

**PREDIKSI KINERJA KARYAWAN BERDASARKAN
PROSES TRAINER MENGGUNAKAN DATA MINING**

SKRIPSI



Oleh:

Yunita Tri Utami

171510082

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

2022

**PREDIKSI KINERJA KARYAWAN BERDASARKAN
PROSES TRAINER MENGGUNAKAN DATA MINING**

SKRIPSI

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



Oleh:

Yunita Tri Utami

171510082

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2022**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Yunita Tri Utami
NPM : 171510082
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

PREDIKSI KINERJA KARYAWAN BERDASARKAN PROSES TRAINER MENGUNAKAN DATA MINING

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, 28 Januari 2022



Yunita Tri Utami
171510082

PREDIKSI KINERJA KARYAWAN BERDASARKAN PROSES TRAINER MENGGUNAKAN DATA MINING

SKRIPSI

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

Oleh

Yunita Tri Utami

171510082

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

Batam, 28 Januari 2022



**Erlin Elisa, S. Kom., M. Kom
Pembimbing**

ABSTRAK

Banyak kasus dijumpai bahwa sulitnya pihak HRD memprediksi kinerja karyawan dikarenakan bagian HRD hanya dapat melihat rekapan data dari segi absensi ataupun keterlambatan karyawan. Dengan menggunakan algoritma hingga mendapatkan karyawan mana yang berhak mendapatkan reward, sehingga tidak menyulitkan pihak perusahaan. Algoritma C4.5 digunakan untuk karyawan mana yang berhak mendapatkan reward dari perusahaan. Variabel nya ialah pengecekan kinerja kerja, kedisiplinan karyawan, target yang telah di capai, kerjasama antar team dan kemampuan dalam berkerja. Hasil uji evaluasi memperlihatkan Algoritma Decision Tree C4.5 cocok dengan menggunakan aplikasi WEKA diterapkan untuk penentuan karyawan mana yang berhak mendapatkan reward dari perusahaan.

Kata Kunci : Kinerja Karyawan, Damining, Algoritma C.45

ABSTRACT

In many cases, it is found that it is difficult for HRD to predict employee performance because the HRD section can only see data recap in terms of employee attendance or delays. By using an algorithm to get which employees are entitled to get rewards, so it doesn't make it difficult for the company. The C4.5 algorithm is used to determine which employees are entitled to a reward from the company. The variables are checking work performance, employee discipline, targets that have been achieved, cooperation between teams and ability to work. The results of the evaluation test show that the Decision Tree C4.5 Algorithm is suitable for using the WEKA application to determine which employees are entitled to receive rewards from the company.

Keywords: Employee Performance, Damining, Algorithm C.45

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala berkat dan anugerahNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Dengan segala keterbatasan, penulis juga menyadari 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. Rektor Universitas Putera Batam Ibu Nur Elfi Husda, S.kom., M.SI.,
2. Bapak Welly Sugianto, S.T., M.M., selaku Dekan Fakultas Teknik Dan Komputer di Universitas Putera Batam, dan Pembimbing Akademik pada Program Studi Sistem Informasi Universitas Putera Batam,
3. Ketua Program Studi Sistem Informasi Bapak Muhammat Rasid Ridho, S.Kom., M.SI.,
4. Ibu Erlin Elisa, S.Kom., M.Kom. selaku Pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam
5. Saut Pintubipar Saragih, S.Kom., M.MSI. Pembimbing Akademik penulis periode 2017-2022 yang telah memberikan motivasi selama masa kuliah,
6. Dosen dan staff Universitas Putera Batam, yang memberikan pengetahuan selama perkuliahan berlangsung,

7. Kedua orang tua penulis, Bapak dan mamak , yang selalu berdoa dan menyemangati penulis,
8. Suami dan juga anak yang telah menjadi support bagi penulis hingga dapat menyelesaikan perkuliahan dan skripsi ini
9. Dan Terimakasih untuk Semua pihak yang telah ikut membantu dan juga memeberikan ilmu yang tak bisa penulis sebutkan satu persatu.

Semoga Tuhan membalas kebaikan dan selalu mencurahkan berkat yang melimpah. Penulis berharap skripsi ini dapat menjadi ilmu yang berharga bagi para pembaca di masa yang akan datang. Penulis juga berharap kritik dan saran yang bisa membangun dari para pembaca.

Batam, 28 Januari 2022

Yunita Tri Utami

DAFTAR ISI

| | |
|----------------------------------------------------|-------------|
| HALAMAN SAMPUL | i |
| HALAMAN JUDUL | ii |
| SURAT PERNYATAAN ORISINALITAS | iii |
| HALAMAN PENGESAHAN | iv |
| ABSTRAK | v |
| ABSTRACT | vi |
| KATA PENGANTAR | vii |
| DAFTAR ISI | ix |
| DAFTAR GAMBAR | xii |
| DAFTAR RUMUS | xiii |
| DAFTAR TABEL | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Identifikasi Masalah..... | 3 |
| 1.3 Batasan Masalah | 4 |
| 1.4 Rumusan Masalah..... | 4 |
| 1.5 Tujuan Penelitian | 4 |
| 1.6 Manfaat Penelitian | 5 |
| 1.6.1 Manfaat Teoritis | 5 |
| 1.6.2 Manfaat Praktis..... | 5 |
| BAB II | 6 |
| TINJAUAN PUSTAKA | 6 |
| 2.1. Teori Umum | 6 |
| 2.1.1 <i>Knowledge Discovery In Database</i> | 6 |

| | | |
|----------------------------------------|----------------------------------------------------------------|-----------|
| 2.1.2 | Data Mining | 8 |
| 2.1.3 | Kelompok <i>Data Mining</i> | 8 |
| 2.1.4 | Algoritma C4.5 | 10 |
| 2.1.5 | <i>WEKA (Waikato Environment for Knowledge Analysis)</i> | 11 |
| 2.2. | Teori Khusus | 12 |
| 2.2.1 | Karyawan | 12 |
| 2.2.2 | Kinerja Karyawan..... | 12 |
| 2.2.3. | <i>Trainer</i> | 14 |
| 2.3. | Penelitian Terdahulu | 15 |
| 2.4. | Kerangka Pemikiran | 18 |
| BAB III METODE PENELITIAN | | 20 |
| 3.1 | <i>Design</i> Penelitian | 20 |
| 3.2 | Objek Peneltian | 22 |
| 3.2.1. | Lokasi Penelitian | 22 |
| 3.3 | Populasi dan Sampel | 22 |
| 3.3.1 | Populasi..... | 23 |
| 3.3.2 | Sampel..... | 23 |
| 3.4 | Variabel Penelitian..... | 24 |
| 3.5 | Teknik Pengumpulan Data..... | 25 |
| 3.6 | Model Penelitian..... | 25 |
| BAB IV | | 27 |
| HASIL DAN PEMBAHASAN..... | | 27 |
| 4.1 | Pembahasan | 27 |
| 4.2 | Hasil Penelitian..... | 33 |
| 4.3 | Analisis Data | 34 |
| 4.3.1 | Pembersihan Data..... | 34 |

| | | |
|----------------------------|--------------------------------------------------------|---------------|
| 4.3.2 | Seleksi Data..... | 34 |
| 4.3.3 | Transformasi Data | 34 |
| 4.3.4 | Pohon Keputusan..... | 40 |
| 4.3.5 | Hasil perhitungan <i>Entropy</i> dan <i>Gain</i> | 55 |
| 4.3.6 | Pengujian menggunakan <i>Software WEKA</i> | 56 |
| BAB V..... | | 67 |
| PENUTUP..... | | 67 |
| 5.1 | Kesimpulan..... | 67 |
| 5.2 | Saran | 69 |
| DAFTAR PUSTAKA..... | | 70 |
| LAMPIRAN 1..... | | lxxiii |
| LAMPIRAN 2..... | | lxxvii |
| LAMPIRAN 3..... | | lxxix |
| LAMPIRAN 4..... | | lxxx |

DAFTAR GAMBAR

| | |
|---------------------------------------------------|----|
| Gambar 2. 1 Langkah langkah KDD | 8 |
| Gambar 2. 2 Pengelompokan Data Mining | 10 |
| Gambar 2. 3 Kerangka Berfikir | 19 |
| Gambar 3. 1 Design Penelitian | 20 |
| Gambar 3. 2 Lokasi Objek Penelitian | 22 |
| Gambar 3. 3 Model Penelitian..... | 26 |
| Gambar 4. 1 Hasil Perhitungan Node 1 | 46 |
| Gambar 4. 2 Pohon Keputusan Node 2..... | 51 |
| Gambar 4. 3 Node 3..... | 55 |
| Gambar 4. 4 Format Data Microsoft Excel | 57 |
| Gambar 4. 5 Format Data Notepad..... | 58 |
| Gambar 4. 6 Software WEKA..... | 59 |
| Gambar 4. 7 Data Pada Menu Explorer | 60 |
| Gambar 4. 8 Memilih Tabel yg akan di Input | 61 |
| Gambar 4. 9 Data pada Menu Classify | 62 |
| Gambar 4. 10 Visualize Tree pada WEKA..... | 64 |
| Gambar 4. 12 Hasil Visualize Tree pada WEKA | 65 |

DAFTAR RUMUS

| | |
|-------------------------------------------|----|
| Rumus 2. 1 Rumus Entropy | 10 |
| Rumus 2. 2 Perhitungan Gain..... | 11 |
| Rumus 3. 1 Perhitungan rumus slovin | 23 |

DAFTAR TABEL

| | |
|------------------------------------------------------|----|
| Tabel 4. 1 Format Data Karyawan..... | 27 |
| Tabel 4. 2 Klasifikasi Atribut Kinerja..... | 35 |
| Tabel 4. 3 Klasifikasi Atribut Kedisiplinan | 35 |
| Tabel 4. 4 Klasifikasi Atribut Target | 35 |
| Tabel 4. 5 Klasifikasi Atribut Kerjasama..... | 36 |
| Tabel 4. 6 Klasifikasi Atribut Kemampuan Kerja | 36 |
| Tabel 4. 7 Data Pra-Proses Karyawan | 37 |
| Tabel 4. 8 Hasil Perhitungan Node 1 | 45 |
| Tabel 4. 9 Perhitungan hasil Node 2..... | 50 |
| Tabel 4. 10 Perhitungan hasil Node 3..... | 54 |

BAB I

PENDAHULUAN

1.1 Latar Belakang

Wabah virus corona pada tahun 2020 adalah tahun terhambatnya perekonomian global di dunia termasuk di Indonesia, hal ini mengakibatkan banyak perusahaan yang mengurangi produksi secara signifikan karena sulitnya mendapatkan bahan baku. Oleh karena itu, perusahaan – perusahaan tersebut melakukan pemutusan hubungan kerja ataupun mengurangi tenaga kerja karyawan untuk mengurangi beban dalam hal penggajian karyawan dikarenakan pemasukan yang minim yang dihasilkan oleh perusahaan tersebut. Hasil survey Badan Pusat Statistik 2019 tercatat ada sebanyak 8,31% pengangguran di Indonesia, lalu masuk ke awal tahun 2020 pengangguran tersebut peningkatannya menjulang tinggi hingga berada di angka 11,79%. Dinas Tenaga Kerja (Disnaker) mengungkapkan, suasana pencari kerja di Batam saat ini atau sepanjang tahun 2021 dari periode januari hingga maret ada 7.946 pencaker di kota Batam, yang terdiri dari 3.783 orang pencaker laki - laki dan 4.163 orang pencaker perempuan.

Dari data diatas dapat disimpulkan bahwa masa pandemi saat ini mengakibatkan semakin meningkatkan angka pengangguran di Indonesia Tak terkecuali di Kota Batam. Seperti yang kita ketahui, bahwa Kota Batam adalah kota industri yang mempunyai banyak penduduk dikarenakan banyaknya para pencari kerja yang mengadu nasib ke Batam, ditambah dengan masa pandemi seperti sekarang ini membuat bertambahnya jumlah pengangguran di Kota Batam ditandai dengan banyaknya pencaker, para pengusaha industri dituntut untuk

mampu mendapatkan karyawan yang kompeten atau mampu menyeleksi karyawan yang kinerjanya memuaskan. Kinerja karyawan tersebut dapat di nilai dari kehadiran, sopan santun, produktivitas dalam bekerja, dan sebagainya. Salah satu industri yang menerapkan sistem seleksi kinerja karyawan tersebut adalah PT.Schneider Manufacturing Batam.

PT.Schneider Manufacturing Batam berlokasi di kota Batam, kepulauan Riau. PT yang bergerak di pendistribusian, otomasi industri dan kontrol. Perusahaan ini pertama kali didirikan pada tahun 1836 di Perancis. Kemudian membuka cabang di Indonesia yang bertepatan di Jakarta dan Batam. PT.Schneider Manufacturing Batam salah satu perusahaan yang baru baru ini dikabarkan menjadi percontohan industri 4.0 di dunia. Sistem pembagian jam kerja di PT.Schneider Manufacturing Batam menerapkan beberapa sistem *shift* kerja yaitu pagi jam 06.00 - 15.00 wib, siang 15.00 - 23.30, dan Malam 23.30 - 06.00.

Permasalahan yang ada pada PT.Schneider Manufacturing Batam sekarang ini adalah sulitnya pihak HRD memprediksi kinerja karyawan dikarenakan bagian HRD hanya dapat melihat rekapan data dari segi absensi ataupun keterlambatan karyawan. Untuk faktor – faktor kinerja karyawan lainnya seperti kualitas yang dihasilkan, *team work*, profesionalisme dan semacamnya hanya dapat dilihat oleh bagian produksi seperti *team leader* atau *team trainer*. Dikarenakan bagian *team leader* sudah mengemban banyak pekerjaan maka prediksi kinerja karyawan diserahkan ke bagian *team trainer*. *Trainer* di PT.Schneider Manufacturing Batam merupakan bagian atau *team* yang bertugas menjelaskan proses kerja secara *detail* atau mengenal suatu produk yang akan di kerjakan oleh karyawan. Bagian *trainer*

juga bertugas dalam memantau kegiatan pegawai. Untuk penyelesaian masalah yang sudah di kemukakan maka cara yang dapat dilakukan ialah menerapkan *datamining* dengan memprediksi kinerja karyawan berdasarkan proses *trainer* dengan menggunakan algoritma c45. Algoritma ini memakai cara *datamining* yang Jalankan proses penggalian informasi tentang keputusan sebelumnya. Informasi ini digunakan sebagai informasi untuk membentuk pola pohon keputusan.. Pada penelitian terdahulu yang dikatakan oleh (Sunge 2018), Algoritma C4.5 dapat bermanfaat dalam melakukan prediksi kompetensi atau kinerja karyawan, dan juga bisa untuk sebagai acuan dalam pengambilan keputusan bagi pihak perusahaan.

Dari uraian yang diatas, maka pengkaji memilih judul dalam bentuk skripsi yaitu **"PREDIKSI KINERJA KARYAWAN BERDASARKAN PROSES TRAINER MENGGUNAKAN DATA MINING"**.

1.2 Identifikasi Masalah

Dari latar belakang yang telah dipaparkan, maka identifikasi masalah yg dibuat yaitu :

1. PT.Schneider Manufacturing Batam masih kesulitan menemukan karyawan yang kompeten dalam bekerja sehingga menyulitkan untuk menerima *reward*.
2. PT.Schneider Manufacturing Batam belum menemukan metode untuk memprediksi kinerja pekerjanya.

1.3 Batasan Masalah

Agar penelitian ini mengarah dengan satu tujuan, maka pengkaji membuat batasan masalahnya yaitu :

1. Pengkajian dilakukan pada PT Schneider Manufaturing Batam yang berlokasi di Batamindo Industrial Park 01, Muka Kuning, Batam.
2. Metode yang dipakai dalam penelitian ini ialah algoritma C4.5 dan pengujian hasilnya menggunakan weka .

1.4 Rumusan Masalah

Rumusan masalah yang ambil pada penelitian ini adalah :

1. Bagaimana menganalisis kinerja karyawan pada proses *training* dengan menggunakan data mining algoritma C4.5?
2. Bagaimana hasil pengujian analisis *datamining* algoritma C4.5 dalam memprediksi kinerja karyawan berdasarkan proses *training* pada PT.Schneider Manufaturing Batam ?

1.5 Tujuan Penelitian

Berikut tujuan dalam pengkajian nya ialah :

1. Buat menganalisa kinerja karyawan pada proses *training* dengan menggunakan *datamining* algoritma C4.5.
2. Untuk mengetahui hasil analisa *datamining* algoritma C4.5 dalam memprediksi kinerja karyawan berdasarkan proses *training* pada PT.Schneider Manufaturing Batam.

1.6 Manfaat Penelitian

1.6.1 Manfaat Teoritis

Adapun Manfaat nya sebagai berikut :

1. Penelitian bisa berguna sebagai pengetahuan baru terkait dengan algoritma C4.5 .
2. Penelitian juga bermanfaat sebagai referensi oleh peneliti selanjutnya.

1.6.2 Manfaat Praktis

Berikut manfaat nya dari pengkajian ini adalah :

1. Untuk Penulis
Dapat mengembangkan pemahaman penulis mengenai prediksi kinerja karyawan dengan algoritma C4.5.
2. Bagi Kampus
Menambah referensi bagi mahasiswa Universitas Putera Batam dalam meneliti terkait dengan algoritma c4.5.
3. Bagi PT.Schneider Manufacturing Batam
Memberikan informasi terkait dengan karyawan yang kompeten dalam bekerja melalui prediksi kinerja pegawai.
4. Untuk Pengkaji selanjutnya
Pengkajian ini bisa berguna untk menambah informasi penelitian lanjutan yang berhubungan dengan kinerja karyawan pada proses *training* menggunakan algoritma C4.5.

BAB II

TINJAUAN PUSTAKA

2.1. Teori Umum

Bab ini mengenai istilah dan pengertian secara umum yang berhubungan dengan pelaksana penelitian.

2.1.1 *Knowledge Discovery In Database*

KDD adalah cara yang menemukan informasi yg diperoleh dari database. Tabel yang saling berhubungan dalam database. Informasi yang dihasilkan pada proses ini dapat digunakan sebagai basis pengetahuan untuk tujuan pengambilan keputusan. Istilah penemuan pengetahuan (KDD) dan penambangan data dalam basis data sering Dipertukarkan digunakan untuk menggambarkan proses penggalian informasi yang tersembunyi dalam database besar. Padahal, konsep kedua istilah ini berbeda tetapi saling berhubungan. Datamining adalah fase dari keseluruhan proses KDD. (Mardi 2017) Proses KDD secara garis besar dapat dijelaskan sebagai berikut:

1. Penyeleksian

Sebelum fase pembekalan dimulai pada KDD harus melakukan seleksi data dari operasional dataset. Biarkan data spesifik yang digunakan dalam proses penambangan data dalam file yang berbeda dari database operasional.

2. Pembersihan

Sebelum proses penambangan data, perlu untuk membersihkan data yang berfokus pada penemuan pengetahuan (KDD) dalam database. Proses pembersihan meliputi penghapusan data duplikat, pengecekan data yang tidak konsisten, dan koreksi kesalahan pada data, seperti kesalahan pencetakan.

Pengayaan juga dilakukan, yaitu proses “memperkaya” data yang ada dengan data atau informasi lain yang relevan yang diperlukan untuk Knowledge Discovery (KDD) dalam database, seperti data atau informasi eksternal lainnya yang diminta.

3. Perubahan

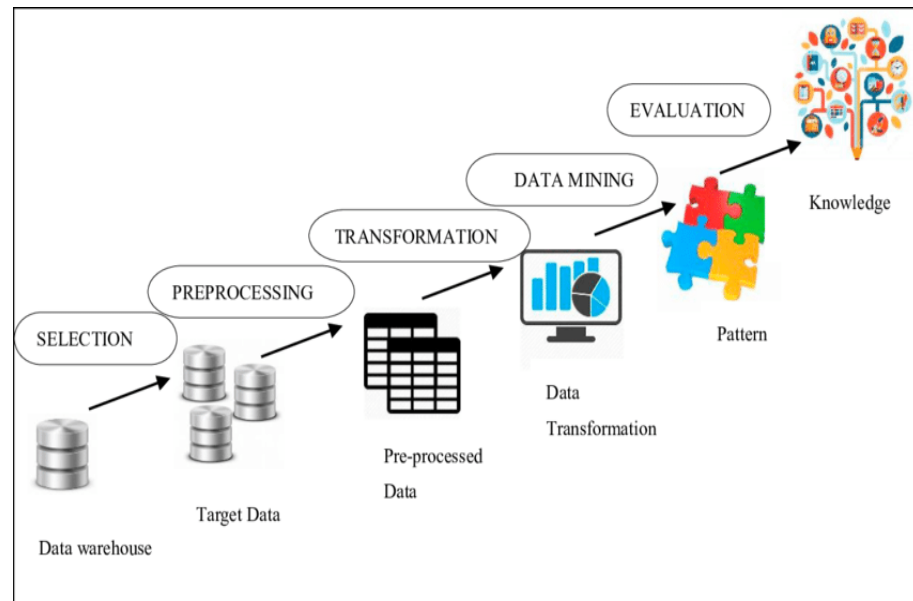
Transformasikan yang dipilih untuk membuat Data yang sesuai untuk operasi penambahan data. Pengkodean dalam Basis Data Penemuan Pengetahuan (KDD) adalah proses kreatif yang sangat bergantung pada jenis atau gaya informasi yang akan dicari dalam basis data.

4. *Data Mining*

Data mining adalah proses menggunakan teknik atau metode tertentu untuk menemukan pola atau informasi menarik dalam data yang dipilih. Teknik, metode atau algoritma dalam data mining sangat beragam. Pilihan metode atau algoritma yang benar sebenarnya tergantung pada tujuan dan proses penemuan pengetahuan yang komprehensif (KDD) dalam database.

5. Evaluasi

Pola data yang di hasilkan dari langkah informasi *data mining* di tampilkan ke wujud yang mudah di pahami pengguna .(Alwafi Ridho Subarkah 2018)



Gambar 2. 1 Langkah langkah KDD

2.1.2 Data Mining

Datamining ialah mempunyai peranan sebagai prediksi, deskripsi, klasifikasi serta asosiasi. Tidak hanya itu peranannya yang merupakan analisis asosiasi klasifikasi serta memprediksi analisis kluster, analisis *outlier*, analisis *trend* serta evolusi. (Novia, Hidayah, and Rozi 2018)

Datamining menganalisis serangkaian tampilan data untuk membuat tautan tak terduga dan merangkum data dengan cara yang berbeda dari sebelumnya. Ini mudah dan nyaman bagi pemilik data. (Utomo and Mesran 2020)

2.1.3 Kelompok *Data Mining*

Berikut ini Berdasarkan tugasnya :

1. Estimasi

Estimasi dan klasifikasi itu sama, kecuali bahwa variabel objektif penilaian merupakan tipe numerik. Model dikonfigurasi dengan harapan nilai yang diberikan sesuai dengan variabel target.

2. Memprediksi

Peramalan, penilaian dan estimasi adalah, kecuali untuk memprediksi hasil nilai di masa depan.

3. PengKlasifikasian

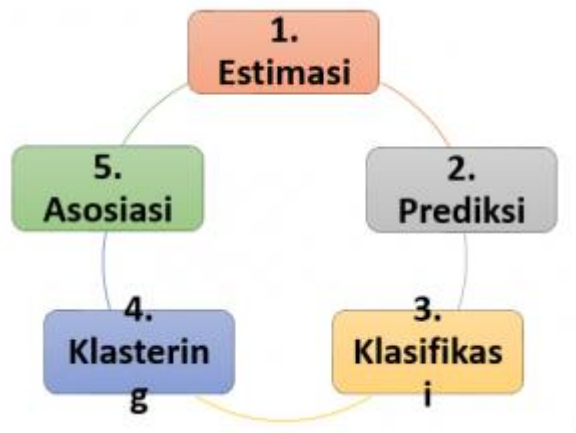
Dalam taksonomi, ada variabel kategoris target. Misalnya, klasifikasi pendapatan yaitu hasil penjualan tinggi, hasil penjualan menengah, dan hasil penjualan rendah.

4. Kluster

Cluster adalah kumpulan *record*, mengamati atau mengamati, dan membuat ruang objek sejenis. Blok adalah rangkaian record yang memiliki kesamaan dan memiliki perbedaan dengan record pada kelompok lainnya. Rakitan tidak sesuai dengan deskripsi, yaitu tidak ada variabel target dalam rakitan.

5. Mengasosiasi

Tugasnya untuk mendapatkan pola yang timbul pada suatu saat. pada global usaha sering dianggap analisis keranjang belanja. (Mardi 2017)



Gambar 2. 2 Pengelompokan Data Mining

2.1.4 Algoritma C4.5

Algoritma C4.5 artinya bentuk algoritma yg dipergunakan sebagai pembagian terstruktur yang serta bersifat prediktif yang pembentukannya adalah pohon keputusan (*Decision Tree*). Berikut langkah – langkahnya :

1. Siapkan data pelatihan yang didapatkan pada historis yang sebelumnya sudah terjadi dan dikompilasi ke dalam bab-bab eksklusif.
2. Tentukan akar pohon yang berasal dari Rute atribut yang didapat dari menghitung nilai gain untuk tiap atribut, dengan nilai gain maksimum menjadi rute pertama. Sebelum menghitung nilai gain dari atribut, terlebih dahulu menghitung nilai entropi. Berikut rumus yang dipakai:
 - a. Rumus menghitung nilai entropi

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

Rumus 2. 1 Rumus Entropy

Keterangan :

S = himpunan

A = kelengkapan yang dipakai

n = jumlah partisi pada bagian S

p_i = proporsi dari S_i ke S

b. Kemudian hitung nilai gain menggunakan persamaan

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

Rumus 2. 2 Perhitungan Gain

Keterangan :

S = himpunan

A = keterangan

n = jumlah partisi di keterangan atau kelengkapan bagian A

$|S_i|$ = jumlah yang ada pada partisi i

$|S|$ = jumlah yang ada pada S

c. Kembali lagi ke cara 2 sampai semua *record* terpartisi.

d. Jika semua record memiliki kelas N yang sama, atribut dalam record tidak dipartisi ulang, dan record di cabang tidak kosong, proses partisi pohon keputusan berakhir..(Mardi 2017)

2.1.5 WEKA (*Waikato Environment for Knowledge Analysis*)

Menurut pendapat WEKA (Lingkungan Waikato untuk Analisis Pengetahuan) sebagai bagian dari arti pembelajaran mesin yang belum pernah dirilis, WEKA didistribusikan di Universitas Waikato di Selandia Baru yang digunakan untuk pendidikan dan penelitian. WEKA adalah perangkat lunak yang bersifat open source, sehingga siapa pun dapat mengakses perangkat lunak ini.

Pengujian dengan WEKA akan jauh lebih mudah dan lebih mudah untuk membuat pohon keputusan, hampir seperti perhitungan yang dilakukan secara manual. (Elisa & Ervina, 2018) .

2.2. Teori Khusus

Teori ini membahas istilah dan pengertian secara khusus berkaitan dengan pelaksanaan penelitian.

2.2.1 Karyawan

Menurut Ayat 2 Pasal 1 Undang-Undang Nomor 13 Tahun 2003 tentang Ketenagakerjaan, pekerja adalah orang yang dapat menghasilkan barang dan jasa yang dapat memenuhi kebutuhannya dan kebutuhan masyarakat, baik di dalam maupun di luar hubungan kerja. Ini menyatakan Dari pengertian di atas, jelas bahwa penggunaan tenaga kerja adalah tenaga kerja yang mempekerjakan dalam bentuk apapun perusahaan atau individu dengan menggunakan upah. (Novia et al. 2018)

2.2.2 Kinerja Karyawan

Memahami kinerja karyawan Keberhasilan kegiatan organisasi sangat tergantung pada peran karyawan di tempat kerja. Semakin baik Sebaliknya semakin baik kinerja karyawan maka semakin baik pula kinerja perusahaan, tetapi semakin buruk kinerja karyawan maka semakin jauh tujuan yang ditetapkan perusahaan dari tujuan yang dipersyaratkan. Kinerja organisasi berarti bahwa kegagalan merupakan bagian yang tidak terpisahkan dari pelaksanaan tugas organisasi baik di forum pemerintah maupun swasta. Achievements berasal dari bahasa achievement atau pencapaian aktual (actual achievement atau achievement by individu atau organisasi). Kinerja seorang pegawai (job performance) adalah

hasil kerja secara kuantitatif dan kualitatif yang dilakukan oleh pegawai dalam melaksanakan pekerjaan sesuai dengan tanggung jawab yang diberikan kepada pegawai tersebut. (Onainor 2019)

Faktor Kinerja Karyawan adalah kegiatan yang paling umum dalam suatu organisasi, untuk melakukan sesuatu yang berhubungan dengan pekerjaan, jabatan, atau peran dalam organisasi. Faktor kinerja terdiri dari faktor internal dan eksternal dan faktor internal (kecenderungan), yaitu faktor yang berhubungan dengan karakteristik seseorang. Sedangkan faktor eksternal adalah faktor yang mempengaruhi kinerja seseorang dari lingkungan. Sikap, sikap dan perilaku rekan kerja, bawahan atau manajer, fasilitas kerja dan budaya organisasi, dll. Berikut faktor yang memengaruhi hasil kinerja karyawan, yaitu :

1. Kualitas kerja adalah tingkat baik buruknya pekerjaan yang diterima seorang pegawai yang dapat ditinjau dari segi ketelitian, keanggunan, keterampilan, dan kemampuan.
2. Kerja Tim (Teamwork), Cari tahu bagaimana seorang karyawan bekerja dengan menggunakan orang lain untuk menyelesaikan suatu pekerjaan.
3. Kreativitas adalah kemampuan pegawai untuk menyelesaikan pekerjaannya dengan caranya sendiri atau atas prakarsa sendiri, yang disebut kemampuan efektif, efisien dan mampu menciptakan perubahan baru bagi pemulihan dan kemajuan organisasi.
4. Inovasi adalah tentang mewujudkan sesuatu yang baru atau memperbarui yang sudah ada. Inovasi dapat diartikan sebagai suatu bentuk kreativitas yang tidak hanya dapat menghasilkan ide-ide baru dan bermanfaat, tetapi

juga mengimplementasikan, mengkomersilkan dan memberikan dampak yang besar.

5. Mengambil inisiatif adalah melakukan sesuatu atau melakukan sesuatu tanpa harus memahami apa yang harus dilakukan sebelumnya..(Novia et al. 2018)

2.2.3. *Trainer*

Pelatih Hal ini penting untuk pengembangan profesional yang sukses. Penguasaannya terhadap materi latihan tidak cukup menjadi andalan jika tidak dibarengi dengan penggunaan beberapa keterampilan lainnya. Keterampilan lain ini berkaitan dengan penggunaan pemahaman tentang metode dan materi dalam pelatihan dengan menggunakan kegiatan profesional pelatih. Selain itu, memahami pengembangan karir itu sendiri merupakan kompetensi penting dari seorang pelatih yang sukses. Pembina atau coach berperan penting dalam pengembangan dan perubahan organisasi, pencapaian tujuan, kompetensi, pengetahuan, dan kapabilitas yang dibutuhkan peserta. Seorang pelatih memiliki dampak pada pelatihan peserta. Pelatih adalah seseorang yang mengajarkan keterampilan orang untuk mempersiapkan pekerjaan. Pelatih adalah seorang pelatih atau sering disebut menggunakan pelatih yang menerima dan menguji pengetahuan dan kemudian menghasilkan hasil tes dengan menggunakan beberapa elemen pelengkap (yang dapat dikurangi atau ditambahkan) sehingga merupakan metode pembelajaran yang efektif dan tepat, sehingga peserta pelatihan mampu menyerap pengetahuan ini dengan cara yang dapat diterapkan dan dapat melakukan hal itu dengan menggunakan apa yang disajikan atau diarahkan oleh pelatih. (Heriyanto et al. 2020)

2.3. Penelitian Terdahulu

Peneliti juga mencapai beberapa hasil penelitian sebelumnya terkait analisis menggunakan algoritma C4.5, yang dirangkum sebagai berikut:

1. pengkajian penggunaan Data Mining dengan c4.5 untuk mendapatkan hasil kinerja karyawan terbaik yang dilakukan oleh (Novia et al. 2018) Penelitian dilakukan pembagian terstruktur mengenai atau segmentasi bersifat prediktif yang digunakan untuk pohon keputusan untuk membantu mempermudah pihak pusat penjaminan mutu (PPM) dengan memiliki taraf akurasi yang termasuk penjabarannya akurat.
2. Dari Penelitian menentukan kinerja karyawan dalam berkerja dengan perusahaan memakai algoritma C4.5 (Ferdiansyah and Goeirianto 2020) Survey karyawan serta hasil yang didapatkan akurat menjadi pengambil Sebuah keputusan yang memprediksi loyalitas karyawan kepada perusahaan.
3. Dari Survey Dampak Kompensasi Terhadap Kinerja Karyawan di Kantor Cabang Persada dan PT Arminareka Perdana Surabaya di Indonesia (Wati 2019) Menggunakan hipotesis asosiatif dan komperatif membuktikan nilai yang besar dan signifikan .
4. Dari Penelitian Algoritma Klasifikasi Kandidat C4.5 untuk Kompetisi Cerdas di SMP dengan Aplikasi weka (Ardiansyah 2019) Klasifikasi data kandidat potensial sebagai sarana penerapan algoritma C4.5 terbukti memiliki akurasi yang diperoleh dari evaluasi klasifikasi tersebut.

5. Dari Penelitian Analisa Kinerja Algoritma C4.5 untuk menentukan target keuntungan (2017, Nofitri, and Sembiring 2017) Mendapatkan informasi tambahan bagi pengelola perusahaan di saat menggunakan rule menjadi alat dalam memprediksi pencapaian profit
6. Dari Penelitian Menerapkan penambahan data untuk memprediksi hasil penjualan menggunakan algoritma C4.5 (Eska 2018) Dalam memakai data mining algoritma C4.5 menggunakan cara dagang yang meliputi pemilihan atribut menjadi akar dan cabang-cabang dan setiap cabang akan pada ulangi sampai Beberapa aturan dapat diperoleh dari penyelesaian pohon keputusan dengan kelas yang sama.
7. Sebuah studi dari tentang penggunaan data mining dengan algoritma C4.5 pada klasifikasi fasilitas medis di negara bagian Indonesia. (Algoritma et al. 2020) Dengan analisis data mining, hasilnya menjadi model pohon keputusan, menentukan peringkat Indeks Pembangunan Manusia, dan menggunakan KFold Cross Validation untuk mendapatkan skor akurasi.
8. Dari Penelitian penggunaan Data Mining pada menentukan mobil yang akan dipilih Dengan memakai Algoritma C4.5 (Harahap 2018) Pada hal ini penjabaran pembelian mobil berdasarkan penerapan algoritma C4.5 membantu memberi bahan pertimbangan pembeli buat memilih mobil yang akan di belinya.
9. Dari Penelitian Penerapan algoritma data mining C4.5 untuk mendapatkan catatan kinerja instruktur STT Harapan Medan, Pemilihan catatan kinerja instruktur menurut tiga atribut input (publikasi,

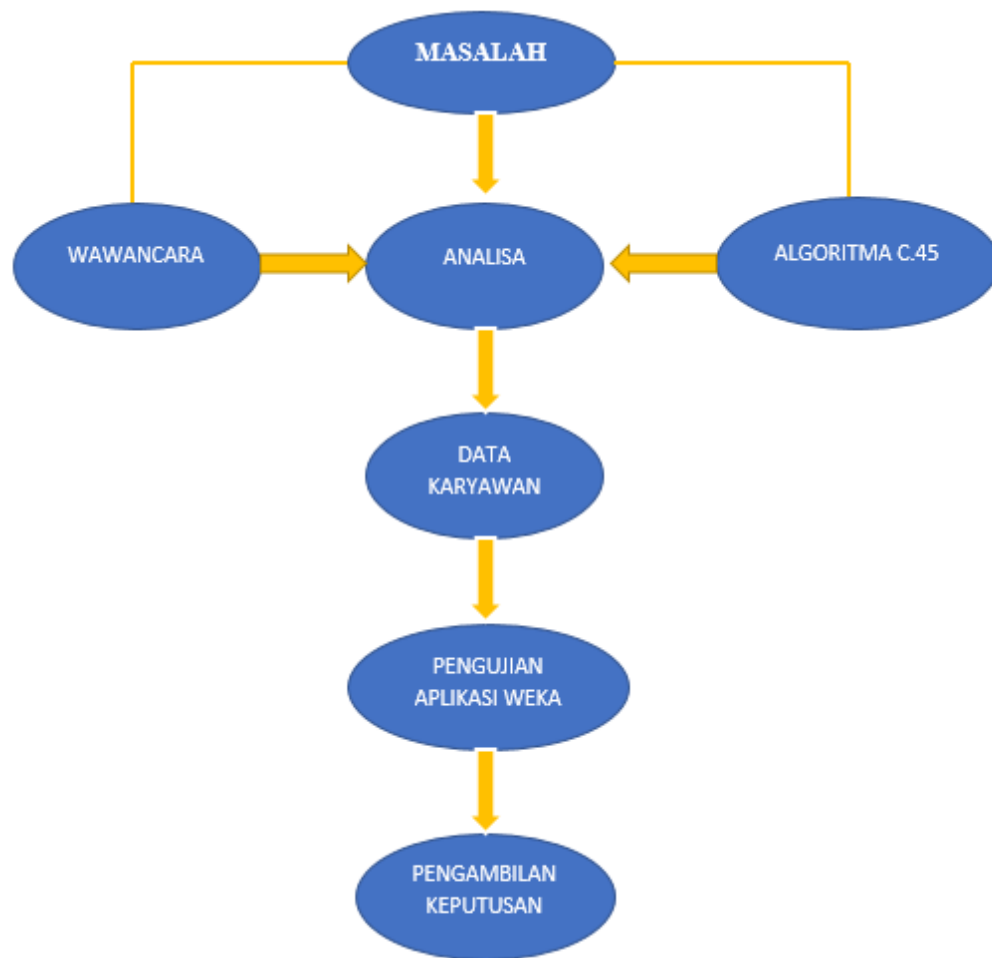
komitmen, pendidikan) dan satu atribut output (nilai kinerja (rendah, cukup, tinggi)).

10. Dari Survey penerapan algoritma data mining C4.5 untuk mengetahui kepuasan pelanggan di Grand Zuri Hotel Dumai (Desyanti 2019) Setelah di lakukan penelitian menggunakan algoritma C4.5 diperoleh yang akan terjadi Jika pelayanannya sangat puas.
11. Dari Penelitian Penerapan Algoritma C4.5 Dalam Memprediksi Ketersediaan Uang Pada Mesin ATM (Syahputra, Hartono, and Rosnelly 2021) Model pohon keputusan yang di hasilkan merupakan menjadikan variabel saldo sebagai akar. Lalu melakukan uji coba dengan metode *K-Fold Cross Validation* hasilnya tingkat akurasi lebih tinggi.
12. Dari Penelitian Menentukan faktor kesesuaian penerimaan karyawan menggunakan algoritma C4.5 (pohon keputusan) pada perusahaan outsourcing staff PT (Sahputra et al. 2020) Dengan memakai metode algoritma C4.lima atribut yang di pakai dalam peneltian berdasarkan Usia, tempat tinggal, jenis kelamin, jabatan, bidang studi, referensi, tingkat pendidikan, dan tes. Saat menggunakan metode solusi, tingkat akurasinya sangat tinggi.
13. Dari penelitian *Predict lung cancer survival based on performance using support vector machine C4.5 algorithm classification techniques for health care analysis* (Pradeep and Naveen 2018) Deretan data penyakit kanker paru-paru beserta dengan data pasien barudi pakai buat mengevaluasi kinerja, perbandingan pada dasarnya akurasi dengan menggunakan algoritma C4.5.

14. Dari Penelitian *Digital transformation in managing agile and employee engagement C4.5 Algorithm* (Elacio, Balazon, and Lacatan 2020)
Penggunaan algoritma C4.5 serta deretan data yang pada kalibrasi membarui pendekatan manual berasal retensi karyawan menjadi pendekatan digital memberikan akurasi yang sangat tinggi.
15. Dari Penelitian *Predicting employee attendance factors using the C4.5 algorithm, random tree, and random forest* (Fahlapi et al. 2020) Dari hasil analisa penulis faktor keterlambatan karyawan menggunakan 3 prosedur pemecahan yaitu bernilai akurasi.

2.4. Kerangka Pemikiran

Dibawah ini adalah kerangka pemikiran dari pengkajian dan juga penjelasannya



Gambar 2. 3 Kerangka Berfikir

BAB III

METODE PENELITIAN

3.1 *Design* Penelitian

Pada *design* pengkajian memakai jabaran penelitiannya yang dapat dilihat berikut ini :



Gambar 3. 1 Design Penelitian

Keterangan:

Penjelasan *design* dari penelitian sesuai dengan model atau gambar di atas, yaitu:

1. Deskripsi masalah

Langkah awal penelitian ini penulis akan melakukan survey awal guna mengetahui permasalahan objek kemudian masalah tersebut akan di deskripsikan sesuai dengan topik penelitian.

2. Melakukan Analisa Terhadap Masalah

pengkaji menganalisa dari berbagai persoalan yang akan penulis kaji berdasarkan tindak lanjut setelah menggambarkan persoalan yang di hadapi.

3. Mempelajari Literatur

pengkaji berusaha mengkaji dan juga mempelajari berbagai asal teori yang mendukung penelitian jurnal supaya dapat menjadi panduan bagi penulis pada melakukan penelitian.

4. Pengumpulan Data

peneliti mewawancarai subjek penelitian menggunakan menanyai pemilik dan karyawan PT Schneider Manufacturing Batam untuk mendapatkan beberapa data serta informasi yang diharapkan.

5. Analisa Menentukan Kinerja Karyawan Dengan Jenis Algoritma Dari C4.5

Pada tahap ini penggunaan jenis algoritma dari C4.5 di mulai dari proses pembuatan pohon keputusan sehingga membuat penentuan kinerja karyawan dalam bentuk aturan.

6. Implementasi Jenis Algoritma C4.5

Penulis menggunakan data buat menentukan *integritas* menjadi *root* dari pohon keputusan, sehingga menghitung nilai gain tertinggi berasal seluruh komponen yang ada pada data yg ditentukan, kemudian menentukan nilai *gain* pada proses pembentukan struktur pohon keputusan.

7. Pengujian Hasil

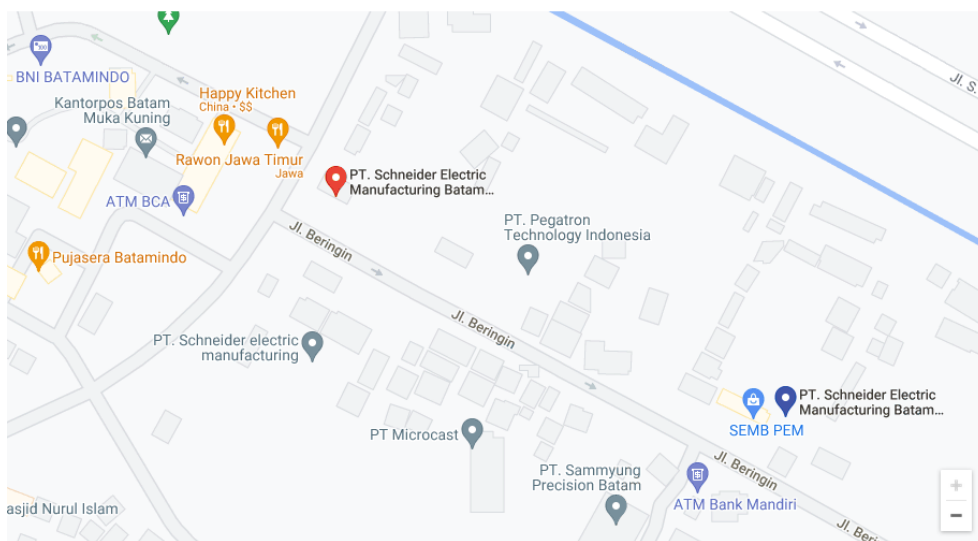
Sehingga di tahap akhir,peneliti melakukan pengujian setelah menyelesaikan perhitungan nilai gain dan menghasilkan pohon keputusan menggunakan menggunakan *software WEKA*.

3.2 Objek Penelitian

Pada pokok bahasan penelitian ini, penulis menjelaskan tempat penelitian dan jadwal selama kegiatan penelitian sebagai berikut.

3.2.1. Lokasi Penelitian

Penelitian ini berlokasi di PT Schneider Manufacturing Batam yang beralamat di Jalan Beringin Lot 1, Batamindo Industrial Park Batam.



Sumber :*Google Maps*(2021)

Gambar 3. 2 Lokasi Objek Penelitian

3.3 Populasi dan Sampel

Dalam tahap ini penulis akan menjelaskan mengenai keterangan yang di miliki oleh populasi dan sampel yang di jadikan bahan dalam penelitian sebagai berikut.

3.3.1 Populasi

Populasi dalam penelitian ini adalah data kinerja karyawan di PT Schneider Manufacturing Batam dalam periode 6 bulan terakhir yaitu tahun 2021, yaitu sebanyak 210 data sampel.

3.3.2 Sampel

Data yang di ambil sampel pada penelitian ini, peneliti akan mengambil data kinerja karyawan sebanyak 210 data sampel yang mana semua data di ambil dari seluruh data yang ada pada populasi dalam periode 6 bulan terakhir yaitu tahun 2021. Kemudian dilakukan pengambilan sampel secara acak dengan teknik *simple random sampling* menggunakan rumus slovin berikut ini :

$$n = \frac{N}{1 + Ne^2}$$

Rumus 3. 1 Perhitungan rumus slovin

Keterangan :

n = Sampel yang akan di cari

N = Jumlah populasi

e = Batas toleransi kesalahan

Telah di tetapkan populasi pada penelitian ini sebanyak 210 data sampel, serta batas toleransi kesalahannya yaitu 10%, sehingga perhitungannya sebagai berikut :

$$n = \frac{210}{1 + 210 (0,1)^2}$$

$$n = \frac{210}{1 + 210 \times 0,01}$$

$$n = \frac{210}{3}$$

$$n = 70$$

Dengan rumus slovin hasil perhitungan sudah di dapatkan, maka sampel yang akan di gunakan adalah sebanyak 70 data sampel.

3.4 Variabel Penelitian

Dalam sebuah penelitian untuk menganalisa data membutuhkan variabel-variabel, data yang akan di olah adalah data kinerja karyawan periode 3 tahun terakhir yaitu tahun 2019, 2020 dan 2021 dari PT Schneider Manufacturing Batam. Beberapa indikator yang di gunakan dalam penelitian khususnya yang berkaitan dengan variabel untuk menentukan kinerja karyawan, sebagai berikut :

1. Kinerja
2. Kedisiplinan
3. Target
4. Kerja sama
5. Kemampuan Kerja

Untuk tabel keputusannya adalah

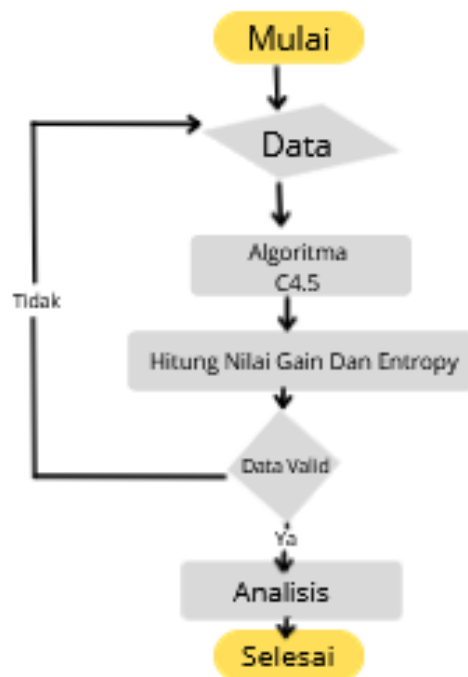
1. Layak
2. Tidak Layak.

3.5 Teknik Pengumpulan Data

Memperoleh informasi yang dibutuhkan dengan cara mengumpulkan data untuk mencapai tujuan penelitian, atau mengetahui informasi dan keadaan tempat penelitian sesuai dengan kebutuhan penelitian. Metode akuisisi data seperti:

1. Teknik studi literatur adalah mengkaji tentang sastra dalam bidang penelitian. Teknik ini menjelaskan perihal topik yang telah di pelajari, di temukan serta di tentukan, disusun berdasarkan waktu atau topik. Peneliti mengkaji banyak sekali jurnal yg berhubungan menggunakan persoalan yang di hadapi.
2. observasi adalah pengumpulan data pengamatan langsung oleh peneliti lapangan, peneliti langsung di PT Schneider manufacturing Batam.
3. Teknik Dokumentasi adalah teknik yang membahas informasi tertulis yang ada dilokasi penelitian, yang bekerjasama tentang masalah data yang pada hadapi atau masalah yang ada, kemudian di kumpulkan data tersebut sesuai kebutuhan penelitian yg dilakukan. Peneliti melakukan pengamatan serta meneliti data kinerja karyawan PT Schneider Manufacturing Batam.
4. Metode wawancara adalah memberikan peneliti pertanyaan dan jawaban secara lisan dan mendengarkan informasi dan komentar secara langsung. Peneliti menanyakan kepada pemilik dan karyawan PT Schneider Manufacturing Batam.

3.6 Model Penelitian



Gambar 3. 3 Model Penelitian

Dari tahapan ini penulis melakukan penelitian menggunakan data mining dan jenis algoritma C4.5. Awalnya dilakukan klasifikasi sesuai dengan kategori data karyawan, selanjutnya dilakukan perhitungan algoritma C4.5 untuk mendapatkan hasil dari nilai *gain* dan *entropy*, jika data yang digunakan valid maka dilakukan analisis dengan membuat struktur dari *decision tree* secara berulang dan pada simpul *decision* tidak memiliki keturunan lagi sampai menghasilkan pola berupa aturan dasar penilaian menentukan kinerja karyawan pada PT Schneider Manufacturing Batam.