksanaan penelitian yang terdiri dari pengertian dan istilah secara umum.

2.1.1.1. Sistem Pendukung Keputusan (SPK)

Menurut (Sri Lestari & Saefudin, 2015), Sistem pendukung keputusan merupakan bagian dari sistem informasi berbasis komputer, termasuk sistem berbasis pengetahuan atau manajemen pengetahuan, yang digunakan untuk mendukung pengambilan keputusan dalam suatu organisasi atau perusahaan. Pada dasarnya, SPK merupakan pengembangan lebih lanjut dari sistem informasi manajemen terkomputerisasi, yang dirancang sedemikian rupa sehingga dapat berinteraksi dengan pemakainya. Sistem pendukung keputusan terdiri dari tiga komponen utama yaitu subsistem pengolah data (database), subsistem pengolah model (modelbase) dan subsistem pengolah dialog.

2.1.1.2.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.

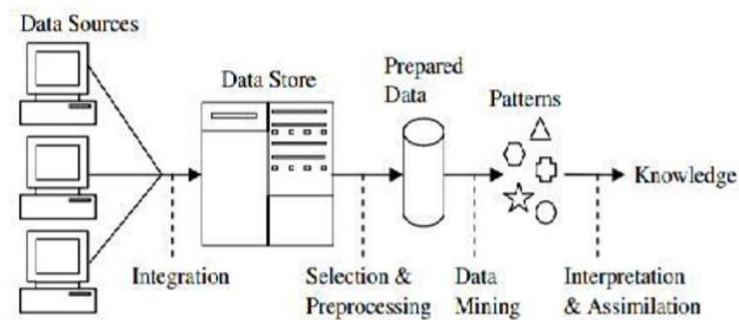
4. *Data Mining*

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:

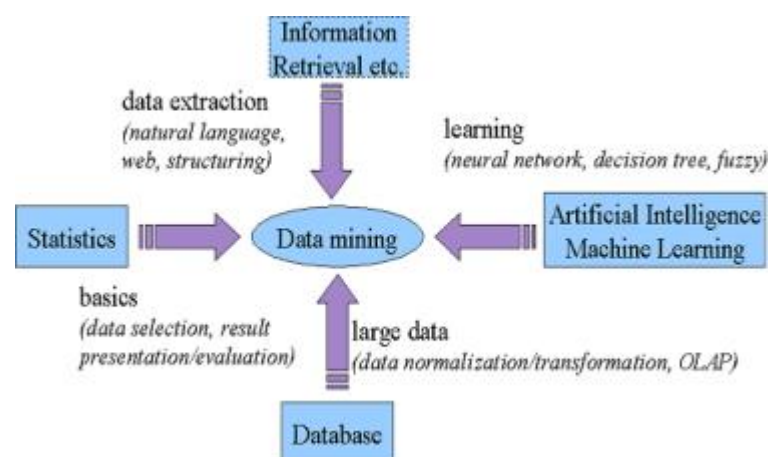


Gambar 2.1 Proses Dalam KDD atau *Knowledge Discovery in Database*

Sumber: (Mardi, 2016)

2.1.1.3. Data Mining

Data mining atau machine learning dan data mining bisa dimanfaatkan dan sering digunakan dalam melakukan berbagai bentuk kegiatan analisis untuk menjelaskan ilmu dan pengetahuan secara jelas dan mudah dipahami oleh pikiran yang dimiliki manusia (Hermawan, Sukma, & Halfis, 2019).



Gambar 2.2 Bidang Ilmu Data Mining

Sumber: (Mardi, 2016)

Dalam prosesnya, data mining memiliki teknik dalam pengolahan data tersendiri hingga terbentuk sebuah pola, kemudian pola tersebut difungsikan untuk dapat mengenali pola lain yang tidak tersimpan dalam lokasi penyimpanan data dan informasi yang sama. Dalam kegiatan atau proses melakukan prediksi, banyak peneliti melakukan kegiatan tersebut dengan memanfaatkan fungsi dari Teknik data mining. Awal mula istilah data mining yaitu karena data dan informasi yang akan diproses sudah lama dan cara pengolahannya dilakukan seperti kegiatan menambang hingga mendapatkan hasil yang berharga dan bermanfaat (Hermawan et al., 2019).

2.1.1.4. Pengelompokan *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 pengelompokan mengenai penghasilan yang dimiliki oleh pekerja dalam sebuah perusahaan, pengelompokan 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 pengelompokan 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:

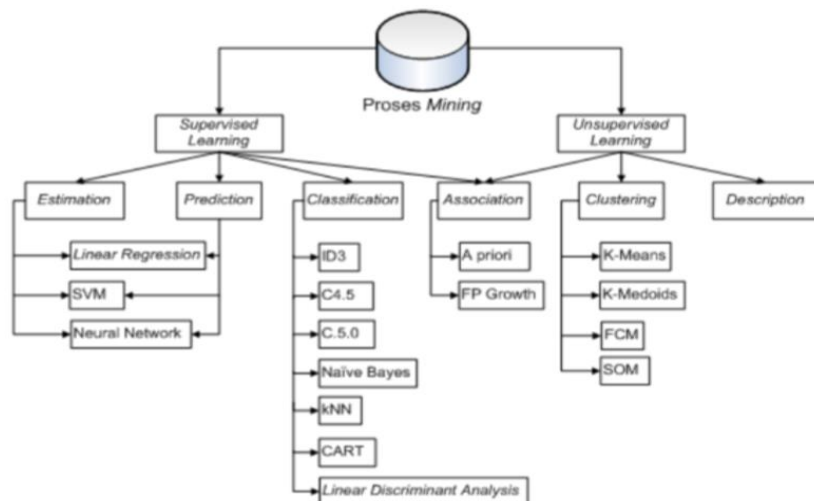
- a. Pada saat melakukan pengelompokan terhadap jenis-jenis dari konsumen yang dijadikan target dari kegiatan pemasaran dari sebuah produk yang dilakukan suatu perusahaan.
- b. Pada saat melakukan kegiatan audit, dilakukan pemisahan berdasarkan transaksi yang diduga mencurigakan dan transaksi yang diduga 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 pengelompokan beberapa proses atau tahapan berdasarkan tugas saat seseorang peneliti melakukan penelitian menggunakan teknik dengan *data mining*, sebagai berikut.



Gambar 2.3 Pengelompokan Teknik *Data Mining*

Sumber: (Hermawan, Sukma, & Halfis, 2019)

2.1.1.5. *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:

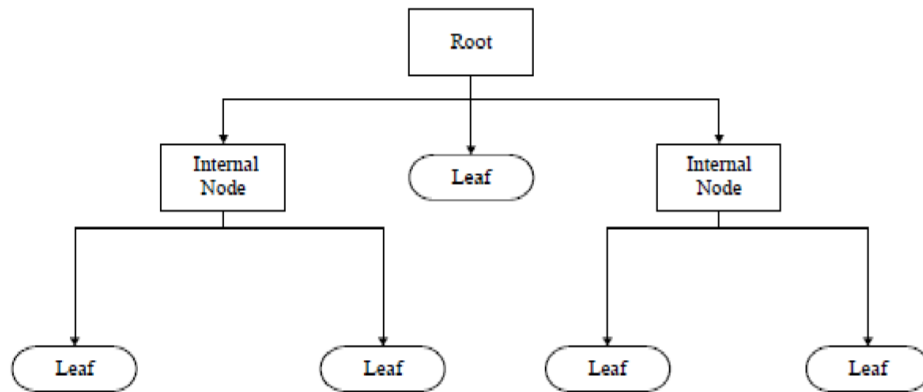
1. Kelebihan dari *decision tree*

- 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 penggunaan 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 2.4 Model *Decision Tree*

Sumber: (Takalapeta, 2018)

2.1.1.6. 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).

2.1.2. Teori Khusus

Penulis membahas beberapa teori khusus yang berkaitan dengan pelaksanaan penelitian yang terdiri dari pengertian dan istilah secara khusus.

2.1.2.1. Kinerja Karyawan

Kinerja karyawan adalah kuantitas dan / atau kualitas pekerjaan yang dihasilkan oleh karyawan PT Indoland Batam, seperti standar kerja, dalam kurun waktu tertentu, berdasarkan peraturan yang ditetapkan atau berlaku, prosedur operasi, standar, dan fungsi di 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

2.1.2.2.Reward

Reward merupakan salah satu cara untuk memotivasi seseorang agar berbuat baik dan meningkatkan prestasinya. *Reward* adalah imbalan, ganjaran, hadiah, atau penghargaan yang dirancang untuk membuat seseorang lebih aktif bekerja untuk meningkatkan atau meningkatkan kinerja yang dicapai. *Reward* dapat diartikan sebagai bentuk penghargaan atas prestasi tertentu yang diberikan oleh individu atau organisasi, biasanya dalam bentuk materi atau verbal (Fitri, Ludigdo & Djamhuri, 2013). *Reward* yang diberikan kepada karyawan akan berdampak signifikan terhadap kinerja.

2.1.2.3.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.

2.1.2.4.Karyawan

Karyawan merupakan faktor pendukung perusahaan atau instansi, karena dengan karyawan dengan standar kualifikasi perusahaan maka produktivitas perusahaan pasti akan terjaga dan ditingkatkan. Proses pemilihan karyawan yang unggul adalah proses kompleks yang membutuhkan pertimbangan yang cermat. Untuk memperoleh informasi yang cepat dan akurat tentang kinerja karyawan yang benar (sesuai dengan standar yang diharapkan) diperlukan proses otomatis dengan menggunakan teknologi. Oleh karena itu untuk memenuhi kebutuhan akan informasi diperlukan suatu sistem yang berbasis komputer (Afriany & Aisyah, 2019).

2.1.2.5.Insentif

Insentif merupakan reward yang diberikan kepada pegawai oleh pimpinan organisasi berupa uang agar pegawai memiliki motivasi dan prestasi yang tinggi dalam mencapai tujuan organisasi. Dengan kata lain insentif kerja adalah uang disamping gaji yang dibayarkan oleh pimpinan organisasi menurut (Sahir &

Panjaitan, 2019) untuk mengevaluasi prestasi kerja dan kontribusi karyawan terhadap organisasi.

Insentif adalah daya tarik yang disengaja yang diberikan oleh perusahaan, manajer dan karyawan. Tujuannya untuk ikut serta dalam membangun, memelihara dan memperkuat ekspektasi manajer dan karyawan agar memiliki semangat yang lebih besar untuk berpartisipasi di perusahaan dalam hal meningkatkan efisiensi kerja.

Oleh karena itu, tindakan insentif akan mendorong peningkatan prestasi kerja karyawan yang sangat membantu pencapaian tujuan perusahaan, pada akhirnya tindakan insentif akan membawa manfaat bagi kedua belah pihak, yaitu meningkatkan prestasi kerja bagi karyawan berarti meningkatkan pendapatan.

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

Waikato Environment for Knowledge Analysis atau WEKA adalah seperangkat alat pembelajaran mesin praktis. *Waikato Environment for Knowledge Analysis (WEKA)* dikembangkan oleh Universitas Waikato di Selandia Baru untuk penelitian, pendidikan, dan berbagai aplikasi (Pujiono, Amborowati, & Suyanto, 2013). WEKA adalah software open source, jadi siapapun bisa menambahkan software ini (Haryati, Sudarsono, & Suryana, 2015).

WEKA dapat digunakan untuk memecahkan masalah data mining di dunia nyata, khususnya klasifikasi dengan menggunakan metode pembelajaran mesin. WEKA mencakup pemrosesan awal data, klasifikasi, regresi, pengelompokan, aturan asosiasi, dan alat visualisasi. Pengujian dengan WEKA lebih fleksibel dan

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



Gambar 2.5 Tampilan Utama WEKA

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 *Decission 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 “Prediksi Waktu Kuliah Dengan Algoritma Classifier” yang dilakukan oleh (Rika Harman & Evan Rosiska, 2020) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti

dapat membantu mahasiswa mencari akar permasalahan dan menentukan penyebab utama mahasiswa tidak dapat menyelesaikan masa studinya tepat waktu.

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 penyelarasan 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 in 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.

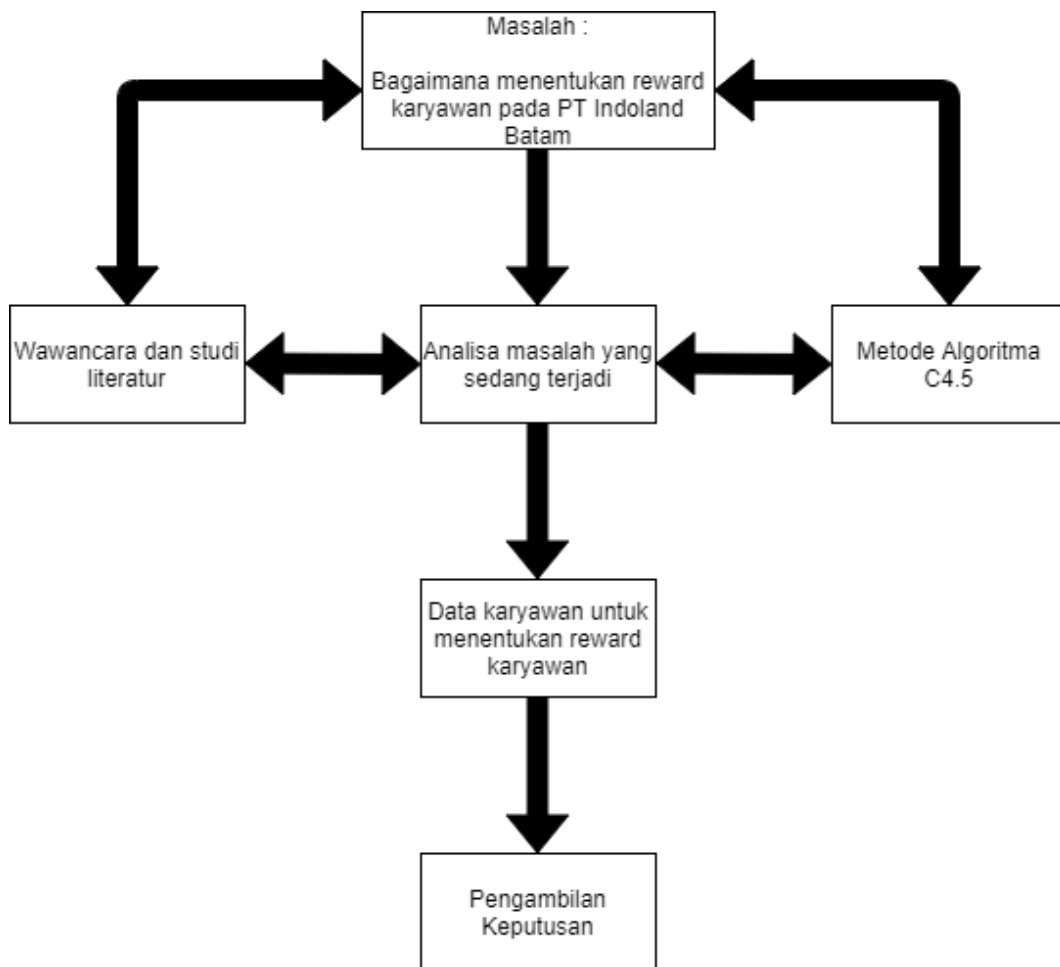
9. Dari penelitian “Analisis Dengan Metode Klasifikasi Menggunakan *Decision Tree* Untuk Prediksi Penentuan Resiko Kredit Pada Bank Bukopin Batam” yang dilakukan oleh (Intan Utma Sari, 2015) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 diharapkan dapat membantu pihak bank agar lebih 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.
10. Penelitian yang berjudul “implementation of decision trees algorithm C4.5” yang dilakukan oleh (Chauhan & Chauhan, 2013) menunjukkan bahwa algoritma C4.5 menunjukkan hasil yang lebih akurat dan pengimplementasian pohon keputusan dengan metode algoritma C4.5 sangat mudah dan cepat. Algoritma ini lebih cocok untuk data sets yang mengandung tidak terlalu banyak data.
11. Penelitian yang berjudul “Performance evaluation of the GIS-based datamining techniques of best-first decision tree, random forest, and naïve bayes tree for landslide suscepttibility modeling” yang dilakukan

oleh (Chen, Zhang, Li, & Shahabi, 2018) menghasilkan bahwa random forest model merupakan yang paling akurat prediksinya.

12. Dari penelitian "Reservoir Inflow Forecasting Using ID3 and C4.5 Decision Tree Model" yang dilakukan oleh (Charoenporn, 2017) didapatkan informasi²⁹ bahwa jenis algoritma dari ID3 dan juga algoritma dari C4.5 merupakan salah satu cara dalam membuat gudang dalam melakukan pengujian dengan suatu model atau tahapan kerja. Maksudnya adalah jenis algoritma dari ID3 dan juga jenis algoritma dari C4.5 merupakan bentuk perhitungan untuk decision tree. Pada akhirnya dapat menghasilkan solusi atau keputusan akhir yang terbaik berdasarkan berbagai proses penyesuaian data.

2.3. Kerangka Pemikiran

Berikut adalah kerangka pemikiran dalam penelitian ini, yang antara lain dapat digambarkan sebagai berikut:



Gambar 2.6 Kerangka Berpikir

Sumber: (Penulis, 2020)



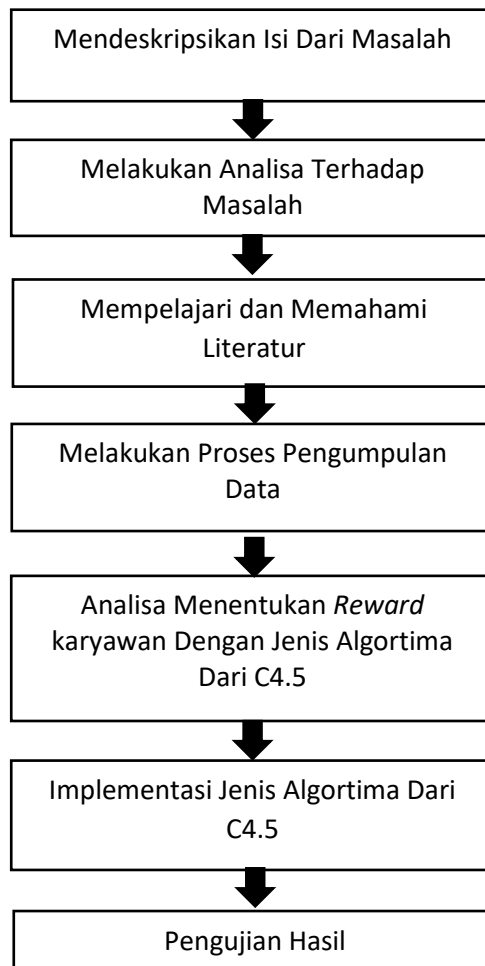
Universitas Putera Batam

BAB III

METODE PENELITIAN

3.1. Desain Penelitian

Berikut ini adalah uraian tentang desain penelitian yang akan dilakukan, desain tersebut secara garis besar dapat dilihat antara lain sebagai berikut:



Gambar 3.1 Desain Penelitian

Sumber: (Penulis, 2020)

Keterangan:

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

1. Mendeskripsikan isi dari masalah

Penulis berusaha keras menjelaskan masalah yang akan diteliti dengan menggunakan algoritma C4.5 terkait menentukan reward karyawan.

2. Melakukan analisa terhadap masalah

Tahapan ini merupakan kegiatan tindak lanjut setelah mendeskripsikan masalah. Penulis mencoba menganalisis menurut beberapa pertanyaan yang ingin penulis angkat dan teliti.

3. Mempelajari dan memahami literatur

Penulis berupaya keras untuk meneliti dan memahami beberapa sumber teori yang mendukung penelitian jurnal agar dapat dijadikan sebagai pedoman bagi penulis dalam melakukan penelitian.

4. Melakukan proses pengumpulan data

Pada tahap ini peneliti melakukan wawancara dengan mewawancarai subjek penelitian dan menanyai pemilik dan pekerja PT Indoland Batam untuk mendapatkan lebih banyak data dan informasi yang dibutuhkan.

5. Analisa menentukan *reward* karyawan dengan jenis algoritma dari C4.5.

Penggunaan jenis algoritma dari C4.5 mulai dari proses pembuatan struktur pohon keputusan, hingga menghasilkan pola penentuan *reward* karyawan dalam bentuk aturan.

6. Implementasi jenis algoritma dari C4.5

Penulis menentukan integritas data yang digunakan sebagai root dari pohon keputusan, kemudian menghitung nilai gain tertinggi dari semua komponen

yang terdapat dalam data tersebut, kemudian menggunakan nilai gain maksimum tersebut dalam proses pembentukan struktur pohon keputusan.

7. Pengujian Hasil

Pada tahap akhir, peneliti melakukan pengujian setelah menyelesaikan perhitungan nilai gain tertinggi dan membentuk pohon keputusan dengan 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 PT Indoland Batam yang beralamat di Citra Buana Park 1, Blok H No 2, Jalan Yos Sudarso, Kampung Seraya Batam.

3.2.2. Jadwal Penelitian

Demikian penjelasan penulis mengenai rencana waktu yang akan dilakukan dalam melaksanakan kegiatan analisa mengenai menentukan *reward* karyawan pada PT Indoland Batam, sebagai berikut.

No	Kegiatan	Sep-20	Okt-20	Nov-20	Des-20	Jan-21
1	Bimbingan					
2	Identifikasi Masalah					
3	Studi Literatur					
4	Penentuan Judul					
5	Pengumpulan Data					
6	Pengolahan Data					
7	Pengujian Hasil					
8	Penyusunan Laporan					

Tabel 3.1 Jadwal Penelitian

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 PT Indoland Batam dalam periode 3 tahun terakhir yaitu tahun 2018, 2019, dan 2020.

3.3.2. Sampel

Data yang akan dijadikan sampel pada penelitian ini, peneliti akan mengambil data kinerja karyawan sebanyak 50 data sampel yang mana semua data tersebut diambil dari seluruh data yang ada pada populasi dalam periode 3 tahun terakhir yaitu tahun 2018, 2019, dan 2020.

3.4. Variabel Penelitian

Untuk menganalisis data dalam sebuah penelitian memerlukan variable-variabel, data yang akan diolah adalah data kinerja karyawan periode 3 tahun terakhir yaitu 2018, 2019, dan 2020 dari PT Indoland Batam. Beberapa indikator yang digunakan dalam penelitian khususnya yang berkaitan dengan variabel untuk menentukan *reward* karyawan, sebagai berikut:

1. Masa Kerja
2. Loyalitas
3. Kehadiran

4. Kerajinan
5. Sikap
6. Inisiatif
7. 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 sastra dalam bidang penelitian tertentu. Kajian ini adalah deskripsi singkat dari topik yang telah dipelajari, diperdebatkan, dan ditentukan, biasanya disusun berdasarkan waktu atau topik. Dalam hal ini peneliti mempelajari berbagai jurnal yang berkaitan dengan masalah yang dihadapi.
2. Teknik observasi adalah teknik pengumpulan data berasal dari observasi langsung oleh peneliti di bidang ini, dalam hal ini peneliti terjun langsung ke PT Indoland Batam untuk melihat proses kinerja karyawan PT Indoland Batam.
3. Teknik dokumentasi adalah peneliti mempelajari informasi tertulis yang ada di lokasi penelitian, yang berkaitan dengan masalah data sebelumnya, atau melakukan proses penggalian data yang ada, kemudian mengelompokkan data tersebut sesuai dengan kebutuhan penelitian yang akan dilakukan. Dalam hal ini peneliti akan mengamati dan mengkaji data kinerja karyawan PT Indoland Batam.

4. Teknik wawancara adalah proses tanya jawab bagi peneliti yang melakukan penelitian lisan, di mana dua orang atau lebih secara langsung mendengarkan informasi atau pernyataan secara langsung. Dalam hal ini peneliti menanyakan kepada pemilik dan pekerja PT Indoland Batam.

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 karyawan, 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 *reward* karyawan pada PT Indoland Batam.