

**IMPLEMENTASI *DATA MINING* UNTUK MENENTUKAN  
KUALITAS WORTEL DI PT PADMA MEKAR LESTARI  
MENGUNAKAN METODE *C4.5***

**SKRIPSI**



**Disusun Oleh :**

**Nama : REYNALDI GUNAWAN**

**NIM : 20181000008**

**TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS BUDDHI DHARMA**

**TANGERANG**

**2022**

**IMPLEMENTASI DATA MINING UNTUK MENENTUKAN  
KUALITAS WORTEL DI PT PADMA MEKAR LESTARI  
MENGUNAKAN METODE C4.5**

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk kelengkapan gelar kesarjaan pada**

**Program Studi Teknik Informatika**

**Jenjang Pendidikan Strata 1**



**Disusun Oleh :**

**Nama : REYNALDI GUNAWAN**

**NIM : 20181000008**

**TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS BUDDHI DHARMA**

**TANGERANG**

**2022**

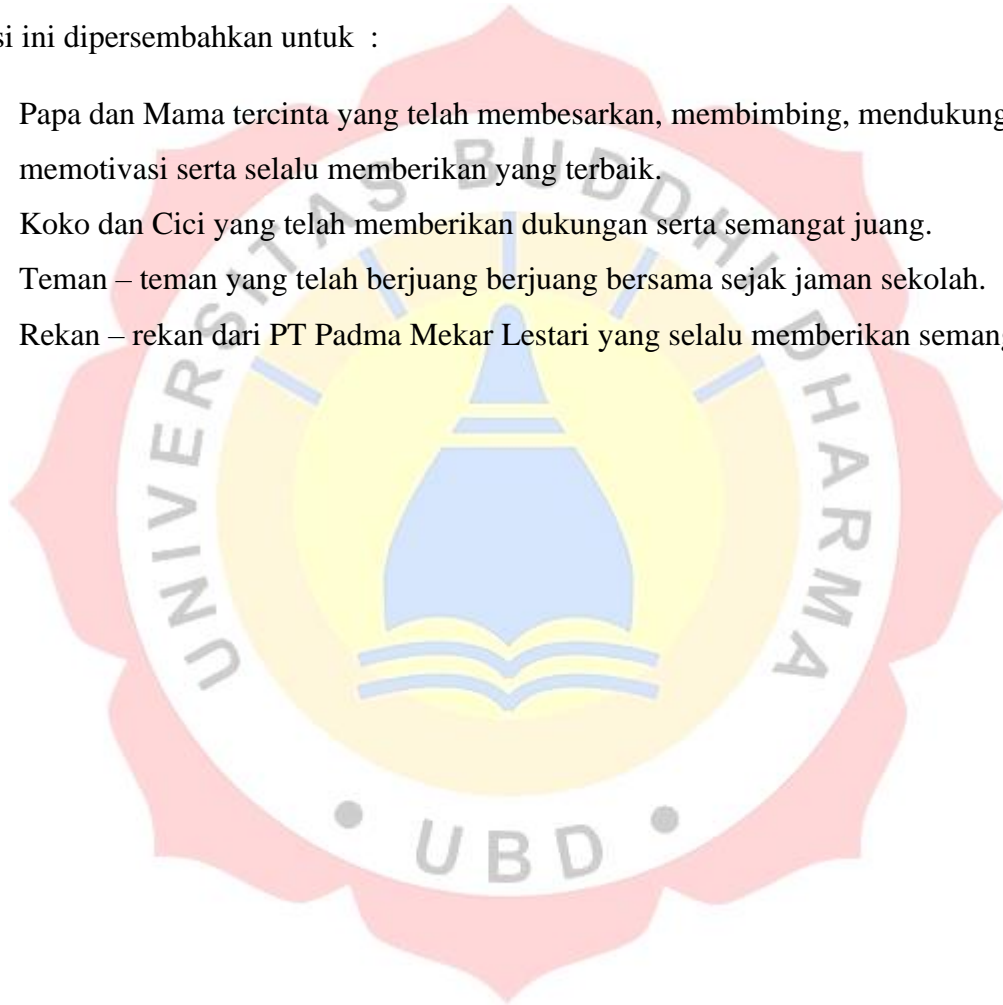
## LEMBAR PERSEMBAHAN

*“Start where you are. Use what you have. Do what you can.”*

-Arthur Ashe-

Dengan mengucapkan puji dan syukur kepada Tuhan Yang Maha Esa dan Sang Triratna, Skripsi ini dipersembahkan untuk :

1. Papa dan Mama tercinta yang telah membesarkan, membimbing, mendukung, memotivasi serta selalu memberikan yang terbaik.
2. Koko dan Cici yang telah memberikan dukungan serta semangat juang.
3. Teman – teman yang telah berjuang bersama sejak jaman sekolah.
4. Rekan – rekan dari PT Padma Mekar Lestari yang selalu memberikan semangat.



**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PERNYATAAN KEASLIAN SKRIPSI**

Yang bertanda tangan dibawah ini,

NIM : 20181000008  
Nama : Reynaldi Gunawan  
Jejang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : Database Development

Dengan ini, saya menyatakan bahwa:

- a. Skripsi ini adalah skripsi asli yang belum pernah diajukan untuk mendapatkan gelar akademik. Baik di Universitas Buddhi Dharma maupun di perguruan tinggi atau kampus lainnya.
- b. Skripsi ini dibuat sendiri tanpa ada bantuan pihak lain, terkecuali arahan dari dosen pembimbing.
- c. Dalam skripsi ini tidak terdapat karya ataupun pendapat yang ditulis maupun dipublikasikan orang lain kecuali, tertulis jelas dan juga dicantumkan.
- d. Dalam skripsi ini tidak terdapat pemalsuan seperti artikel, jurnal, buku serta tanda tangan yang ada didalamnya.
- e. Lembar pengesahan ini saya buat dengan sesungguhnya tanpa ada paksaan.

Sekali lagi saya menyatakan bahwa, skripsi ini saya buat dengan sungguh – sungguh tanpa ada paksaan atau unsur lain didalamnya.

Tangerang, 02 Agustus 2022

Yang menyatakan,



Reynaldi Gunawan

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH**

NIM : 20181000008  
Nama : Reynaldi Gunawan  
Jenjang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : Database Developmen

Dengan ini saya menyampaikan bahwa, demi pengembangan ilmu pengetahuan generasi berikutnya saya Reynaldi Gunawan menyetujui serta memberikan izin kepada pihak kampus Universitas Buddhi Dharma atas hak – hak bebas royalti dan sebagainya atas skripsi yang saya buat dengan judul “IMPLEMENTASI *DATA MINING* UNTUK MENENTUKAN KUALITAS WORTEL DI PT PADMA MEKAR LESTARI MENGGUNAKAN METODE *C4.5*” untuk disimpan , dikelola di dalam penyimpanan data serta di publikasikannya di internet atau media lainnya untuk kepentingan akademik kedepannya selama serta mencantumkan nama saya sebagai pemilih hak cipta.

Sekali lagi saya sampaikan bahwa pernyataan ini saya buat dengan sungguh – sungguh tanpa ada paksaan dari pihak lain ataupun hal lainnya.

Tangerang, 02 Agustus 2022

Yang menyatakan,



Reynaldi Gunawan

**UNIVERSITAS BUDDHI DHARMA**

**LEMBAR PENGESAHAN PEMBIMBING**

**IMPLEMENTASI *DATA MINING* UNTUK MENENTUKAN  
KUALITAS WORTEL DI PT PADMA MEKAR LESTARI  
MENGUNAKAN METODE *C4.5***

Dibuat oleh:

NIM : 20181000008

Nama : Reynaldi Gunawan

Telah disetujui untuk dipertahankan dihadapan Tim Penguji Ujian

Komprehensif

Program Studi Teknik Informatika

Peminatan *Database Development*

2022

Tangerang, 27 Juni 2022

Disahkan oleh

**Pembimbing,**



**Susanto Hariyanto, S.Kom., M.Kom.**

**NIDN. 0428128601**

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PENGESAHAN SKRIPSI**  
**IMPLEMENTASI *DATA MINING* UNTUK MENENTUKAN**  
**KUALITAS WORTEL DI PT PADMA MEKAR LESTARI**  
**MENGGUNAKAN METODE *C4.5***

Dibuat Oleh:

NIM : 20181000008

Nama : Reynaldi Gunawan

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian

Komprehensif

Program Studi Teknik Informatika

Database Development

2022

Disahkan oleh,

Tangerang, 02 Agustus 2022

Dekan,



Dr.Eng. Ir. Amin Suyitno, M.Eng

NIDK: 8826333420

Ketua Program Studi,



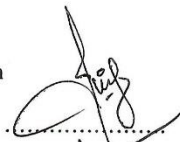
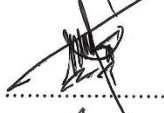

Hartana Wijaya, S.Kom., M.Kom.

NIDN: 0412058102

## LEMBAR PENGESAHAN TIM PENGUJI

Nama : Reynaldi Gunawan  
NIM : 20181000008  
Fakultas : Sains dan Teknologi  
Judul Skripsi : IMPLEMENTASI *DATA MINING* UNTUK MENENTUKAN  
KUALITAS WORTEL DI PT PADMA MEKAR LESTARI  
MENGUNAKAN METODE *C4.5*

Dinyatakan LULUS setelah mempertahankan di depan Tim Penguji pada hari Selasa,  
02 Agustus 2022

	Nama Penguji :	Tanda Tangan :
Ketua Sidang	: <b>Dicky Surya Dwi Putra, S.Kom., M.Kom</b> NIDN : 0427068701	
Penguji I	: <b>Hartana Wijaya, S.Kom., M.Kom</b> NIDN : 0412058102	
Penguji II	: <b>Susanto Hariyanto, S.Kom., M.Kom</b> NIDN : 0428128601	

Mengetahui,

**Dekan Fakultas Sains Dan Teknologi**

  
**Dr. Eng. Ir. Amin Suvitno, M.Eng**

**NIDK: 8826333420**



## KATA PENGANTAR

Dengan mengucapkan Puji Syukur ke hadapan Tuhan Yang Maha Esa, yang telah memberikan Rahmat dan karunia-Nya kepada penulis sehingga dapat menyusun dan menyelesaikan skripsi ini, dengan judul **IMPLEMENTASI DATA MINING UNTUK MENENTUKAN KUALITAS WORTEL DI PT PADMA MEKAR LESTARI MENGGUNAKAN METODE C4.5**. Tujuan utama dari pembuatan skripsi ini adalah sebagai salah satu syarat kelengkapan dalam menyelesaikan program pendidikan Strata I Program Studi Teknik Informatika di Universitas Buddhi Dharma. Dalam penyusunan skripsi ini penulis banyak menerima bantuan dan dorongan baik moril maupun materil dari berbagai pihak, maka pada kesempatan ini penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Suryadi Winata, SE, MM, M.Si., Ak., CA, CMA, CBV, ACPA, CPA (Aust.), sebagai Plt. Rektor Universitas Buddhi Dharma.
2. Bapak Dr. Eng, Ir. Amin Suyitno, M.Eng., Dekan Fakultas Sains dan Teknologi.
3. Bapak Hartana Wijaya, S.Kom., M.Kom., sebagai Ketua Program Studi Teknik Informatika.
4. Bapak Susanto Hariyanto, S.Kom., M.Kom, sebagai pembimbing yang telah membantu dan memberikan dukungan serta harapan untuk menyelesaikan penulisan Skripsi ini.
5. Orang tua dan keluarga yang selalu memberikan dukungan baik moril dan materil.
6. Teman-teman yang selalu membantu dan memberikan semangat.

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan skripsi ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan dimasa yang akan datang. Akhir kata semoga skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Tangerang, 02 Agustus 2022



Reynaldi Gunawan

# IMPLEMENTASI DATA MINING UNTUK MENENTUKAN KUALITAS WORTEL DI PT PADMA MEKAR LESTARI MENGGUNAKAN METODE C4.5

(104 Halaman + xv / 14 tabel / 32 gambar / 26 pustaka / lampiran)

## ABSTRAK

Wortel merupakan salah satu sayuran yang menyimpan karbohidrat yang sangat besar yang pada umumnya digunakan untuk pengolahan sebagai makanan maupun minuman. Pengolahan atau penyortiran ulang wortel harus dilakukan agar dapat mengetahui mana kualitas wortel yang terbaik, agar dapat ditentukan harga jual wortel berdasarkan tingkat kualitas yang dimilikinya. Maka dari itu untuk dapat membantu dalam menentukan kualitas wortel yang ada, dapat menggunakan bidang ilmu data mining yang dimana pada penelitian ini akan menggunakan algoritma C4.5 dengan serta menggunakan aplikasi Rapidminer yang mendapatkan accuracy sebesar 82,95%.

**Kata Kunci** : Wortel, Data Mining, Algoritma C4.5, Rapidminer

*IMPLEMENTATION OF DATA MINING TO DETERMINE THE QUALITY OF CARROTS  
IN PT PADMA MEKAR LESTARI USING THE C4.5 METHOD*

*(104 pages + xv / 14 tables / 32 images / 26 libraries / appendices)*

**ABSTRACT**

*Carrots are one of the vegetables that store very large carbohydrates which are generally used for processing as food and drinks. Processing or re-sorting of carrots must be done in order to find out which carrots are the best quality, so that the selling price of carrots can be determined based on the level of quality they have. Therefore, to be able to assist in determining the quality of existing carrots, you can use the field of data mining science which in this study will use the C4.5 algorithm and use the Rapidminer application which gets an accuracy of 82.95%.*

**Keywords** : *Carrots, Data Mining, C4.5 Algorithm, Rapidminer*

## DAFTAR ISI

**LEMBAR JUDUL LUAR SKRIPSI**

**LEMBAR JUDUL DALAM SKRIPSI**

**LEMBAR PERSEMBAHAN**

**LEMBAR PERNYATAAN KEASLIAN SKRIPSI**

**LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH**

**LEMBAR PENGESAHAN PEMBIMBING**

**LEMBAR PENGESAHAN PENGUJI**

**LEMBAR PENGESAHAN SKRIPSI**

KATA PENGANTAR.....	v
ABSTRAK.....	vi
<i>ABSTRACT</i> .....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN.....	xiv
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang Masalah .....	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	4
1.4 Ruang Lingkup .....	4
1.5 Tujuan dan Manfaati Penelitian .....	4
1.5.1 Tujuan.....	5
1.5.2 Manfaat.....	5
1.6 Metode Penelitian .....	6
1.6.1 Teknik Pengumpulan Data .....	6
1.7 Sistematika Penulisan .....	6
<b>BAB II LANDASAN TEORI.....</b>	<b>8</b>
2.1 Teori Umum.....	8

2.1.1	<i>Database</i> .....	8
2.1.2	Sistem .....	10
2.1.3	Internet.....	10
2.1.4	<i>Data</i> .....	13
2.1.5	Komputer .....	15
2.1.6	Implementasi .....	16
2.1.7	Aplikasi.....	17
2.2	Teori Khusus.....	19
2.2.1	Klasifikasi.....	19
2.2.2	Algoritma.....	20
2.2.3	<i>Data Mining</i> .....	21
2.2.4	<i>Decision Tree</i> atau Pohon Keputusan.....	26
2.2.5	Algoritma <i>C4.5</i> .....	29
2.3	Teori Analisa dan Perancangan .....	32
2.3.1	XAMPP .....	32
2.3.2	Netbeans .....	32
2.3.3	Java.....	34
2.3.4	phpMyAdmin .....	35
2.3.5	MySQL .....	36
2.3.6	JDK ( <i>Java Development Kit</i> ).....	37
2.3.7	Rapidminer .....	39
2.3.8	UML ( <i>Unified Modeling Language</i> ) .....	39
2.3.9	ERD ( <i>Entity Relationship Diagram</i> ).....	49
2.4	Tinjauan Studi.....	49
2.4.1	Penelitian Fauriatun Helmiah .....	49
2.4.2	Penelitian Asep Muhidin .....	51
2.4.3	Penelitian Mhd Wendico Herdian .....	53
2.4.4	Rangkuman Model Penelitian .....	55
2.5	Kerangka Pemikiran .....	58
<b>BAB III ANALISA SISTEM YANG BERJALAN .....</b>		<b>59</b>
3.1	Tinjauan Umum Perusahaan .....	59
3.1.1	Sejarah Perusahaan .....	59
3.1.2	Visi Perusahaan .....	61
3.1.3	Misi Perusahaan.....	61
3.1.4	Struktur Organisasi .....	61

3.1.5	Fungsi Setiap Bagian Organisasi.....	62
3.2	Identifikasi Kebutuhan Sistem.....	64
3.2.1	Kebutuhan Sumber <i>Data</i> .....	64
3.3	Alternatif Pemecahan Masalah .....	66
3.3.1	Konstruksi Algoritma <i>C4.5</i> .....	66
3.3.2	Alur Metode.....	74
3.3.3	Konstruksi Metode .....	75
3.4	Perancangan UML .....	76
3.4.1	<i>Use Case</i> .....	76
3.5	Perancangan Layar, <i>Menu</i> dan <i>Database</i> .....	77
3.5.1	Perancangan Halaman <i>Login</i> .....	77
3.5.2	Perancangan Halaman <i>Register</i> .....	79
3.5.3	Perancangan Halaman <i>Log Out</i> .....	80
3.5.4	Perancangan Halaman <i>Menu</i> .....	80
3.5.5	Perancangan Halaman <i>Input Data</i> .....	81
3.5.6	Perancangan Halaman Test Klasifikasi .....	81
3.5.7	Perancangan <i>Database</i> .....	82
BAB IV PENGUJIAN DAN IMPLEMENTASI .....		Error! Bookmark not defined.
4.1	Pembahasan Metode dan Algoritma .....	<b>Error! Bookmark not defined.</b>
4.1.1	Algoritma <i>C4.5</i> .....	<b>Error! Bookmark not defined.</b>
4.1.2	<i>Decision Tree</i> atau Pohon Keputusan .....	<b>Error! Bookmark not defined.</b>
4.2	Spesifikasi <i>Hardware</i> dan <i>Software</i> .....	<b>Error! Bookmark not defined.</b>
4.2.1	Spesifikasi <i>Hardware</i> .....	<b>Error! Bookmark not defined.</b>
4.2.2	Spesifikasi <i>Software</i> .....	<b>Error! Bookmark not defined.</b>
4.3	Tampilan Program .....	<b>Error! Bookmark not defined.</b>
4.3.1	Tampilan Halaman <i>Login</i> .....	<b>Error! Bookmark not defined.</b>
4.3.2	Tampilan Halaman <i>Logout</i> .....	<b>Error! Bookmark not defined.</b>
4.3.3	Tampilan Halaman <i>Regis</i> .....	<b>Error! Bookmark not defined.</b>
4.3.4	Tampilan Halaman <i>Menu</i> .....	<b>Error! Bookmark not defined.</b>
4.3.5	Tampilan Halaman <i>Input</i> .....	<b>Error! Bookmark not defined.</b>
4.3.6	Tampilan Halaman Klasifikasi .....	<b>Error! Bookmark not defined.</b>
4.4	Pengujian Aplikasi.....	<b>Error! Bookmark not defined.</b>
4.4.1	Pengujian <i>Confussion Matrix</i> .....	<b>Error! Bookmark not defined.</b>
4.4.2	Pengujian <i>Black Box</i> .....	<b>Error! Bookmark not defined.</b>
BAB V SIMPULAN DAN SARAN .....		Error! Bookmark not defined.

5.1 Simpulan ..... **Error! Bookmark not defined.**  
5.2 Saran ..... **Error! Bookmark not defined.**

**DAFTAR PUSTAKA**

**LAMPIRAN**

**RIWAYAT HIDUP**



## DAFTAR GAMBAR

GAMBAR 2. 1 PROSES KDD( <i>KNOWLEDGE DISCOVERY IN DATABASE</i> ) (MAIMON & ROKACH, 2014, P. 5) .....	26
GAMBAR 2. 2 <i>DECISION TREE</i> ATAU POHON KEPUTUSAN (MAIMON & ROKACH, 2014, P. 14) .....	28
GAMBAR 2. 3 KERANGKA PEMIKIRAN .....	58
GAMBAR 3. 1 PRODUK .....	60
GAMBAR 3. 2 PRODUK .....	60
GAMBAR 3. 3 STRUKTUR ORGANISASI .....	62
GAMBAR 3. 4 HASIL POHON KEPUTUSAN ATAU <i>DECISION TREE</i> .....	73
GAMBAR 3. 5 KONTRUKSI ALGORITMA DAN METODE .....	75
GAMBAR 3. 6 PROSES PENGOLAHAN RAPIDMINER.....	75
GAMBAR 3. 7 IMPLEMENTASI ALGORITMA <i>C4.5</i> ATAU <i>DECISION TREE</i> DI RAPIDMINER.....	76
GAMBAR 3. 8 <i>USE CASE</i> .....	77
GAMBAR 3. 9 PERANCANGAN HALAMAN <i>LOGIN</i> .....	77
GAMBAR 3. 10 PERANCANGAN HALAMAN <i>LOGIN</i> GAGAL.....	78
GAMBAR 3. 11 PERANCANGAN HALAMAN <i>LOGIN</i> TIDAK DITEMUKAN .....	78
GAMBAR 3. 12 PERANCANGAN HALAMAN <i>REGISTER</i> .....	79
GAMBAR 3. 13 PERANCANGAN HALAMAN <i>REGISTER</i> BERHASIL .....	80
GAMBAR 3. 14 PERANCANGAN HALAMAN <i>LOG OUT</i> .....	80
GAMBAR 3. 15 PERANCANGAN HALAMAN <i>MENU</i> .....	81
GAMBAR 3. 16 PERANCANGAN HALAMAN <i>INPUT DATA</i> .....	81
GAMBAR 3. 17 PERANCANGAN HALAMAN <i>TEST</i> KLASIFIKASI .....	82
GAMBAR 4. 1 PENERAPAN ALGORITMA <i>C4.5</i> PERTAMA . <b>ERROR! BOOKMARK NOT DEFINED.</b>	
GAMBAR 4. 2 PENERAPAN ALGORITMA <i>C4.5</i> KEDUA ..... <b>ERROR! BOOKMARK NOT DEFINED.</b>	
GAMBAR 4. 3 <i>DECISION TREE</i> ATAU POHON KEPUTUSAN RAPIDMINER ..... <b>ERROR! BOOKMARK NOT DEFINED.</b>	
GAMBAR 4. 4 TAMPILAN HALAMAN <i>LOGIN</i> ..... <b>ERROR! BOOKMARK NOT DEFINED.</b>	
GAMBAR 4. 5 TAMPILAN HALAMAN <i>LOGIN</i> GAGAL ..... <b>ERROR! BOOKMARK NOT DEFINED.</b>	
GAMBAR 4. 6 TAMPILAN HALAMAN <i>LOGIN</i> TIDAK DITEMUKAN..... <b>ERROR! BOOKMARK NOT DEFINED.</b>	



GAMBAR 4. 7 TAMPILAN HALAMAN <i>LOGOUT</i> .....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
GAMBAR 4. 8 TAMPILAN HALAMAN <i>REGIS</i> .....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
GAMBAR 4. 9 TAMPILAN HALAMAN <i>MENU</i> .....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
GAMBAR 4. 10 TAMPILAN HALAMAN <i>INPUT</i> .....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
GAMBAR 4. 11 TAMPILAN HALAMAN KLASIFIKASI	<b>ERROR! BOOKMARK NOT DEFINED.</b>
GAMBAR 4. 12 HASIL PENGUJIAN <i>CONFUSION MATRIX</i> .	<b>ERROR! BOOKMARK NOT DEFINED.</b>

### DAFTAR TABLE

TABEL 2. 1 TABEL SIMBOL <i>USE CASE DIAGRAM</i> .....	40
TABEL 2. 2 TABEL SIMBOL <i>ACTIVITY DIAGRAM</i> .....	43
TABEL 2. 3 TABEL SIMBOL <i>SEQUENCE DIAGRAM</i> .....	45
TABEL 2. 4 TABEL SIMBOL <i>CLASS DIAGRAM</i> .....	46
TABEL 2. 5 TABEL SIMBOL <i>STATEMACHINE DIAGRAM</i> .....	48
TABEL 2. 6 PENELITIAN PERTAMA.....	49
TABEL 2. 7 PENELITIAN KEDUA.....	51
TABEL 2. 8 PENELITIAN KETIGA.....	53
TABEL 2. 9 RANGKUMAN PENELITIAN .....	55
TABEL 3. 1 <i>TABLE DATASET</i> .....	64
TABEL 3. 2 <i>ENTROPY</i> DAN <i>GAIN</i> PERTAMA.....	69
TABEL 3. 3 <i>ENTROPY</i> DAN <i>GAIN</i> KEDUA.....	72
TABEL 3. 4 PERANCANGAN TABEL <i>REG</i> .....	82
TABEL 3. 5 PERANCANGAN TABEL <i>INPUT</i> .....	83

## **DAFTAR LAMPIRAN**

LAMPIRAN 1 SURAT IJIN RISET DATA

LAMPIRAN 2 SURAT BIMBINGAN

LAMPIRAN 3 SCRIPT WAWANCARA

LAMPIRAN 4 *REQUIREMENT ELECITATION*



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Wortel atau (*Daucus Carota L*) merupakan salah satu sayuran yang menyimpan karbohidrat yang sangat besar didalam nya untuk sayuran tersebut agar dapat berbunga di tahun kedua pada proses penumbuhan sayuran tersebut. Wortel juga merupakan sayuran yang paling sering dikonsumsi oleh masyarakat dengan cara di rebus, di buat untuk minuman jus, serta di goreng dicampur dengan sayuran lainnya. Wortel Berastagi adalah jenis wortel yang biasanya dihasilkan disebuah dataran tinggi yang dimana wortel ini kaya akan *vitamin A* yang bermanfaat untuk menyehatkan kedua mata sehingga masyarakat banyak mengkonsumsi wortel berastagi ini untuk mencegah dari berbagai macam penyakit mata yang akan diderita.

Konsumen dari wortel berastagi ini berasal dari semua golongan masyarakat yang menjadikan wortel berastagi ini memiliki jangkauan pemasaran yang sangat luas dibandingkan dengan jenis wortel lainnya. Perkembangan zaman sekarang, tidak mempengaruhi jumlah konsumsi dimasyarakat. Malah dengan perkembangan zaman yang terjadi ini , konsumsi dimasyarakat menjadi sangat meningkat karena munculnya olahan- olahan makanan dan minuman baru yang menggunakan wortel didalam nya. Tentunya untuk menghasilkan makanan dan minuman yang baik, diperlukan kualitas wortel yang terbaik.

Produsen serta konsumen terkadang dapat menyimpan wortel bisa sampai satu minggu pada ruangan tertutup bersuhu dingin ataupun tidak bersuhu dingin. Hal ini menyebabkan kualitas wortel yang akan dikonsumsi atau diolah ini mengalami penurunan kualitas. Semakin lama waktu wortel ini disimpan akan mengakibatkan

perubahan seperti penurunan berat wortel, timbul bercak–bercak hitam pada kulit wortel serta kulit wortel menjadi licin, pucat dan menjadi basah. Maka dari itu, pengolahan atau penyortiran ulang wortel harus dilakukan agar dapat mengetahui mana kualitas wortel yang terbaik, sehingga ketika mendapatkan kualitas wortel yang terbaik tersebut dengan mudah dapat ditentukan harga jual wortel berdasarkan tingkat kualitas yang dimilikinya.

Dalam mengatasi permasalahan tersebut, akan menggunakan data mining yang dimana data mining adalah sebuah proses penambangan atau penggalian yang terkandung didalam sebuah ilmu komputer yang biasanya proses tersebut digunakan untuk melakukan sebuah proses dalam yang bertujuan untuk mencari sebuah informasi atau knowledge yang dimana didalam nya terkandung sebuah tahap–tahap yang berguna mencari sebuah pola tertentu yang ada didalam sebuah data yang tersimpan didalam database, data mining ini banyak digunakan didalam bidang seperti statistika, mesin pencarian dan *artificial intelligence*. (Xu et al., 2013, p. 4)” Data mining adalah sebuah proses analisis dan analitik untuk mengungkapkan sebuah pola maupun sebuah tren yang tersembunyi serta tercantum sekumpulan data yang banyak serta luas, yang dimana telah melalui paradigma kecerdasan komputasi yang sudah mutakhir”.

Data Mining ini memiliki banyak fungsi seperti mengetahui sebuah trend yang sedang berlangsung maupun memiliki fungsi untuk memprediksi keputusan yang akan diambil dimasa depan, didalam data mining ini terdapat beberapa metode yang dimana metode– metode yang digunakan tersebut akan disesuaikan dengan tujuan yang ingin dicari. Salah satunya adalah metode klasifikasi atau *classification* yang dimana metode ini adalah salah satu metode yang paling umum digunakan oleh banyak masyarakat yang biasanya digunakan untuk mencari serta menemukan sebuah model–model atau

fungsi yang didalam nya menjelaskan tentang ciri-ciri dari sebuah kelas maupun konsep data yang ada.

Didalam data mining juga terdapat algoritma yang dimana algoritma itu sendiri adalah sebuah alur pemikiran untuk mencari pemecahan masalah yang dimana didalam nya disusun secara logis dan sistematis untuk menyelesaikan sebuah masalah yang ada. Menurut (Sismoro & Amikom, n.d., p. 34) "Algoritma adalah pemikiran yang berisi langkah-langkah dalam menyelesaikan sebuah masalah, yang dimana langkah-langkah tersebut merupakan urutan aksi seperti kejadian, tindakan serta pengulangan dari aksi tersebut".

Salah satu algoritma yang terkandung didalam metode klasifikasi ini adalah algoritma C4.5. Algoritma C4.5 adalah sebuah algoritma yang biasanya digunakan untuk membuat sebuah *Decision Tree* atau yang biasanya akrab disebut dengan pohon keputusan yang dimana proses dari pohon keputusan ini adalah mengubah bentuk data yang ada di sebuah table menjadi sebuah model data yang berbentuk pohon. Berdasarkan latar belakang yang penulis buat diatas, maka penulis bertujuan membuat tugas akhir dengan judul : **"IMPLEMENTASI DATA MINING UNTUK MENENTUKAN KUALITAS WORTEL DI PT PADMA MEKAR LESTARI MENGGUNAKAN METODE C4.5"**.

## 1.2 Identifikasi Masalah

Berdasarkan penjelasan dari latar belakang masalah diatas, identifikasi masalah adalah sebagai berikut:

- a. Penentuan kualitas wortel yang baik yang dimana proses penentuannya tidak memiliki tolak ukur karena yang dilakukan sekarang masih kurang baik.

- b. Sering terjadi miss dalam bagian quality control dikarenakan penentuan kualitas wortel yang masih kurang baik.

### 1.3 Rumusan Masalah

Berdasarkan penjelasan dari latar belakang masalah diatas, maka perumusan masalah adalah sebagai berikut:

- a. Apakah implementasi data mining menggunakan metode C4.5 bisa membantu PT Padma Mekar Lestari dalam menentukan kualitas wortel yang baik ?
- b. Bagaimana cara menentukan kualitas wortel di PT Padma Mekar Lestari?
- c. Bagaimana cara klasifikasi kualitas wortel di PT Padma Mekar Lestari?

### 1.4 Ruang Lingkup

Adapun ruang lingkup dari penelitian ini adalah sebagai berikut:

1. Penelitian ini dilakukan di PT Padma Mekar Lestari.
2. Penelitian ini dilakukan untuk menentukan kualitas wortel di PT Padma Mekar Lestari.
3. Data set yang digunakan adalah data set yang didapat dari PT Padma Mekar Lestari.

### 1.5 Tujuan dan Manfaati Penelitian

Tujuan dan manfaat dari penelitian yang diadakan ini adalah untuk memperoleh pengetahuan yang baru, sebagaimana untuk membuktikan tentang kebenaran suatu penelitian yang ada. Adapun tujuan dan manfaat dari penelitian ini adalah sebagai berikut:

### 1.5.1 Tujuan

- a. Melakukan implementasi data mining menggunakan metode C4.5 untuk melakukan penentuan kualitas wortel di PT Padma Mekar Lestari agar dapat mempermudah dalam melakukan proses penentuan kualitas.

### 1.5.2 Manfaat

Bagi penulis:

- a. Sebagai bentuk implementasi atau penerapan atas ilmu – ilmu yang sudah penulis peroleh selama melakukan masa perkuliahan.
- b. Memberikan manfaat untuk penulis agar dapat menyelesaikan tugas akhir.

Bagi PT Padma Mekar Lestari:

- a. Untuk dapat mengetahui dengan implementasi data mining menggunakan metode C4.5 ini apakah dapat membantu PT Padma Mekar Lestari dalam mengatasi klasifikasi kualitas wortel.
- b. Meningkatkan proses efektivitas serta efisiensi dalam klasifikasi kualitas wortel di PT Padma Mekar Lestari.

Bagi pembaca:

- a. Dengan adanya penelitian ini semoga dapat dijadikan sebagai referensi untuk melakukan pengembangan dari penelitian ini sehingga mendapatkan hasil yang lebih maksimal dari yang sebelumnya.
- b. Serta semoga dapat dijadikan sebagai perbandingan dalam melakukan penelitian lainnya.

## 1.6 Metode Penelitian

### 1.6.1 Teknik Pengumpulan Data

#### a. Studi Pustaka

Teknik pengumpulan data ini dilakukan untuk mengumpulkan semua data yang terkait dengan penelitian agar dapat mendukung penelitian ini baik dari jurnal ilmiah, artikel, buku, internet dan lainnya.

#### b. Wawancara

Teknik pengumpulan data ini dilakukan untuk mengumpulkan semua data dengan cara melakukan sebuah wawancara kepada pemilik perusahaan serta tim yang bekerja dalam menentukan kualitas wortel di PT Padma Mekar Lestari.

#### c. Observasi

Teknik pengumpulan data ini dilakukan dengan cara mengamati secara langsung proses yang berjalan, untuk dapat memahami aktivitas serta kondisi yang terjadi pada saat itu.

## 1.7 Sistematika Penulisan

Adapun sistematika penulisan dari penelitian ini adalah sebagai berikut :

### a. BAB I PENDAHULUAN

Bab ini akan menjelaskan tentang latar belakang masalah, identifikasi masalah, rumusan masalah, ruang lingkup, tujuan dan manfaat penelitian, metode penelitian dan juga sistematika penulisan.



## **b. BAB II LANDASAN TEORI**

Bab ini akan menjelaskan tentang gambaran dari teori–teori yang digunakan untuk dilakukannya penelitian ini. Yang dimana teori–teori ini berisi teori umum, teori khusus, teori perancangan dan tinjauan studi.

## **c. BAB III ANALISA MASALAH & PERANCANGAN APLIKASI**

Bab ini akan menjelaskan tentang tinjauan umum perusahaan, kebutuhan apa yang akan dibutuhkan oleh perusahaan, pemecahan masalah, perancangan dari UML dan juga perancangan aplikasi.

## **d. BAB IV PENGUJIAN DAN IMPLEMENTASI**

Bab ini akan menjelaskan tentang pengujian dan juga implementasi dari *system* yang dibuat didalam penelitian yang dimana apakah sudah sesuai dengan kebutuhan dari perusahaan atau tidak.

## **e. BAB V KESIMPULAN DAN SARAN**

Bab ini akan menjelaskan tentang kesimpulan dan juga saran dari penelitian ini agar dapat dijadikan sebagai referensi untuk penelitian lainnya.

## **f. DAFTAR PUSTAKA**

Didalam ini akan berisi tentang daftar pustaka yang berhubungan dengan kutipan yang dicantumkan di penulisan penelitian ini.

## **g. LAMPIRAN**

Didalam ini akan berisi tentang lempiran-lampiran yang berhubungan dengan penelitian yang dilakukan.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Teori Umum**

##### **2.1.1 Database**

*Database* atau yang biasa disebut sebagai basis data adalah sebuah kumpulan dari informasi–informasi yang disimpan didalam sebuah komputer secara sistematis yang dimana ketika data tersebut dikelola akan menggunakan sebuah ketentuan yang tertentu untuk mendapatkan informasi yang dicari, membuang informasi maupun menyimpan informasi tersebut.

Menurut (Asrianda, 2012, p. 1) ”Database adalah sekumpulan dari table-table yang saling berhubungan, yang dimana hubungan pada table-table yang tersebut dapat dilihat disetiap masing-masing tablenya”. Biasanya database digunakan untuk mempermudah dalam proses mengidentifikasi sebuah data yang diminta untuk mendapatkan sebuah informasi yang cepat dan juga akurat.

Database memiliki beberapa jenis, yaitu:

##### **2.1.1.1 Database Warehouse**

*Database Warehouse* adalah sebuah sistem basis data yang dimana, didalam nya itu menggabungkan data dari berbagai macam sumber menjadi satu di tempat penyimpanan yang terpusat untuk mendukung dilakukannya sebuah data mining, analisis data, machine learning dsb.

Menurut (Jollyta et al., 2020, p. 5) ”*Database Warehouse* merupakan sekumpulan data dari sebuah organisasi yang telah diolah

sehingga telah menghasilkan sebuah informasi yang dibutuhkan untuk pengambilan sebuah keputusan kedepannya”. *Database Warehouse* mengatur, menyimpan dan juga mengelola data historis dalam jumlah yang sangat besar, yang dimana *Database Warehouse* ini menyimpan dan juga mengelola data menurut tingkat kedetailan data atau biasa disebut *granularity* yang berbeda-beda.

#### **2.1.1.2 *Distributed Database***

*Distributed Database* adalah sebuah basis data yang dimana media penyimpanannya itu tidak terpasang secara langsung pada sebuah komputer atau laptop yang ada. *Distributed Database* ini disimpan di beberapa tempat yang tersebar lewat jaringan maupun di beberapa tempat yang sama. Yang dimana basis data disini akan terdistribusi melewati suatu situs atau media yang sudah tergabung tetapi situs atau media tersebut tidak mempunyai komponen secara bentuk fisik.

#### **2.1.1.3 *Relational Database***

*Relational Database* atau database relasional adalah sebuah kumpulan item maupun variable sebuah data yang dimana hubungannya sudah ditentukan diawal. Contohnya seperti table-table yang berisi sebuah informasi. yang dimana *relational database* disini akan memberikan sebuah skema gambaran tentang hubungan antara table-table tersebut didalam sebuah database.

Menurut (Dr. Henderi, 2020, p. 55) "Database relasional terdiri dari sebuah kumpulan table-tabel, yang dimana table-table tersebut diberikan sebuah nama yang unik yang dimana didalam nya menyimpan informasi tentang instruktur".

### **2.1.2 Sistem**

Sistem adalah sebuah sesuatu atau kesatuan yang terdiri dari sebuah struktur dan juga proses yang berisi sebuah prinsip-prinsip serta aturan-aturan yang ada yang biasanya digunakan untuk mengerjakan sesuatu. Berasal dari bahasa Yunani *sistema* dan bahasa Latin *systema* yang mempunyai arti bagian himpunan atau komponen yang saling terhubung teratur secara keseluruhan.

Menurut (Sumampouw, 2019, p. 124) "Sistem adalah sebuah tatanan yang berisi gambaran dari adanya sebuah rangkaian berbagai macam komponen-komponen yang saling memiliki hubungan serta tujuan bersama secara serasi, terkoordinasi yang cara kerja atau berjalannya itu berdasarkan jangka waktu yang sudah tertentu dan direncanakan sejak awal".

### **2.1.3 Internet**

Internet adalah suatu jaringan pada sebuah komputer yang sudah terhubung serta memiliki kemampuan untuk menangkap dan juga dapat membaca sebuah *ip* (*internet protocol*) dan juga *tcp* (*transmission control protocol*) yang dimana memiliki manfaat seperti menjadi sebuah sarana untuk melakukan sebuah komunikasi jarak jauh dan juga cepat serta untuk mencari sebuah informasi yang tersebar.

Pada tahun 1969, sebuah badan pertahanan yang bertempat di *US(United States)* atau Amerika Serikat membiayai sebuah proyek yang bernama ARPANET, yang dimana ARPANET itu sendiri adalah sebuah proyek yang dimana proyek tersebut memungkinkan sebuah sistem untuk saling terhubung satu dengan yang lainnya. Proyek ini dirancang menggunakan sebuah jaringan, informasi yang dapat dipindahkan serta standar lainnya yang menjadi awal terciptanya *tcp(transmission control protocol)*.

ARPANET tersebut pertama kali dikenalkan kepada umum pada tahun 1972, yang dimana disini ARPANET menggunakan sebuah basis teknologi yang bernama ALOHAnet. Kemudian pada tahun 1981, muncul sebuah proyek baru yang dimana didalam proyek ini ada sebanyak 213 jaringan yang saling terhubung yang membuat ARPANET bingung bagaimana cara untuk mengaturnya sedemikian rupa.

Kemudian ARPANET membuat sebuah informasi untuk memecah menjadi 2 bagian, yang dimana menjadi MILNET yang lebih bergerak untuk keperluan dalam bidang militer di Amerika Serikat dan ARPANET yang lebih bergerak untuk keperluan non militer seperti universitas. Dari sini lah, ARPANET dan MILNET dikenal dengan nama DARPA Internet atau yang sekarang disebut sebagai internet.

Menurut (Media, 2016, p. 6) "Internet adalah sebuah jaringan atau sebuah sistem yang terdapat pada sebuah jaringan komputer yang sudah saling terhubung dengan menggunakan sebuah *system global transmission control / internet protocol suite (tcp/ip)* sebagai protokol utama dalam

melakukan sebuah pertukaran *packet switching communication protocol* untuk melayani semua pengguna internet tersebut dimanapun berada.

Internet memiliki sebuah fasilitas yang dapat digunakan, diantaranya adalah:

1. Browsing

Browsing adalah sebuah fasilitas yang dimana, fasilitas ini memungkinkan pengguna internet untuk melakukan sebuah kegiatan berselancar didalam internet seperti mencari sebuah berita terbaru maupun mencari sebuah informasi untuk membantu sebuah penelitian.

2. Chatting

Chatting adalah sebuah fasilitas yang dimana, fasilitas ini memungkinkan pengguna untuk melakukan sebuah percakapan menggunakan media yang ada seperti text, suara maupun video yang tidak memiliki batasan jarak untuk melakukan sebuah percakapan.

3. E-mail

E-mail adalah sebuah fasilitas yang dimana, fasilitas ini memungkinkan pengguna untuk melakukan pengiriman surat dengan orang lain.

4. Download

Download adalah sebuah fasilitas yang dimana, fasilitas ini memungkinkan pengguna untuk melakukan sebuah proses pengambilan file dalam bentuk gambar, suara dan video dari komputer pengguna.

## 5. Upload

Upload adalah sebuah fasilitas yang dimana, fasilitas ini memungkinkan pengguna untuk melakukan sebuah proses peletakan file yang dimana file tersebut setelah diletakan dapat di download oleh orang lain.

### 2.1.4 Data

Data adalah sebuah kumpulan dari berbagai macam informasi yang ada atau juga sekumpulan macam keterangan dari banyal hal yang diperoleh dengan cara melakukan pencarian ke sumber yang ada. Lebih sederhana nya, data adalah sebuah kumpulan informasi yang memberikan gambaran tentang suatu keadaan.

Menurut (Indah Werdiningsih et al., 2020, p. 1) "Data adalah sebuah kumpulan dari objek dan atribut yang telah menjadi satu, yang dimana atribut tersebut menunjukkan setiap buah karakteristik nya tersendiri dan juga sebuah objek yang dapat digambarkan secara jelas".

#### 2.1.4.1 Fungsi Data

Berdasarkan fungsi dari sebuah data, diantaranya adalah:

1. Dapat dijadikan sebagai sebuah acuan untuk melakukan sebuah kegiatan.
2. Dapat dijadikan sebagai sebuah perencanaan dalam melakukan sebuah hal.
3. Dapat dijadikan sebagai sebuah dasar dalam membuat sesuatu dan pengambilan keputusan.

4. Dapat dijadikan sebagai bahan untuk melakukan evaluasi yang bertujuan untuk meningkatkan sebuah kualitas.

#### **2.1.4.2 Jenis Data**

Data memiliki 3 buah jenis, ketiganya di kelompokkan berdasarkan kelompok nya masing–masing.

##### **2.1.4.2.1 Berdasarkan Cara Mendapatkannya**

- Data Premier

Data Primer adalah sebuah data yang di dapatkan melalui proses pengumpulan secara langsung dari sumber yang sebelumnya sudah dilakukan penelitian didalam sebuah organisasi , maupun penelitian perorangan.

- Data Sekunder

Data Sekunder adalah sebuah data yang di dapatkan dari sumber yang sudah ada

##### **2.1.4.2.2 Berdasarkan Sumbernya**

- Data Eksternal

Data Eksternal adalah sebuah data hasil penelitian yang didapatkan dari luar tempat atau diluar organisasi penelitian itu dilakukan.

- Data Internal

Data Internal adalah sebuah data hasil penelitian yang didapatkan langsung dari organisasi tempat penelitian itu dilakukan.



### 2.1.4.2.3 Berdasarkan Sifatnya

- Data Kuantitatif

Data Kuantitatif adalah sebuah data yang dihasilkan dengan melakukan sebuah survey yang dimana hasil survey tersebut berisi jawaban berbentuk sebuah angka.

- Data Kualitatif

Data Kualitatif adalah sebuah data yang dihasilkan dengan cara seperti melakukan kuisioner, wawancara , overservasi dan juga yang lainnya yang dimana data ini berbentuk seperti gambar, simbol maupun berbentuk verbal lainnya.

### 2.1.5 Komputer

Komputer berasal dari bahasa latin dari "*computare*" yang memiliki sebuah arti menghitung. Ditemukan pertama kali pada tahun 1822 oleh seorang ahli yang bernama Charles Babbage yang berasal dari negara inggris yang ingin menciptakan sebuah mesin hitung menggunakan uap yang ia berinama "*Difference Engine 0*" yang dapat menghitung sebuah angka.

Adalah sebuah perangkat elektronik yang dapat melakukan sebuah perhitungan dengan cepat seperti perhitungan aritmatika yang dimana selain dapat melakukan sebuah perhitungan, komputer juga dapat menerima data lalu serta memprosesnya yang menghasilkan sebuah informasi untuk penggunanya.

Menurut (Fachri et al., 2020, p. 13) "Komputer adalah mesin hitung elektronik yang dapat menerima dengan cepat sebuah informasi digital dan dapat mengolah informasi tersebut menurut instruksi yang sudah tersimpan didalam komputer tersebut yang dapat menghasilkan informasi setelah diolah".

Generasi pertama untuk kebutuhan umum dibangun pada tahun 1946 oleh John Mauchly dan partner nya John Presper Eckert di kampus Universitas Pennsylvania dengan nama ENIAC (Electronic Numerical Integrator and Computer) yang berfungsi untuk membantu kebutuhan umum pertama kali di dunia.

Menurut (fauzi, 2018, p. 1) "Komputer adalah sebuah alat elektronik yang dapat melakukan sebuah proses input, proses pengolahan dan juga memberikan sebuah informasi dengan menggunakan aplikasi maupun program yang tersimpan didalam nya dan dapat menyimpan aplikasi maupun program hasil pengolahan secara otomatis tersebut".

#### **2.1.6 Implementasi**

Implementasi adalah sebuah kegiatan yang dilakukan untuk melakukan perencanaan yang mengacu kepada suatu aturan tertentu untuk mencapai tujuan dari dilaksanakan nya perencanaan tersebut dengan suatu rencana maupun konsep yang akan diterapkan di berbagai macam bidang tanpa ada batasan apapun didalam nya.

Seseorang melakukan implementasi biasanya untuk mencapai sebuah tujuan-tujuan yang diinginkanya seperti menciptakan sebuah rancangan

*system* sambil menganalisa serta melakukan penelitian terhadap rancangan *system* tersebut, menyempurnakan sebuah *system* yang sudah direncanakan akan dibuat dan sampai untuk memprediksi berapa banyak kebutuhan yang dibutuhkan terhadap *system* yang ada.

### 2.1.7 Aplikasi

Aplikasi adalah sebuah perangkat lunak atau *software* yang didalamnya berisi bermacam-macam fitur tertentu yang bisa diakses oleh semua pengguna yang memilikinya. Seperti yang tersebar didalam *App Store* maupun *Play Store* yang terdapat didalam smartphone pengguna. Aplikasi ini telah menjadi sumber penghasilan beberapa *industry* yang ada, seperti *industry game* yang telah meraup keuntungan sangat banyak dari pengguna yang memainkannya.

Menurut (fauzi, 2018, p. 12) "Perangkat lunak aplikasi adalah sebuah bagian dari perangkat lunak yang ada pada sebuah komputer yang dimana telah dibuat menggunakan sebuah program yang bertujuan untuk dapat digunakan untuk menjalankan suatu tugas yang telah diinginkan oleh pengguna atau *user* yang telah membuatnya".

Tidak hanya *industry* yang bergerak di bidang game, *industry* lainnya yang bergerak dibidang social media juga memberikan keuntungan bagi pengguna nya. Seperti Instagram, yang dimana sudah banyak user aplikasi instagarm tersebut bisa mendapatkan penghasilan dari menggiati serta menjadi terkenal pada aplikasi tersebut.

Dalam kehidupan sehari-hari, aplikasi ini memiliki berbagai macam fungsi yang bermanfaat untuk kehidupan seperti menjadi sebuah sarana ilmu pengetahuan dalam melakukan sebuah tugas maupun penelitian, menjadi sebuah sarana pendidikan seperti penggunaan aplikasi seperti zoom dan google meet yang dimana membantu dalam proses pembelajaran jarak jauh serta menjadi sarana bisnis seperti penggunaan aplikasi seperti Tokopedia dan shopee yang dimana pengguna atau *user* dapat menjual barang pada aplikasi tersebut yang dapat menghasilkan sebuah keuntungan.

Berikut adalah jenis-jenis aplikasi, antara lain adalah sebagai berikut:

1. Aplikasi Pengolah Angka

Aplikasi ini biasanya digunakan untuk mengolah atau memproses serta dapat digunakan untuk menghitung sebuah angka, biasanya aplikasi yang digunakan disini seperti Google Sheet, Microsoft Excel serta aplikasi kalkulator.

2. Aplikasi Pengolah Kata

Aplikasi ini biasanya digunakan untuk mengolah sebuah kata, aplikasi yang biasanya digunakan disini seperti Notepad dan Microsoft Word.

3. Aplikasi Pengolah Grafis

Aplikasi ini biasanya digunakan untuk mengolah sebuah grafik atau gambar yang ada, biasanya aplikasi pengolah grafis yang digunakan seperti Corel Draw dan Adobe Photoshop.

4. Aplikasi Pengolah Data

Aplikasi ini biasanya digunakan untuk mengolah atau memproses sebuah data yang dimana hasilnya akan memberikan

sebuah informasi, biasanya aplikasi pengolah data yang digunakan seperti Rapidminer dan Microsoft Access.

## 2.2 Teori Khusus

### 2.2.1 Klasifikasi

Menurut (Wanto, 2020, p. 103) "Klasifikasi atau *classification* adalah sebuah proses pembelajaran yang dimana mempunyai satu buah target yang dimana target tersebut memetakan setiap himpunan ke sebuah label kelas yang sudah dilakukan proses indentifikasi". Fungsi dari sebuah klasifikasi atau *classification* ini dapat dilihat dari ditemukannya sebuah model atau fungsi yang menggambarkan sebuah class dari suatu data yang ada.

Klasifikasi adalah sebuah proses untuk mencari sebuah model dan fungsi dari sebuah data yang dimana tujuan dari mencari model atau fungsi itu sendiri adalah untuk memisahkan sebuah class dari suatu data, lalu menjelaskan class dari suatu data itu tersebut agar dapat memprediksi sebuah class data baru yang belum pernah dimiliki. Biasanya hasil dari klasifikasi ini ditampilkan ke dalam sebuah bentuk pohon keputusan atau nama lainnya adalah *decision tree* maupun dalam bentuk lainnya seperti *neural networks*.

Contoh dari sebuah klasifikasi adalah sebuah task machine learning yang mempunyai tipe supervised learning untuk melakukan sebuah prediksi yang ingin dicari seperti melakukan penentuan email spam atau ada juga penentuan seseorang menghidap penyakit diabetes serta bisa digunakan untuk penentuan kemenangan dari sebuah lomba olahraga. Di klasifikasi ini, yang akan diprediksi adalah sebuah class yang dapat dibagi menjadi sebuah level.

Sampai saat ini pengembangan klasifikasi tetap berlanjut untuk menyelesaikan sebuah masalah seperti 1 class, 2 class dan maupun lebih dari 2 class, untuk kasus dalam menyelesaikan kasus lebih dari 2 class biasanya digunakan sebuah metode yang dapat menyelesaikan permasalahan lebih dari 2 class tersebut.

Menurut (Id, 2021, p. 7) "Klasifikasi adalah sebuah pengelompokan sebuah data yang mempunyai sebuah *class* label, yang dimana komputer akan melakukan penentuan atau memilih sebuah class dari data tersebut yang sudah ditentukan sejak awal.

### **2.2.2 Algoritma**

Algoritma adalah sebuah metode atau proses yang berisi serangkaian aturan yang harus dilakukan dalam melakukan sebuah perhitungan dalam sebuah pemecahan masalah yang ditujukan ke komputer, yang dimana susunan logis akan diurutkan berdasarkan sistematika yang sudah ditentukan serta digunakan untuk memecahkan sebuah masalah. Umumnya intruksi yang digunakan seperti "*If*" dan "*Else*" untuk memecahkan sebuah masalah yang dimiliki.

#### **2.2.2.1 Fungsi Algoritma**

Beberapa fungsi dari sebuah algoritma, diantaranya adalah:

1. Dapat digunakan untuk mengatasi sebuah kesalahan eror pada saat melakukan pemograman.
2. Dapat digunakan untuk mencari sebuah masalah terhadap program yang sudah dibuat.

3. Dapat membantu pada saat melakukan sebuah penulisan bahasa pemrograman yang dimana, tidak perlu mengulang-ulang dalam melakukan penulisan tersebut.
4. Dapat membantu pada saat pembuatan program agar lebih logis dan juga terstruktur rapih.

#### **2.2.2.2 Karakteristik Algoritma**

Beberapa karakteristik dari sebuah algoritma, diantaranya adalah:

1. Sebuah algoritma jelas dan tidak bersifat ambigu, dalam artian intruksi yang digunakan harus bisa langsung dipahami.
2. Sebuah algoritma memiliki input maupun output yang jelas, dalam artian hasil input dan output yang ada bisa ditentukan dengan jelas.
3. Sebuah algoritma tidak bergantung dengan satu bahasa pemrograman saja, dalam artian algoritma dapat memahami bahasa pemrograman apapun.
4. Sebuah algoritma terbatas, dalam artian secepat mungkin algoritma akan menghasilkan sebuah output yang diinginkan.

#### **2.2.3 Data Mining**

Data Mining adalah sebuah proses yang dimana untuk menemukan pola pada serta korelasi sebuah kumpulan data yang sangat banyak dan juga besar yang melibatkan sebuah metode statistik, system database untuk mencari sebuah hasil yang diinginkan. Biasanya, sebuah proses data mining ini pada

umumnya sering sekali menggunakan metode matematika, statistika hingga menggunakan *artificial intelligence*.

Menurut (Indah Werdiningsih et al., 2020, p. 17) "Data Mining adalah sebuah bidang ilmu yang digunakan untuk menyelesaikan sebuah masalah dalam pengambilan informasi dari sebuah database yang berukuran besar dengan menggabungkan teknik dari statistic, mesin pembelajaran, visualisasi data, pengenalan pola dan database. Tujuan dari data mining ini adalah untuk mengekstrak sebuah informasi dengan metode dari kumpulan data kemudian mengubah informasi tersebut menjadi lebih terstruktur yang dapat dipahami".

Menurut (Fitri Marisa et al., 2021, p. 6) "Data mining merupakan sebuah alat yang memungkinkan para pengguna untuk mengkases secara cepat sebuah data dalam ukuran yang besar, yang dimana sebuah data mining dapat menggambarkan sebuah pengumpulan teknik yang bertujuan untuk menemukan sebuah pola yang tidak diketahui terhadap data yang ada".

Data mining digunakan karena banyaknya ketersediaan data dalam jumlah yang sangat besar serta kebutuhan yang sangat mendesak untuk menghasilkan sebuah pengetahuan dan juga informasi yang berguna, sehingga dengan menggunakan data mining untuk kebutuhan yang mendesak seperti menghasilkan pengetahuan dan informasi serta mengatasi jumlah data yang sangat besar adalah sebuah hal yang tepat.

Menurut (Yusuf Sulisty Nugroho, 2019, p. 133) "Data mining adalah sebuah serangkaian proses untuk mencari nilai tambah seperti informasi yang tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan



diperoleh dengan cara mengekstrak sebuah data, mengenali pola-pola yang penting serta menarik data yang terdapat didalam basis data”.

Data mining mempunyai beberapa metode yang dapat digunakan, diantaranya adalah:

1. *Classification* / Klasifikasi

Klasifikasi atau *Classification* adalah sebuah metode yang paling umum kita jumpai pada sebuah data mining. Yang dimana metode ini bertujuan untuk menemukan definisi antara kesamaan karakteristik dari sebuah *class* data yang belum pernah diketahui.

2. *Clustering*

*Clustering* atau nama lainnya adalah metode segmentasi, sebuah metode yang biasanya digunakan untuk mengidentifikasi sebuah kelompok dari kasus yang ada berdasarkan pada kelompok atribut yang dimiliki atau juga biasanya digunakan untuk mengelompokkan data yang memiliki atribut yang sama.

3. *Association*

*Association* atau nama lainnya adalah metode market basket analisis, adalah sebuah metode yang berhubungan dengan pemasaran sebuah bisnis yang bertujuan untuk menganalisa transaksi penjualan dengan identifikasi produk-produk yang paling sering dibeli secara bersamaan pada saat orang membelinya.

4. *Regression*

*Regression* adalah sebuah metode yang memiliki kemiripan dengan *classification* atau klasifikasi, yang membedakan kedua metode tersebut adalah metode *regression* tidak bisa mencari pola yang ada sebagai *class*,

metode ini biasanya digunakan untuk mencari sebuah pola tertentu dari sebuah nilai-nilai numerik.

#### 5. *Forecasting*

*Forecasting* adalah sebuah metode yang pada umumnya digunakan untuk memprediksi sebuah nilai yang akan diacapai pada sebuah periode yang sudah di tentukan dengan musim, trend, noise data dan nilai dari periode yang lalu yang akan dijadikan sebagai dasar pada saat melakukan sebuah prediksi.

#### 6. *Sequencing*

*Sequencing* adalah sebuah metode yang pada umumnya digunakan untuk mencari pola tahapan terhadap serangkaian suatu kejadian yang ada, metode ini biasanya menggunakan data yang berasal dari periode yang sudah berjalan yang kemudian digabungkan lalu dilakukan proses analisa untuk mengetahui pola tahapan yang ada.

#### 7. *Descriptive / Deskripsi*

*Descriptive / Deskripsi* adalah sebuah metode yang pada umumnya digunakan untuk mencari sebuah perilaku data, yang dimana metode ini secara khusus mengamati data berdasarkan dari sifat data itu tersendiri bukan berfokus pada sebuah angka.

Pada saat melakukan data mining, terdapat beberapa tahapan yang harus dilakukan diantaranya adalah:

#### 1. *Data Selection / Seleksi Data*

Didalam tahap ini akan dilakukan tahap pemilihan data dari data yang ada, dikarenakan ada beberapa data didalam sebuah database tersebut

yang tidak akan terpakai pada saat melakukan data mining, sehingga pemilihan data ini dilakukan untuk mencari data yang sesuai.

## 2. *Data Cleaning* / Pemilihan Data

Didalam tahap ini akan dilakukan tahap pembersihan data yang mencakup seperti membuang data yang terduplikat, memeriksa data yang tidak konsisten, memperbaiki data seperti kesalahan pada saat proses cetak, serta melakukan proses *enrichment* atau proses memperkaya sebuah data yang ada dengan menggunakan data maupun informasi yang relevan seperti data maupun informasi yang bersifat eksternal.

## 3. Transformasi Data

Didalam tahap ini akan dilakukan tahap perubahan skala pengukuran data asli kedalam bentuk lain, sehingga data yang ada dapat memenuhi asumsi yang akan mendasari dari analisis yang akan digunakan sebagai penelitian. Transformasi data ini biasanya membantu dalam melakukan penelitian untuk membuat data asli memenuhi sebuah analisis.

## 4. Data Mining

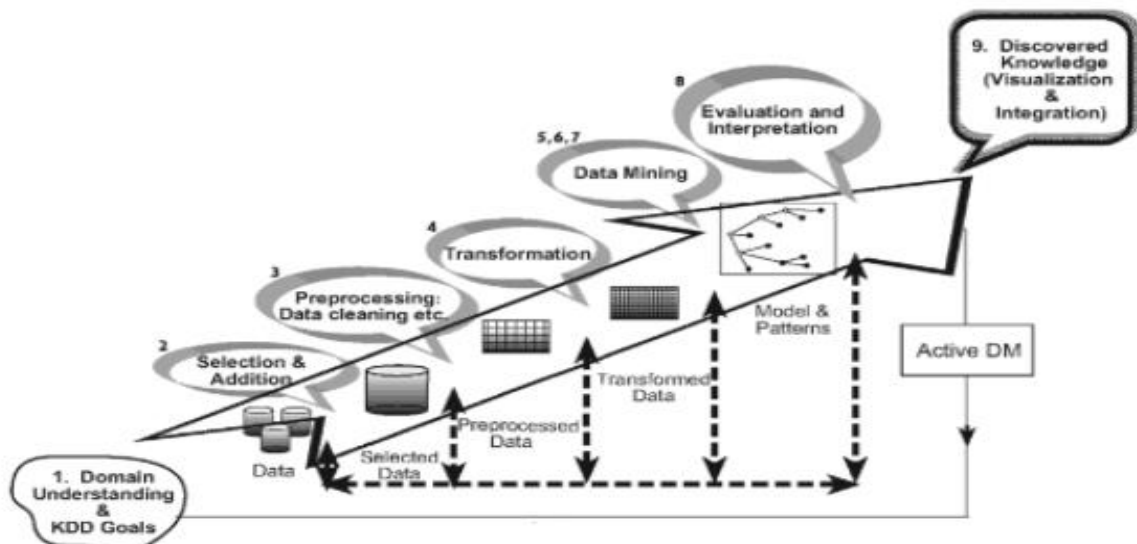
Didalam tahap ini akan dilakukan tahap mining sebuah data guna mencari sebuah pola ataupun informasi menarik dari data tersebut yang akan terpilih dengan metode yang digunakan. Metode yang akan digunakan berdasarkan penelitian yang dilakukan lalu dengan memilih sebuah metode yang tepat.

## 5. Pattern Evaluation

Didalam tahap ini setelah melakukan tahap mining data yang menghasilkan sebuah informasi akan di presentasikan dalam bentuk yang mudah dipahami berdasarkan hasil dari pengukuran yang telah dilakukan.

## 6. Knowledge Representation

Dalam tahap ini adalah tahap terakhir yang dimana pada tahap ini semua informasi berupa pengetahuan yang sudah didapatkan akan divisualisasikan kepada seluruh user.



Gambar 2. 1 Proses KDD(Knowledge Discovery in Database) (Maimon & Rokach, 2014,

p. 5)

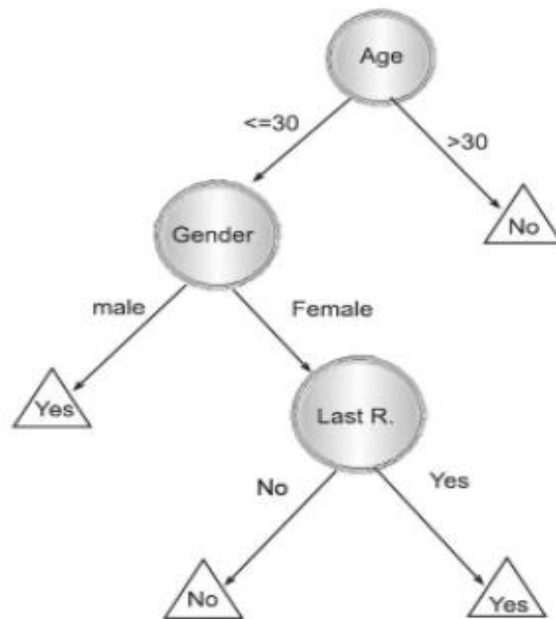
### 2.2.4 Decision Tree atau Pohon Keputusan

Decision tree atau Pohon keputusan adalah sebuah metode yang ada didalam algoritma klasifikasi yang dimana konsepnya yang berbentuk sebuah struktur pada pohon. Node yang paling pertama dari sebuah decision tree atau pohon keputusan biasanya disebut dengan root yang dimana setiap node yang

ada itu akan mempresentasikan sebuah attribute yang ada, bentuk daun yang ada disini akan mempresentasikan sebuah class yang ada, kemudian cabang yang ada akan mempresentasikan sebuah nilai dari attribute itu sendiri,

Decision tree atau pohon keputusan memiliki nama lain yaitu CART (Classification and Regression Tree), yang dimana metode ini menggabungkan Classification tree dan regression tree menjadi satu. Konsep dari sebuah decision tree / pohon keputusan ini adalah untuk mengubah sebuah data menjadi sebuah aturan keputusan.

Menurut (Kretowski, 2019, p. 28) Decision tree atau pohon keputusan adalah sebuah hierarki, struktur berurutan yang terdiri dari simpul dan cabang yang menghubungkan node. node pertama disebut root, dan itu adalah satu-satunya node tanpa predecessor. dua jenis node perlu dibedakan: node internal dan daun. setiap node internal memiliki setidaknya dua atau lebih keturunan, sedangkan daun tidak memiliki keturunan. Decision tree atau pohon keputusan paling sederhana hanya terdiri dari satu node. Dengan setiap node internal yang di tes kemudian sesuai dikaitkan.



Gambar 2. 2 Decision Tree atau Pohon Keputusan (*Maimon & Rokach, 2014, p. 14*)

Menurut (Kashyap, 2018, p. 108) Decision tree atau pohon keputusan adalah sebuah alat atau algoritma untuk pendukung keputusan yang menggunakan grafik mirip sebuah pohon untuk menghasilkan sebuah keputusan. itu adalah metodologi percabangan yang menunjukkan kemungkinan dari hasil sebuah keputusan berdasarkan beberapa kondisi yang ada. Decision tree atau pohon keputusan adalah penggambaran sederhana untuk mengklasifikasikan contoh. algoritma pembelajaran mesin pohon keputusan adalah salah satu teknik yang paling banyak digunakan untuk pembelajaran mesin yang diawasi. pohon keputusan membuat jenis diagram alir, yang terdiri dari node (dikenal sebagai daun) dan serangkaian keputusan yang harus dibuat berdasarkan node (dikenal sebagai cabang). busur yang berasal dari node ditandai dengan fitur dan digunakan untuk mewakili setiap kemungkinan nilai atau hasil fitur. setiap daun dari pohon keputusan ditandai dengan kelas atau distribusi probabilitas di atas kelas. aturan klasifikasi

diwakili melalui jalur dari akar ke node daun. struktur daun dan cabang membentuk representasi hierarki yang menyerupai bentuk pohon. Decision tree atau pohon keputusan adalah salah satu algoritma yang paling banyak digunakan untuk pembelajaran mesin. itu mencerminkan pengambilan keputusan manusia lebih dekat daripada algoritma lainnya. tiga algoritma dasar yang banyak digunakan di bawah ekosistem pohon keputusan adalah id3, c4.5 dan cart.

### 2.2.5 Algoritma C4.5

Algoritma C4.5 adalah sebuah algoritma perluasan dari algoritma ID3 (*Iterative Dichotomiser 3*) yang sudah dikenal dan sering digunakan dalam mengklasifikasi sebuah data yang merupakan struktur pohon keputusan yang dimana teknik penyelesaiannya berupa hasil entropi dan hasil gain. Didalam data yang di klasifikasi tersebut terdapat sebuah atribut seperti atribut numerik dan atribut ketegorial yang dimana setiap cabang akan menggambarkan sebuah hasil dari atribut yang di uji dan setiap daun akan menggambarkan sebuah class.

Menurut (Bulolo, 2020, p. 35) "Algoritma C4.5 adalah algoritma yang digunakan untuk mengeksplorasi sebuah data untuk menemukan hubungan variable input maupun attribute kriteria dengan variable target atau decision attribute. Hasil dari eksplorasi data yang didapat akan diimplementasi kedalam bentuk pohon keputusan atau *decision tree* sehingga dapat mudah untuk dipahami. Algoritma ini dapat mengubah sebuah data yang besar menjadi pohon keputusan atau *decision tree* yang dimana dapat dibaca dalam bentuk sebuah rule".

Menurut (YENDRIZAL, 2022, p. 9) terdapat tiga prinsip kerja dari algoritma C4.5 dari data yang bisa dijelaskan, diantaranya adalah:

1. Pembuatan pohon keputusan

Algoritma C4.5 memiliki sebuah objektif yang dimana pohon keputusan yang ada akan mengkonstruksi struktur data pohon yang akan digunakan untuk memprediksi sebuah class dari kasus baru maupun record baru yang belum memiliki sebuah class. Algoritma C4.5 akan mengkonstruksi pohon keputusan dengan strategi divide dan conquer, yang dimana hanya akan dibuat node akar dengan menerapkan divide dan conquer. Algoritma ini akan memilih pemecahan kasus yang paling terbaik dengan menghitung dan membandingkan dari hasil gain ratio yang dimiliki.

2. Pemangkasan pohon keputusan dan evaluasi (opsional)

Jika hasil pohon keputusan yang sudah dikonstruksi berukuran besar dan sulit untuk dibaca maupun dipahami, algoritma C4.5 bisa menyederhanakan pohon keputusan tersebut dengan melakukan pemangkasan nilai berdasarkan *confidence level* yang ada.

3. Pembuatan aturan-aturan dari pohon keputusan (opsional)

Rules dalam if-then akan diturunkan dari pohon keputusan dengan melakukan penelusuran dari akar sampai ke daun, yang dimana setiap node yang ada serta syarat dari pencabangannya akan diberikan rules if. Sedangkan nilai pada daun akan ditulis di then yang dimana ketika semua aturan tersebut dibuat, aturan tersebut akan di gabungkan.



Menurut (Builolo, 2020, p. 35) dalam melakukan penyelesaian kasus yang ada menggunakan algoritma C4.5 ada elemen-elemen yang harus dicari, diantaranya adalah:

a. Entropy(S)

Entropy(S) adalah jumlah bit yang dibutuhkan untuk mengekstrak suatu class dari jumlah data acak pada jumlah sample yang digunakan untuk menyatakan suatu class. Makin rendah hasil nilai entropy yang didapat, maka tingkat keberagaman data yang digunakan berarti semakin rendah, jika semakin tinggi hasil nilai entropy yang didapat, maka tingkat keberagaman data semakin tinggi.

Entropy memiliki rumus perhitungan, di antaranya adalah:

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

Penjelasan:

S : merupakan jumlah sample data yang digunakan

n : merupakan jumlah partisi dari S

$p_i$  : merupakan proporsi dari S terhadap S.

b. Gain(S, A)

Gain(S, A) digunakan untuk pembetulan sebuah node dan cabang dari pohon keputusan atau decision tree yang dimana didapatkan dari hasil selisih nilai total entropy yang didapat lalu dikurangi nilai entropy dari masing-masing nilai setiap attribute kriteria kemudian dikali dengan proporsi nilai attribute dan diakhiri dengan dibagi jumlah sampel data yang digunakan. Fungsi dari nilai

gain adalah untuk mengukur seberapa tinggi efektivitas dari attribute kriteria dalam mengklasifikasikan sebuah data. Gain memiliki rumus perhitungan, diantaranya adalah:

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

Penjelasan:

S : merupakan jumlah sample data yang digunakan

A: merupakan atribut

n : merupakan jumlah partisi dari S

|S<sub>i</sub>| : merupakan jumlah kasus pada partisi ke-i

|S| : merupakan jumlah kasus dalam S

## 2.3 Teori Analisa dan Perancangan

### 2.3.1 XAMPP

XAMPP berasal dari singkatan X(cross platform), A(Apache), M(MySQL), P(PHP) dan P(Perl) adalah sebuah software atau aplikasi pada komputer yang biasa digunakan didalam dunia website yang digunakan untuk membuat website itu sendiri. XAMPP adalah sebuah software yang bersifat *open source* yang berbasis sebuah web server yang dapat digunakan menggunakan sebuah system operasi seperti Windows, Linux dsb.

### 2.3.2 Netbeans

Netbeans adalah sebuah perangkat lunak atau software bersifat *open source* yang pada umumnya digunakan dalam dunia programmer serta

developer untuk membuat sebuah aplikasi *integrated development environment* atau ide yang dimana dapat digunakan dengan menggunakan bahasa pemrograman java, bahasa pemrograman c, bahasa pemrograman c++ serta dapat menggunakan bahasa pemrograman php.

Menurut (Salter, 2015, p. 297) Netbeans dikembangkan sebagai sebuah rangkaian modul. Yang dimana sifat modular netbeans ini adalah salah satu fitur utama IDE yang memungkinkan modul tambahan ditulis dengan mudah yang dapat memberikan fungsionalitas tambahan kepada pengembang. Sifat modular dari netbeans juga berarti bahwa netbeans sangat cocok sebagai dasar untuk mengembangkan aplikasi dekstop.

Netbeans menyediakan sejumlah besar fungsionalitas sebagai sebuah standar. Namun, masih ada ekosistem yang berkembang di dalam plugin pihak ketiga yang memungkinkan netbeans berinteraksi dengan database noSQL seperti mongoDB atau kerangka kerja web modern seperti NodeJs.

Netbeans memiliki beberapa fitur yang umumnya membantu user pada saat menggunakannya, diantaranya adalah:

1) *Error Stripe*

Fitur ini memungkinkan user untuk mengetahui bagian mana yang terdapat sebuah eror, yang dimana ditandai dengan *highlight* berwarna merah pada sebuah garis.

2) *Smart Code Completion*

Fitur ini memberikan user sebuah rekomendasi kepada user untuk menggunakan nama variable pada suatu tipe dan parameter

pada sebuah *method* yang digunakan, serta melengkapi keyword yang belum lengkap atau selesai.

### 3) *Bookmarking*

Fitur ini memungkinkan user untuk menandai sebuah baris yang sudah ada untuk menjadi pengingat ketika ingin memodifikasi baris tersebut.

### 4) *Code Generator*

Fitur ini memungkinkan user untuk menghasilkan sebuah konstruktor, penyetel dan pengambil pada sebuah *method*.

### 5) *Go to Commands*.

Fitur ini memungkinkan user untuk dapat berpindah ke sebuah deklarasi variable, asal kode atau file yang terdapat pada project yang dijalankan.

## 2.3.3 Java

Java adalah sebuah bahasa pemrograman yang dapat dijalankan di berbagai macam platform-platform yang ada seperti android, linux, windows tanpa harus di modifikasi ulang di karenakan java memiliki sebuah kode pemrograman level tinggi dan sebuah syntax atau aturan yang ketat. Java juga memiliki 2 buah platform seperti *java application programming interface* atau java api dan *java virtual machine* atau jvm, pada umumnya java digunakan untuk membuat sebuah aplikasi berbasis desktop, mobile dan juga website.

Menurut (Dewi et al., 2021, p. 63) “Java merupakan teknologi yang unik serta revolusioner yang menganut semboyan *write once run anywhere*

yang berarti menulis sekali dijalankan dimana saja, sehingga program java dapat dijalankan di berbagai macam platform system operasi seperti linux, windows maupun unix yang memiliki kelebihan dibandingkan dengan bahasa pemrograman lainnya”.

Java pertama kali diciptakan pada tahun 1991 dengan nama Oak, yang dimana diciptakan oleh seorang programmer bernama James Gosling yang bekerja di perusahaan Sun Microsystem. Pada mulanya Gosling mengembangkan bahasa pemrograman ini sebagai otak dari peralatan pintar yang ada menggunakan bahasa pemrograman c++. Tetapi Gosling tidak puas dengan hasil yang didapatnya menggunakan bahasa pemrograman c++, sehingga Gosling pergi menemui rekan-rekannya untuk melakukan pengembangan bahasa pemrograman baru yang dapat memenuhi kebutuhan yang diperlukannya dengan tetap menggunakan aturan atau syntax dasar yang ada pada bahasa pemrograman c++.

Menurut (Rusli et al., 2016, p. 2) “Java adalah sebuah bahasa pemrograman yang tergolong pada sebuah *high level language* yang dimana memiliki arti mudah bagi manusia untuk dapat memahaminya, yang dimana memerlukan sebuah *syntax* atau aturan yang sangat ketat”.

#### **2.3.4 phpMyAdmin**

phpMyAdmin adalah sebuah aplikasi atau software scripting language PHP open source yang berbasis web yang dapat digunakan untuk membantu dalam hal pengelolaan sebuah database pada MySQL agar menjadi lebih mudah. phpMyAdmin disini tidak hanya dapat membantu

MySQL saja, melainkan dapat membantu seperti Drizzle dan MariaDB untuk mengelola sebuah database.

phpMyAdmin pertama kali diciptakan oleh seorang yang bernama Tobias Ratschiller yang bekerja sebagai konsultan di bidang IT. Pada saat itu, pada tahun 1998 Ratschiller yang bekerja sebagai konsultan di bidang IT sedang mengerjakan sebuah program front-end berbasis PHP dengan MySQL yang terinspirasi oleh program buatan Peter Kuppewieser berbasis MySQL-Webadmin yang dimana sudah tidak lagi dikembangkan.

Menurut (Rahman, 2013, p. 12) phpMyAdmin adalah sebuah aplikasi PHP, yang dimana phpMyAdmin disini bergerak sebagai administrator MySQL. phpMyAdmin juga mendukung berbagai aktivitas MySQL seperti pengelolaan data, table, relasi antar table dsb. Tanpa Program seperti phpMyAdmin ini, segala aktivitas yang berhubungan dengan query pada sebuah database MySQL hanya akan dapat dilakukan orang yang sudah handal dibidang tersebut. Tetapi dengan adanya phpMyAdmin disini, jika ingin mengelola sebuah database MySQL akan menjadi sangat lebih mudah dan dapat dilakukan oleh semua kalangan termasuk para pemula yang baru belajar.

### 2.3.5 MySQL

MySQL adalah sebuah DBMS(*database management system*) yang bersifat open source. MySQL memiliki 2 buah lisensi yang terlekat, yang pertama adalah lisensi Free Software dan yang kedua adalah lisensi Shareware, yang dimana kedua lisensi ini termasuk kedalam lisensi *General*

*Public License (GPL)* yang dimana disini artinya dapat digunakan untuk kepentingan pribadi maupun kepentingan komersial yang dimana tidak memerlukan biaya apapun untuk membeli lisensi yang ada.

Awal terciptanya MySQL adalah ketika seorang programmer asal negara Swedia yang bernama Michael “Monty” Widenius yang bekerja di perusahaan yang bernama TcX sedang mengembangkan system database yang sangat sederhana yang diberi nama UNIREG. Pada tahun 1994, perusahaan TcX tempat Monty bekerja ini mulai tertarik dengan project UNIREG yang dikembangkan oleh nya yang dimana akan berencana untuk menjadikan UNIREG sebagai system database untuk project website yang akan mereka luncurkan, tetapi ternyata UNIREG ini tidak terlalu cocok untuk sebuah database dinamis pada sebuah website, kemudian TcX disini mencari sebuah ide baru lalu menemukan sebuah software yang bernama miniSQL yang dikembangkan oleh David Hughes. Tetapi setelah dilakukan percobaan, miniSQL yang dikembangkan oleh David Hughes ini masih belum memenuhi kebutuhan yang diperlukan TcX. Pada Akhirnya UNIREG yang dijalankan oleh Morty kemudian miniSQL yang dijalankan oleh David Hughes serta TcX melakukan sebuah kerjasama untuk mengembangkan sebuah system database yang belum pernah diciptakan yang dimana pada tahun 1995 mereka merilis MySQL.

### **2.3.6 JDK (*Java Development Kit*)**

*Java Development Kit* atau yang kita kenal dengan JDK adalah sebuah paket produk *open source* yang dikeluarkan oleh Oracle Corporation setelah mengakuisisi perusahaan Sun Microsystem. Dimana paket produk ini

berisi komponen yang bertujuan untuk membantu user dalam melakukan pengembangan aplikasi menggunakan bahasa pemrograman java.

Menurut (Komputer, 2014, p. 18) JDK dilengkapi berbagai macam komponen yang dapat membantu dalam proses pemrograman, diantaranya adalah:

1. Java

Tool ini adalah sebuah loader untuk aplikasi java yang dimana dapat menerjemahkan sebuah file class yang sudah di generate oleh javac compiler.

2. Javac

Tool ini adalah sebuah compiler yang mengonversi source code java ke dalam java bytecode.

3. Appletviewer

Tool ini adalah sebuah tool yang memungkinkan user untuk dapat menjalankan java dan debug applet java tanpa web browser.

4. Apt

Tool ini adalah sebuah tool yang digunakan untuk melakukan proses *annotation*

5. Extcheck

Tool ini adalah sebuah tool utility yang dapat mendeteksi konflik di file JAR

6. Javadoc

Tool ini adalah sebuah tool generator dokumentasi yang akan menghasilkan sebuah dokumentasi secara otomatis dari komentar pada source code yang ada.



### 2.3.7 Rapidminer

Rapidminer adalah sebuah aplikasi atau software yang bersifat open source, yang pada umumnya digunakan sebagai alat pembelajaran data mining maupun digunakan untuk melakukan proses pengolahan maupun analisis terhadap data mining. Rapidminer dapat memberikan bantuan dalam pengambilan keputusan yang baik dikarenakan menggunakan teknik prediksi dan teknik deksriptif.

Menurut (Jollyta et al., 2021, p. 87) Rapidminer adalah sebuah alat penambangan data yang digunakan untuk membuat model prediktif dengan sangat cepat yang dapat menampilkan ratusan persiapan data dan algoritma pembelajaran untuk mendukung semua project data mining. Rapidminer adalah paket perangkat lunak yang memungkinkan untuk melakukan penambangan teks, penambangan data dan melakukan analitik prediktif yang memungkinkan pengguna untuk memasukan data mentah seperti database dan teks kemudian akan lakukan proses analisis secara otomatis dalam skala yang besar.

### 2.3.8 UML (*Unified Modeling Language*)

*Unified Modeling Language* atau UML adalah suatu metode yang berorientasi pada objek yang bermodel secara visual, yang dimana UML ini sudah menjadi standart utama dalam visualisasi dalam pembuatan sebuah software dan juga pada saat melakukan dokumentasi software. Dengan menggunakan UML, desain software yang akan dibuat akan ditampilkan kedalam bentuk sebuah diagram dan juga sebuah simbol-simbol yang memiliki makna masing-masing. Desain yang ditampilkan kedalam sebuah

bentuk diagram dan juga simbol-simbol, dapat diterjemahkan menjadi sebuah kode program yang dimana kode program ini dapat ditulis dengan bahasa pemrograman apapun.

Menurut (Sari & Indra, 2021, p. 9) UML adalah sebuah bahasa pemodelan grafik yang umumnya digunakan sebagai membantu serta mendesain sebuah system perangkat lunak atau *software* yang khususnya berorientasi sebuah objek. UML mencakup berbagai macam masalah yang meliputi seperti visualisasi, spesifikasi, konstruksi serta dokumentasi software maupun sebuah model bisnis. UML juga umumnya dapat digunakan dalam pengembangan seperti pengembangan tahapan dari segi analisis kebutuhan system sampai ke pengembangan tahapan dari segi implementasi system tersebut.

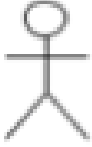
UML memiliki beberapa diagram yang biasanya digunakan, diantaranya adalah:




#### 2.3.8.1 Use Case Diagram


Use Case Diagram adalah sebuah diagram yang terdapat pada UML, yang dimana diagram ini akan menjelaskan atau mendeskripsikan sebuah hubungan antara system dengan pengguna nya.

Tabel 2. 1 Tabel Simbol Use Case Diagram

No	Gambar	Nama	Keterangan
----	--------	------	------------

1		Actor	Menggambarkan tentang user atau individu maupun organisasi ketika berinteraksi dengan use case maupun system.
2		System	Menggambarkan tentang suatu system secara terbatas.
3		Use Case	Menggambarkan tentang abstrak interaksi dari system dengan user atau individu maupun organisasi.
4		Association	Menggambarkan tentang hubungan dengan objek


			satu dengan objek lainnya.
5		Dependency	Menggambarkan tentang perubahan yang terjadi pada elemen yang akan mempengaruhi elemen tersebut.
6		Extend	Menggambarkan tentang spesifikasi use case yang memperluas jika suatu kondisi terpenuhi.
7		Generalization	Menggambarkan tentang spesialisasi user atau individu maupun organisasi yang dapat






			berpartisipasi di use case.
8		Include	Menggambarkan tentang suatu use case yang ada merupakan fungsionalitas dari use case satu dengan yang lainnya.

### 2.3.8.2 Activity Diagram

Activity Diagram adalah sebuah diagram yang terdapat pada UML, yang dimana diagram ini akan menjelaskan atau mendeskripsikan tentang kejadian-kejadian didalam use case diagram.

Tabel 2. 2 Tabel Simbol Activity Diagram

No	Simbol	Nama	Keterangan
1		Status Awal	Menggambarkan sebuah system. activity diagram mempunyai status awal




2		Status Akhir	Menggambarkan sebuah status akhir yang dilakukan system, activity diagram.
3		Aktivitas	Menggambarkan sebuah aktivitas yang dilakukan system, activity diagram.
4		Penggabungan	Menggambarkan sebuah aktivitas satu dengan aktivitas lainnya menjadi satu.
5		Percabangan	Menggambarkan aktivitas yang ada lebih dari satu aktivitas.
6		Swimlane	Menggambarkan pemisahan organisasi yang bertanggung jawab terhadap


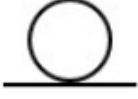


			aktivitas yang dilakukan
--	--	--	--------------------------

### 2.3.8.3 Sequence Diagram

Sequence Diagram adalah sebuah diagram yang terdapat pada UML, yang dimana diagram ini akan menggambarkan kelakuan sebuah objek berurutan sesuai waktu pada use case diagram.

Tabel 2. 3 Tabel Simbol Sequence Diagram

No	Gambar	Nama	Keterangan
1		Actor	Menggambarkan tentang user atau individu yang sedang berinteraksi dengan system yang ada.
2		Objek	Menggambarkan tentang komponen utama pada sequence diagram.
3		Boundary Class	Menggambarkan tentang penjelasan atau penggambaran dari suatu form.

4		Control Class	Menggambarkan tentang hubungan table dengan boundary.
5		Entity Class	Menggambarkan tentang hubungan kegiatan yang ada pada suatu system
6		Life Line	Menggambarkan tentang dimulai nya message dan berakhir nya message.
7		Message	Menggambarkan tentang aktivitas pengiriman pesan.

#### 2.3.8.4 Class Diagram

Class Diagram adalah sebuah diagram yang terdapat pada UML, yang dimana diagram ini akan menggambarkan hubungan antara kelas-kelas yang ada didalam desain model yang ada dari system yang akan dibuat.

Tabel 2. 4 Tabel Simbol Class Diagram

No	Gambar	Nama	Keterangan
----	--------	------	------------



1		Generalization	Menggambarkan tentang relasi antara kelas dengan makna generalisasi spesialis.
2		Class	Menggambarkan tentang objek-objek dari berbagai atribut.
3		Association	Menggambarkan tentang penghubungan objek satu dengan objek lainnya.
4		Collaboration	Menggambarkan tentang urutan aksi yang ditampilkan suatu system yang menghasilkan suatu hasil.
5		Nary Association	Menggambarkan tentang opsi mencegah atau menghindari asosiasi yang terjadi oleh 2 objek lebih.
6		Dependency	Menggambarkan tentang perubahan elemen mandiri yang akan mempengaruhi elemen yang tidak mandiri.
7		Realization	Menggambarkan tentang operasi atau kegiatan dilakukan oleh suatu objek yang ada.

### 2.3.8.5 Statemachine Diagram

Statemachine Diagram adalah sebuah diagram yang terdapat pada UML, yang dimana diagram ini akan menggambarkan perubahan keadaan terhadap suatu objek.

Tabel 2. 5 Tabel Simbol Statemachine Diagram

No	Gambar	Nama	Keterangan
1		Titik Awal	Menggambarkan tentang permulaan kejadian atau kegiatan yang ada pada state diagram.
2		Titik Akhir	Menggambarkan tentang berakhirnya kejadian atau kegiatan yang ada pada state diagram.
3		State	Menggambarkan tentang kondisi yang terjadi.
4		Point	Menggambarkan tentang apakah tahap selanjutnya akan masuk kedalam state atau keluar state.
5		Event	Menggambarkan tentang satu kondisi yang terjadi pada state.

6	[guard]	Guard	Menggambarkan tentang syarat terjadinya transisi yang bersangkutan.
---	---------	-------	---

### 2.3.9 ERD (Entity Relationship Diagram)

*Entity Relationship Diagram* atau ERD adalah sebuah diagram pada pemodelan basis data yang digunakan untuk menggambarkan relasi antar objek pada sebuah data secara detail. Pada umumnya ERD digunakan untuk menyusun atau membuat sebuah database menjadi lebih mudah dalam menampilkan data yang memiliki sebuah relasi.

Menurut (Yanto, 2016, p. 32) "ERD adalah suatu diagram yang menggambarkan desain konseptual dari suatu basis data yang memiliki relasi yang dimana ERD menggambarkan relasi antar objek satu dengan objek lainnya".

## 2.4 Tinjauan Studi

### 2.4.1 Penelitian Fauriatun Helmiah

Tabel 2. 6 Penelitian Pertama

NO	DATA JURNAL/MAKALAH	KETERANGAN
1.	Judul	Penerapan Algoritma C4.5 Dalam Klasifikasi Kualitas Sayur Kol Di Kabupaten Simalungun
2.	Jurnal	J-Com (Journal of Computer)

3.	Volume dan halaman	Vol 1, No. 3
4.	Tanggal dan Tahun	November 2021
5.	Penulis	Fauriatun Helmiah , Yulia Purnama Sari Hutagaol , Sumantri
6.	Penerbit	STMIK Royal , Sumatera Utara
7.	Tujuan Penelitian	Melakukan penelitian yang bertujuan untuk menentukan kualitas telur ayam Australia terbaik
8.	Lokasi dan Subjek Penelitian	CV Bintang Ternak
9.	Perancangan Sistem	<ol style="list-style-type: none"> <li>1. Rapid Miner</li> <li>2. Algoritma C4.5</li> <li>3. Microsoft Excel</li> <li>4. Sistem Informasi</li> </ol>
10.	Teknik Pengumpulan Data	Teknik pengumpulan data yang dipakai adalah , menggunakan data internal dari CV Bintang Ternak
11.	Hasil Penelitian	Terbentuk system informasi yang dapat menentukan kualitas telur ayam Australia menggunakan attribute yang ada.
12.	Kekuatan Penelitian	Attribute yang digunakan oleh penulis , sangat efektif dalam artian attribute yang ada dapat digunakan untuk menentukan kualitas telur ayam Australia yang ada dan menghasilkan sebuah hasil.
13.	Kelemahan Penelitian	Tidak ada metode lain yang digunakan didalam penelitian, sehingga tidak diketahui apakah

		metode yang dipakai sangat efektif dibandingkan metode yang lainnya.
14.	Kesimpulan	Penerapan metode c4.5 ini menghasilkan sebuah system informasi yang dapat menentukan kualitas telur ayam Australia terbaik dengan memilih atribut sebagai akar lalu membuat cabang untuk setiap nilai , mebagi kasus di dalam cabang serta mengulangi proses pada cabang sampai semua kasus memiliki kelas yang sama
15.	Link Jurnal	<a href="https://jurnal.stmikroyal.ac.id/index.php/j-com/article/view/1348">https://jurnal.stmikroyal.ac.id/index.php/j-com/article/view/1348</a>

#### 2.4.2 Penelitian Asep Muhidin

Tabel 2. 7 Penelitian Kedua

NO	DATA JURNAL/MAKALAH	KETERANGAN
1.	Judul	Penerapan Data Mining Untuk Klasifikasi Kualitas Pipa PVC Menggunakan Metode Algoritma C4.5, Studi Kasus PT Cipta Aneka Agung
2.	Jurnal	SIGMA
3.	Volume dan Halaman	Vol 9, No. 4
4.	Tanggal dan Tahun	Juni 2019
5.	Penulis	Asep Muhidin , Ading Bagus Saputra
6.	Penerbit	STT Pelita Bangsa

7.	Tujuan Penelitian	Untuk mengetahui hasil prediksi uji kualitas pipa PVC dengan melihat keakuratan algoritma C4.5 sehingga memudahkan dalam proses QC untuk menentukan kualitas pipa tersebut.
8.	Lokasi dan Subjek Penelitian	PT Cipta Aneka Agung
9.	Perancangan Sistem	<ol style="list-style-type: none"> <li>1) Rapid Miner</li> <li>2) Algoritma C4.5</li> <li>3) Microsoft Excel</li> <li>4) Sistem Informasi</li> </ol>
10.	Teknik Pengumpulan Data	Teknik pengumpulan data yang dipakai adalah , menggunakan data internal dari PT Cipta Aneka Agung
11.	Hasil Penelitian	Mendapatkan hasil pengukuran data accuracy yang diperoleh dari proses data testing mencapai 91.67%
12.	Kekuatan Penelitian	<p>Hasil pengukuran data accuracy yang diperoleh dari data testing mencapai 91.67%. Dari data tersebut diketahui prediksi NG dengan true NG mencapai 30 pipa dan true OK sebanyak 1 pipa, dengan pencapaian class precision 96.77%.</p> <p>Sedangkan untuk prediksi OK dengan true NG sebanyak 4 pipa dan untuk true OK mencapai 25 pipa dengan pencapaian class precision 86.21%.</p> <p>Untuk class recall dengan true NG mencapai</p>

		88.24% sedangkan untuk class recall dengan true OK mencapai 96.15%
13.	Kelemahan Penelitian	Hanya menggunakan satu metode, yaitu metode c4.5.
14.	Kesimpulan	Dari hasil perhitungan data training data uji kualitas pipa dengan algoritma c4.5 menggunakan data training dengan confusion matrix mendapatkan hasil nilai accuracy sebesar 92.92%, precision 87.64% dan recall 93.75% serta kurva ROC optimistic dengan akurasi excellent classification sebesar 0.984
15.	Link Jurnal	<a href="https://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/435">https://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/435</a>

### 2.4.3 Penelitian Mhd Wendico Herdian

Tabel 2. 8 Penelitian Ketiga

NO.	DATA JURNAL/MAKALAH	KETERANGAN
1.	Judul	Penerapan Algoritma C4.5 Dalam Klasifikasi Kualitas Sayur Kol Di Kabupaten Simalungun
2.	Jurnal	RESOLUSI
3.	Volume dan Halaman	Vol 2, No. 3, Hal 132-137
4.	Tanggal dan Tahun	Januari 2022
5.	Penulis	Mhd Wendico Herdian, Riki Winanjaya, Susiani
6.	Penerbit	STIKOM Tunas Bangsa, Sumatera Utara

7.	Tujuan Penelitian	Untuk menentukan sayur kol yang layak dan tidak layak dikarenakan factor cuaca yang tidak menentu dan hama tanaman yang mengancam menjadikan kualitas sayur kol kurang baik sehingga memberikan dampak kepada penghasil kol.
8.	Lokasi dan Subjek Penelitian	Kabupaten Simalungun , Petani Kol , Pedagang Kol , Agen Kol
9.	Perancangan Sistem	<ol style="list-style-type: none"> <li>1) Rapid Miner</li> <li>2) Algoritma C4.5</li> <li>3) Microsoft Excel</li> <li>4) Sistem Informasi</li> </ol>
10.	Teknik Pengumpulan Data	Teknik pengumpulan data yang dipakai adalah , melakukan wawancara kepada petani , pedagang dan agen kol yang ada serta melakukan observasi
11.	Hasil Penelitian	Hasil penelitian ini adalah ditemukan hasil klasifikasi yang membentuk dua kelas yaitu kelas layak dan tidak layak yang dimana menghasilkan tingkat akurasi yang tinggi.
12.	Kekuatan Penelitian	Data yang tersedia dapat diterapkan menggunakan data mining dengan metode c4.5 , menghasilkan tingkat accuracy yang besar , dan yang digunakan validasi dan dapat dibuktikan dengan tools RapidMiner.



13.	Kelemahan Penelitian	Hanya menggunakan satu metode , yaitu metode c4.5.
14.	Kesimpulan	Berdasarkan hasil pengujian software RapidMiner diperoleh hasil yang sama dengan analisis perhitungan Algoritma C4.5 dimana diperoleh tingkat accuracy yaitu sebesar 100%.
15.	Link Jurnal	<a href="http://djournals.com/resolusi/article/view/312">http://djournals.com/resolusi/article/view/312</a>

#### 2.4.4 Rangkuman Model Penelitian

Tabel 2. 9 Rangkuman Penelitian

PENELITI	NAMA JURNAL	TAHUN	INSTITUSI	JUDUL DAN METODE YANG DIGUNAKAN	KESIMPULAN
Fauriatun Helmiah	J-Com (Journal of Computer)	2021	STMIK Royal	Penerapan Algoritma C4.5 Dalam Klasifikasi Kualitas Sayur Kol Di Kabupaten Simalungun	Penerapan metode c4.5 ini menghasilkan sebuah system informasi yang dapat menentukan kualitas telur ayam Australia terbaik dengan

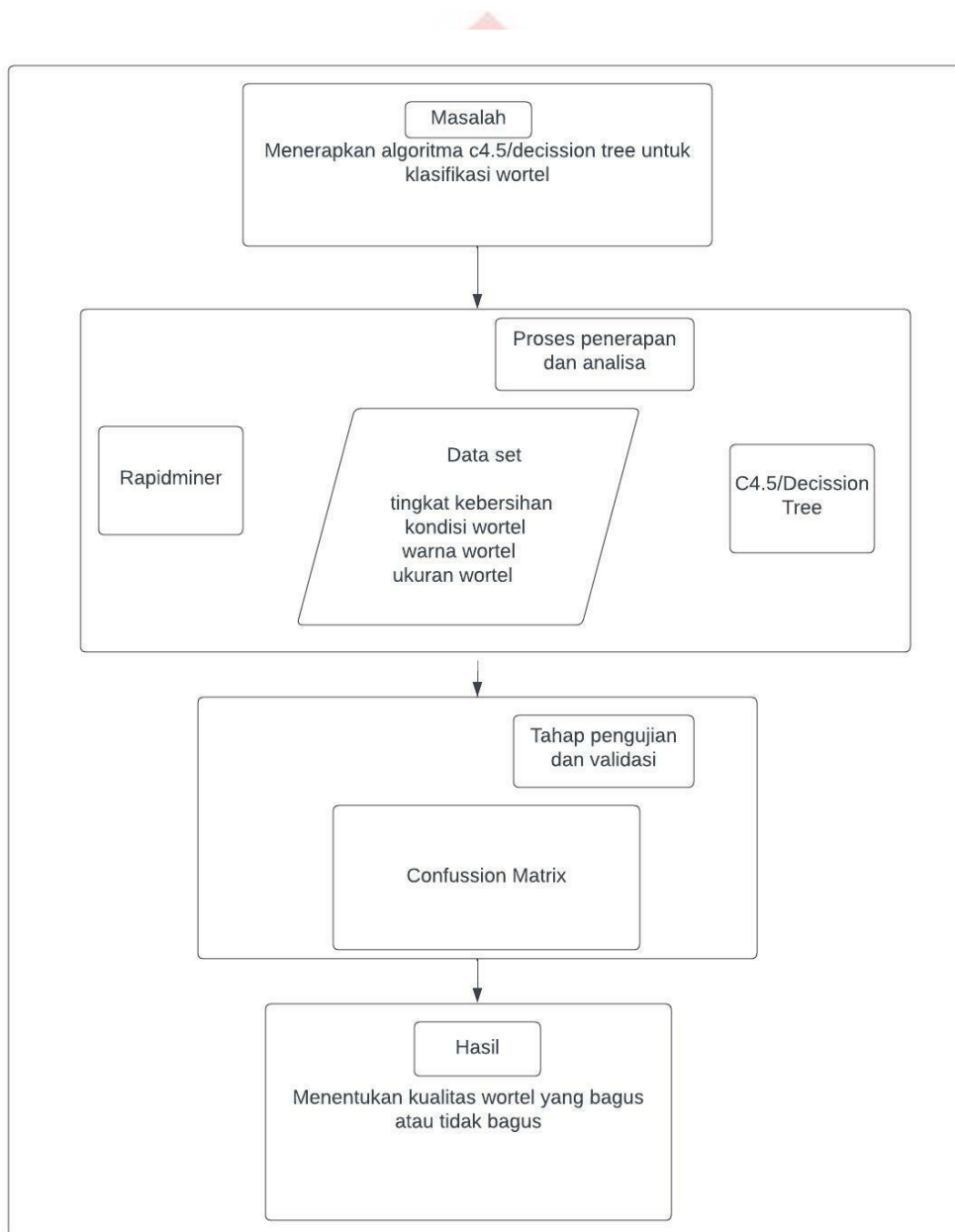
					<p>memilih atribut sebagai akar lalu membuat cabang untuk setiap nilai , mebagi kasus di dalam cabang serta mengulangi proses pada cabang sampai semua kasus memiliki kelas yang sama</p>
Asep Muhidin	SIGMA	2019	STT Pelita Bangsa	<p>Penerapan Data Mining Untuk Klasifikasi Kualitas Pipa PVC Menggunakan Metode Algoritma C4.5, Studi Kasus PT Cipta Aneka Agung</p>	<p>Dari hasil perhitungan data training data uji kualitas pipa dengan algoritma c4.5 menggunakan data training dengan confusion matrix</p>

					<p>mendapatkan hasil nilai accuracy sebesar 92.92% , precision 87.64% dan recall 93.75% serta kurva ROC optimistic dengan akurasi excellent classification sebesar 0.984</p>
Mhd Wendico Herdian	RESOLUSI	2022	STIKOM Tunas Bangsa	<p>Penerapan Algoritma C4.5 Dalam Klasifikasi Kualitas Sayur Kol Di Kabupaten Simalungun</p>	<p>Berdasarkan hasil pengujian software RapidMiner diperoleh hasil yang sama dengan analisis perhitungan Algoritma C4.5 dimana diperoleh tingkat</p>

					accuracy yaitu sebesar 100%.
--	--	--	--	--	------------------------------

## 2.5 Kerangka Pemikiran

Berikut ini adalah gambaran dari kerangka pemikiran keseluruhan dari penelitian yang dilakukan, diantaranya adalah:



Gambar 2. 3 Kerangka Pemikiran

## **BAB III**

### **ANALISA SISTEM YANG BERJALAN**

#### **3.1 Tinjauan Umum Perusahaan**

##### **3.1.1 Sejarah Perusahaan**

PT Padma Mekar Lestari adalah sebuah perusahaan yang bergerak di bidang supplier sayuran atau buah wortel yang memiliki kualitas wortel yang baik, bagus dan juga berkualitas. PT Padma Mekar Lestari membeli wortel dalam bentuk bahan baku yang kemudian akan dijadikan barang jadi untuk dijual kembali di pasar tradisional maupun pasar modern.

Resmi berdiri sejak tahun 2018, yang dimana pada awal mulanya perusahaan ini adalah bisnis perorangan yang sudah dijalankan sejak tahun 1997 dengan modal awal sebanyak Rp. 2.500.000, yang dimana pada saat berdirinya perusahaan ini peluang untuk berkembang di bidang usaha penjualan wortel ini sangatlah luas serta menyakinkan dikarenakan masih belum banyak nya peminat dibidang ini serta kebutuhan sayur atau buah wortel ini sangat banyak kegunaan seperti dijadikan makanan serta dijadikan minuman, sampai saat ini perkembangan usaha ini tetap berjalan dengan rencana awal yang sudah direncanakan.



Gambar 3. 1 Produk



Gambar 3. 2 Produk

Saat ini, PT Padma Mkear Lestari memiliki 6 buah karyawan office dan juga kurang lebih sekitar 40 karyawan produksi dan gudang, yang berlokasi di Pergudangan Modern Business Park Nomor 95, Cipondoh Tangerang. Hingga pada waktu ini, PT Padma Mekar Lestari telah melakukan proses penjualan sekitar jabodetabek hingga diluar jabodetabek.

### 3.1.2 Visi Perusahaan

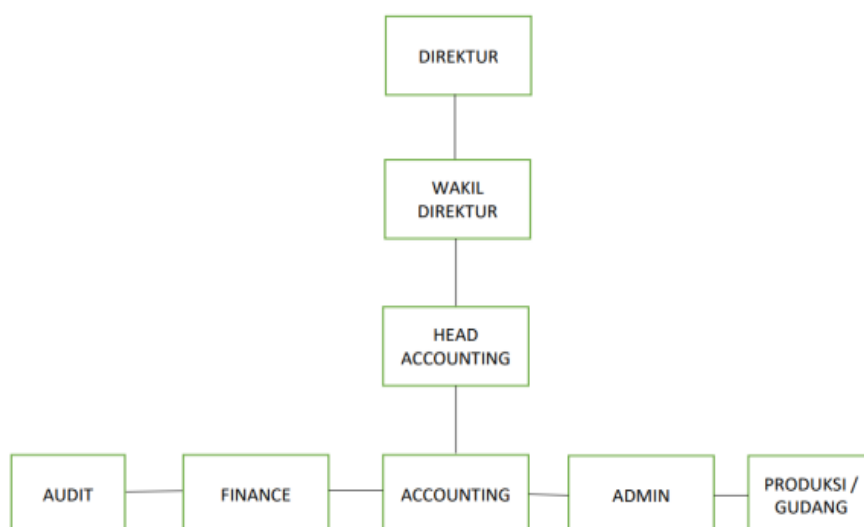
Menjadikan setiap produk memiliki nilai tambah dalam hal penampilan dan juga daya tahan.

### 3.1.3 Misi Perusahaan

Memahami karakter dan penanganan setiap produk dengan baik dan juga tepat.

### 3.1.4 Struktur Organisasi

Struktur Organisasi adalah sebuah hierarki atau sebuah system yang digunakan untuk mengidentifikasi atau menjelaskan sebuah komponen-komponen menjadi satu di sebuah organisasi atau perusahaan yang dimana setiap individunya itu berada didalam ruang lingkup organisasi atau perusahaan yang memiliki sebuah tugas masing-masing berdasarkan posisi.



Gambar 3. 3 Struktur Organisasi

### 3.1.5 Fungsi Setiap Bagian Organisasi

Direktur:

- Membuat strategi untuk membuat bisnis semakin berkembang.
- Mengecek laporan seluruh operasional perusahaan.
- Mengawasi semua kegiatan perusahaan.
- Memberi saran dan masukan kepada karyawan.
- Membuat daftar harga.

Wakil Direktur:

- Mengatur jalannya operasional perusahaan.
- Mengatur pembagian tugas kepada karyawan.
- Mengatur rute pada saat pengiriman.

Head Accounting

- Melakukan pengaturan keuangan.
- Membuat laporan keuangan
- Memberikan laporan stock bulanan.
- Membuat rincian anggaran pengeluaran dan pendapatan.
- Membuat rekonsiliasi pajak

Accounting

- Melakukan journal harian.
- Melakukan rekonsiliasi bank.
- Melakukan fileing data.



### Finance

- Melakukan penagihan piutang kepada customer.
- Membuat rekap piutang customer
- Mengawasi atau mengontrol piutang customer

### Auditor

- Melakukan pengecekan stock di gudang.
- Melakukan stock opname.
- Melakukan daily control untuk stock
- Memonitoring keluar dan masuk barang.

### Admin Penjualan

- Membuat surat jalan.
- Merapihkan surat jalan.
- Membantu membuat rute pengiriman
- Menjawab dan menerima telpon atau email yang masuk.
- Membuat faktur penjualan

### Operasional dan Gudang

- Melakukan pengolahan wortel.
- Melakukan pengamanan stock wortel.
- Melakukan pengeluaran dan pemasukan wortel.
- Menyiapkan barang kiriman.
- Merapihkan Gudang

## 3.2 Identifikasi Kebutuhan Sistem

### 3.2.1 Kebutuhan Sumber Data

- Tahap identifikasi produk dilakukan saat produk pertama kali datang dan juga ketika produk masuk ke proses selanjutnya, yang dimana memberikan identitas pada setiap produk yang ada.
- Setelah dilakukan tahap identifikasi, data kemudian dicatat lalu diinput yang menghasilkan dataset seperti berikut:

Tabel 3. 1 Table Dataset

No	Tingkat Kebersihan	Kondisi Wortel	Warna Wortel	Ukuran Wortel	Kualitas Wortel
1	Kotor	Bintik Hitam	Terang	Besar	Tidak Bagus
2	Bersih	Normal	Terang	Kecil	Tidak Bagus
3	Bersih	Tumbuh	Terang	Normal	Tidak Bagus
4	Bersih	Licin	Pucat	Kecil	Tidak Bagus
5	Bersih	Licin	Pucat	Besar	Tidak Bagus
6	Kotor	Normal	Terang	Kecil	Tidak Bagus
7	Bersih	Licin	Pucat	Normal	Tidak Bagus
8	Bersih	Normal	Terang	Normal	Bagus
9	Kotor	Licin	Terang	Normal	Bagus
10	Kotor	Beku	Terang	Normal	Tidak Bagus
11	Kotor	Licin	Pucat	Besar	Bagus
12	Kotor	Tumbuh	Pucat	Besar	Tidak Bagus
13	Bersih	Bintik Hitam	Terang	Besar	Bagus
14	Bersih	Normal	Terang	Normal	Bagus

15	Bersih	Beku	Pucat	Normal	Bagus
16	Kotor	Licin	Pucat	Kecil	Tidak Bagus
17	Bersih	Normal	Pucat	Besar	Bagus
18	Bersih	Normal	Terang	Normal	Bagus
19	Bersih	Normal	Terang	Normal	Bagus
20	Bersih	Tumbuh	Pucat	Besar	Bagus
21	Kotor	Licin	Pucat	Besar	Tidak Bagus
...	.....	.....	.....	.....	.....
247	Kotor	Licin	Terang	Besar	Tidak Bagus
248	Kotor	Normal	Pucat	Besar	Tidak Bagus
249	Kotor	Beku	Terang	Normal	Tidak Bagus
250	Kotor	Licin	Terang	Besar	Tidak Bagus
251	Bersih	Licin	Terang	Normal	Bagus
252	Bersih	Normal	Terang	Besar	Bagus
253	Bersih	Tumbuh	Terang	Normal	Tidak Bagus
254	Kotor	Licin	Pucat	Normal	Tidak Bagus
255	Bersih	Normal	Terang	Normal	Bagus
256	Bersih	Licin	Terang	Besar	Bagus
257	Bersih	Licin	Terang	Besar	Bagus
258	Kotor	Beku	Terang	Normal	Tidak Bagus
259	Bersih	Normal	Terang	Normal	Bagus

Keterangan data:

1. Tingkat Kebersihan : tingkat kebersihan dari wortel yang berisi bersin dan kotor.

2. Kondisi Wortel : kondisi dari wortel yang berisi normal,licin tumbuh dan bitnik hitam.
3. Warna Wortel : kondisi yang dimana warna dari wortel, yang berisi pucat dan terang.
4. Ukuran Wortel : ukuran dari wortel yang berisi normal, besar dan kecil
5. Kualitas : kualitas dari wortel yang berisi bagus dan tidak bagus.

### 3.3 Alternatif Pemecahan Masalah

#### 3.3.1 Kontruksi Algoritma C4.5

Ketika ingin melakukan perhitungan dengan algoritma C4.5, kita harus menggunakan perhitungan sebagai berikut:

- Pertama mencari entropy

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

Dimana:

S : merupakan jumlah sample data yang digunakan

n : merupakan jumlah partisi dari S

$p_i$  : merupakan propori dari S terhadap S.

- Kedua mencari gain.

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

Dimana:

S : merupakan jumlah sample data yang digunakan  
A: merupakan atribut

n : merupakan jumlah partisi dari S

|S<sub>i</sub>| : merupakan jumlah kasus pada partisi ke-i

|S| : merupakan jumlah kasus dalam S

Dalam kasus ini, contoh data yang dilakukan pengujian adalah:

1. Jumlah data sebanyak = 10
2. Jumlah data bagus sebanyak = 5
3. Jumlah data tidak bagus sebanyak = 5

Disini dilakukan tahap pencarian entropy pertama, yang dimana:

1. Total kasus yang ada sebanyak 10, terdiri dari 5 bagus dan 5 tidak bagus, untuk mencari entropy nya digunakan rumus:

$$Entropy(\text{total}) = \left( \frac{5}{10} \times \log^2 \frac{5}{10} \right) + \left( \frac{5}{10} \times \log^2 \frac{5}{10} \right) = 1$$

2. Tahap selanjutnya adalah mencari gain dari tingkat kebersihan, yang dimana sudah dicari nilai dari entropy tingkat kebersihan yaitu bersih = 0,970950 dan kotor 5 = 0,970950.

$$Entropy(\text{total}) = \left( \frac{2}{5} \times \log^2 \frac{2}{5} \right) + \left( \frac{3}{5} \times \log^2 \frac{3}{5} \right) = 0.970950$$

$$\text{Entropy}(\text{total}) = \left(\frac{3}{5} \times \log^2 \frac{3}{5}\right) + \left(\frac{2}{5} \times \log^2 \frac{2}{5}\right) = 0,970950$$

Lalu memperoleh 5 data bersih dengan hasil 3 bagus dan 2 tidak bagus, kemudian 5 data kotor dengan hasil 2 bagus dan 3 tidak bagus yang dimana perhitungannya mencari gainnya adalah:

*Gain(total, tingkat kebersihan)*

$$= 1 - \left(\frac{5}{10} \times 0,970950\right) + \left(\frac{5}{10} \times 0,970950\right) = 0,029049$$

3. Tahap selanjutnya adalah mencari gain dari kondisi wortel, yang dimana sudah dicari nilai dari entropy kondisi wortel yaitu beku = 0,918295, licin = 1 dan normal = 0,918295.

$$\text{Entropy}(\text{total}) = \left(\frac{2}{3} \times \log^2 \frac{2}{3}\right) + \left(\frac{1}{3} \times \log^2 \frac{1}{3}\right) = 0,918295$$

$$\text{Entropy}(\text{total}) = \left(\frac{2}{4} \times \log^2 \frac{2}{4}\right) + \left(\frac{2}{4} \times \log^2 \frac{2}{4}\right) = 1$$

$$\text{Entropy}(\text{total}) = \left(\frac{1}{3} \times \log^2 \frac{1}{3}\right) + \left(\frac{2}{3} \times \log^2 \frac{2}{3}\right) = 0,918295$$

Lalu memperoleh 3 data beku dengan hasil 1 bagus dan 2 tidak bagus kemudian 4 data licin dengan hasil 2 bagus dan 2 tidak bagus dan yang terakhir 3 data normal dengan hasil 2 bagus dan 1 tidak bagus yang dimana perhitungannya mencari gainnya adalah:

*Gain(total, kondisi wortel)*

$$= 1 - \left(\frac{3}{10} \times 0,918295\right) + \left(\frac{4}{10} \times 1\right) + \left(\frac{3}{10} \times 0,918295\right) = 0,049022$$

4. Tahap selanjutnya adalah mencari gain dari warna wortel, yang dimana sudah dicari nilai dari entropy warna wortel yaitu pucat = 0,918295 dan terang = 0,811278.



	Total		10	5	5	1	
	Tingkat Kebersihan						
		Bersih	5	3	2	0,970950	0,029049
		Kotor	5	2	3	0,970950	
	Kondisi Wortel						
		Beku	3	1	2	0,918295	0,049022
		Licin	4	2	2	1	
		Normal	3	2	1	0,918295	
	Warna Wortel						
		Pucat	6	2	4	0,918295	0,124112
		Terang	4	3	1	0,811278	
	Ukuran Wortel						
		<b>Besar</b>	<b>7</b>	<b>2</b>	<b>5</b>	<b>0,863120</b>	<b>0,395815</b>
		Normal	3	3	0	0	

Kemudian dilakukan tahap pencarian entropy kedua yang dimana ukuran wortel nilai besar manjadi kasus selanjutnya karena hasil gain yang didapat lebih besar dari ukuran wortel nilai normal, yang dimana:

1. Total kasus selanjut nya ada sebanyak 7, terdiri dari 2 bagus dan 5 tidak bagus, untuk mencari entropy nya digunakan rumus:



$$\text{Entropy}(\text{total}) = \left(\frac{5}{7} \times \log^2 \frac{5}{7}\right) + \left(\frac{2}{7} \times \log^2 \frac{2}{7}\right) = 0,863120$$

2. Tahap selanjutnya adalah mencari gain dari tingkat kebersihan, yang dimana sudah dicari nilai entropy tingkat kebersihan yaitu bersih = 0,918295 dan kotor = 0,811278.

$$\text{Entropy}(\text{total}) = \left(\frac{2}{3} \times \log^2 \frac{2}{3}\right) + \left(\frac{1}{3} \times \log^2 \frac{1}{3}\right) = 0,918295$$

$$\text{Entropy}(\text{total}) = \left(\frac{3}{4} \times \log^2 \frac{3}{4}\right) + \left(\frac{1}{4} \times \log^2 \frac{1}{4}\right) = 0,811278$$

Lalu memperoleh 3 data bersih dengan hasil 1 bagus dan 2 tidak bagus, kemudian 4 data kotor dengan hasil 1 bagus dan 3 tidak bagus yang dimana perhitungan mencari gainnya adalah:

*Gain(total, tingkat kebersihan)*

$$= 0,863120 - \left(\frac{3}{7} \times 0,918295\right) + \left(\frac{4}{7} \times 0,811278\right) = 0,005977$$

3. Tahap selanjutnya adalah mencari gain dari kondisi wortel, yang dimana sudah dicari nilai entropy kondisi wortel yaitu licin = 0,918295 kemudian normal = 1 dan beku = 0.

$$\text{Entropy}(\text{total}) = \left(\frac{2}{3} \times \log^2 \frac{2}{3}\right) + \left(\frac{1}{3} \times \log^2 \frac{1}{3}\right) = 0,918295$$

$$\text{Entropy}(\text{total}) = \left(\frac{1}{2} \times \log^2 \frac{1}{2}\right) + \left(\frac{1}{2} \times \log^2 \frac{1}{2}\right) = 1$$

$$\text{Entropy}(\text{total}) = \left(\frac{2}{2} \times \log^2 \frac{2}{2}\right) + \left(\frac{0}{2} \times \log^2 \frac{0}{2}\right) = 0$$

Lalu memperoleh 3 data licin dengan hasil 1 bagus dan 2 tidak bagus, kemudian 2 data normal dengan hasil 1 bagus dan 1 tidak bagus dan 2 data beku dengan hasil 0 bagus dan 2 tidak bagus yang dimana perhitungan mencari gainnya adalah:

*Gain(total, kondisi wortel)*

$$= 0,863120 - \left(\frac{3}{7} \times 0,918295\right) + \left(\frac{2}{7} \times 1\right) + \left(\frac{2}{7} \times 0\right) = 0,183850$$

4. Tahap terakhir mencari gain dari warna wortel, yang dimana sudah dicari nilai entropy warna wortel yaitu pucat = 0,721928 dan terang = 1.

$$\text{Entropy}(\text{total}) = \left(\frac{4}{5} \times \log^2 \frac{4}{5}\right) + \left(\frac{1}{5} \times \log^2 \frac{1}{5}\right) = 0,721928$$

$$\text{Entropy}(\text{total}) = \left(\frac{1}{2} \times \log^2 \frac{1}{2}\right) + \left(\frac{1}{2} \times \log^2 \frac{1}{2}\right) = 1$$

Lalu memperoleh 5 data pucat dengan hasil 1 bagus dan 4 tidak bagus dan 2 data terang dengan hasil 1 bagus dan 1 tidak bagus yang dimana perhitungan mencari gainnya adalah:

*Gain(total, warna wortel)*

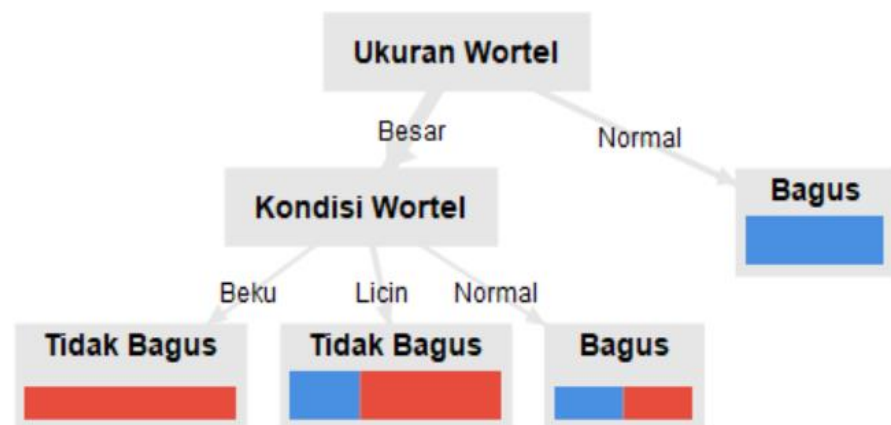
$$= 0,863120 - \left(\frac{5}{7} \times 0,721928\right) + \left(\frac{2}{7} \times 1\right) = 0,061743$$

Tabel 3. 3 Entropy dan Gain Kedua

Node		Jumlah Kasus	Bagus	Tidak Bagus	Entropy	Gain
Total		7	2	5	0.863120	
Tingkat Kebersihan						
	Bersih	3	1	2	0,918295	0,005977
	Kotor	4	1	3	0,811278	
Kondisi Wortel						
	Licin	3	1	2	0,918295	<b>0,183850</b>

		Normal	2	1	1	1	
		Beku	2	0	2	0	
	Warna Wortel						
		Pucat	5	1	4	0,721928	0,061743
		Terang	2	1	1	1	

Setelah melakukan perhitungan *entropy* dan *gain* sebanyak 2 kali, kemudian data diolah menggunakan software rapidminer dengan menggunakan algoritma C4.5 atau *decision tree* yang dimana mendapatkan hasil *decision tree* seperti dibawah ini:



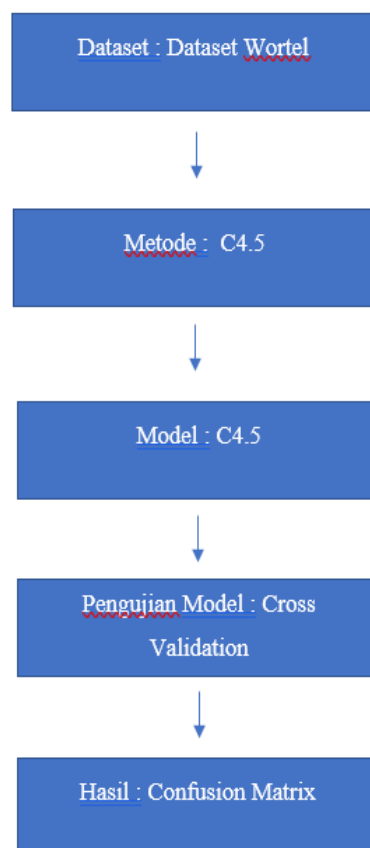
Gambar 3. 4 Hasil Pohon Keputusan atau Decission Tree

Dari gambar pohon keputusan atau *decision tree* pada gambar 3.5 diatas, didapat rule model dalam penentuan kualitas wortel yang dimana sebagai berikut:

1. Ukuran Wortel = Besar
2. | Kondisi Wortel = Beku: Tidak Bagus {Bagus=0, Tidak Bagus=2}
3. | Kondisi Wortel = Licin: Tidak Bagus {Bagus=1, Tidak Bagus=2}
4. | Kondisi Wortel = Normal: Bagus {Bagus=1, Tidak Bagus=1}
5. Ukuran Wortel = Normal: Bagus {Bagus=3, Tidak Bagus=0}

### 3.3.2 Alur Metode

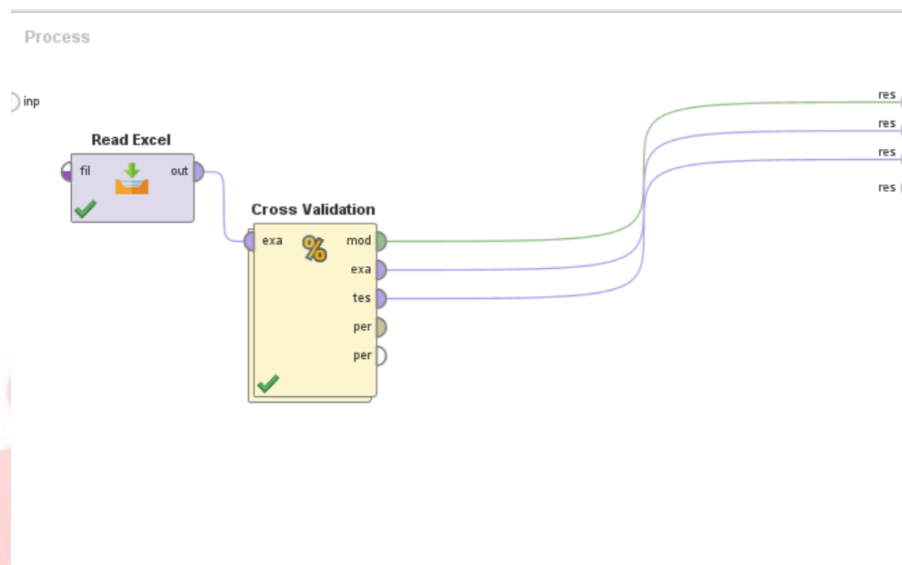
Alur Metode disini akan menggambarkan tahapan algoritma atau metode yang digunakan kedalam rapidminer, diantaranya adalah:



### Gambar 3. 5 Kontruksi Algoritma dan Metode

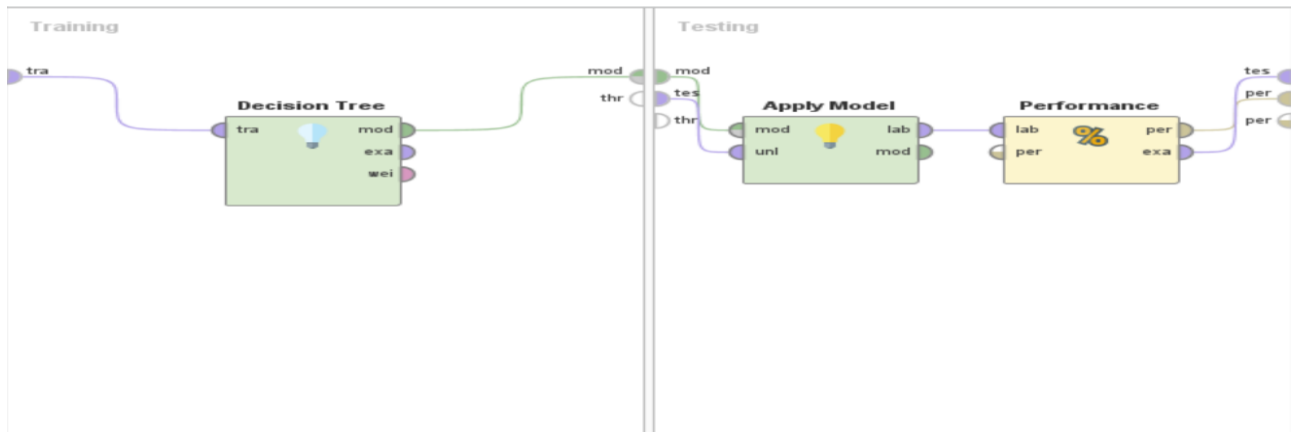
#### 3.3.3 Kontruksi Metode

Pada tahap ini, akan digambarkan model pengolahan data menggunakan metode C4.5 pada software rapid miner.



Gambar 3. 6 Proses Pengolahan Rapidminer

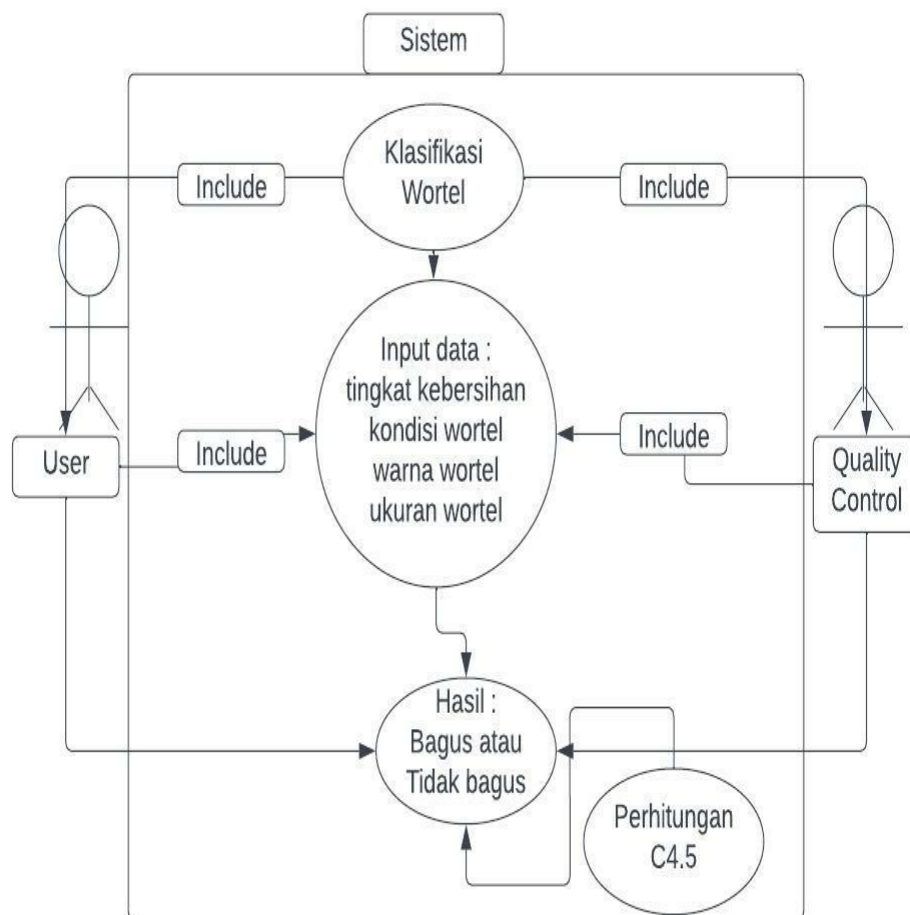
Pada gambar diatas adalah proses yang dimana tahapan awal adalah memasukan dataset kedalam *software* rapidminer yang dimana, data yang digunakan berbentuk excel yang selanjutnya memasukan fungsi *Cross Validation* yang dimana lalu disambungkan dengan data excel tersebut. Kemudian, masuk ke jendela *Cross Validation* tersebut yang dimana memasukan algoritma C4.5 atau *Decision Tree*, lalu selanjutnya memasukan fungsi *Apply Model* lalu *Perfromance Classification*.



Gambar 3. 7 Implementasi Algoritma C4.5 atau Decision Tree di Rapidminer

### 3.4 Perancangan UML

#### 3.4.1 Use Case

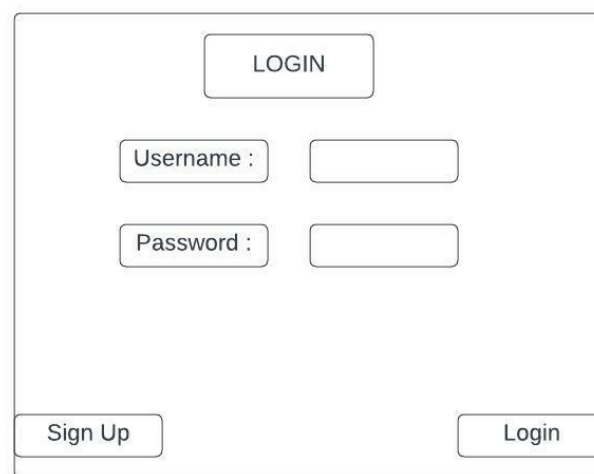


Gambar 3. 8 Use Case

### 3.5 Perancangan Layar, Menu dan Database

#### 3.5.1 Perancangan Halaman Login

Dibawah ini adalah gambaran dari perancangan halaman login yang dimana akan muncul ketika ingin menjalankan aplikasi:

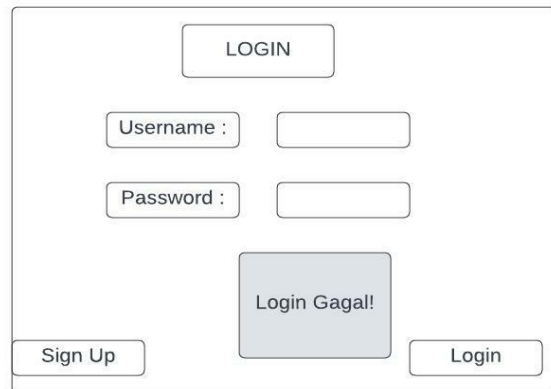


LOGIN	
Username :	<input type="text"/>
Password :	<input type="password"/>
Sign Up	Login

Gambar 3. 9 Perancangan Halaman Login

##### 3.5.1.1 Perancangan Halaman Login Gagal

Dibawah ini adalah gambaran dari perancangan halaman login gagal, yang dimana akan muncul ketika sudah memasukan username dan password tetapi tidak terhubung dengan database:

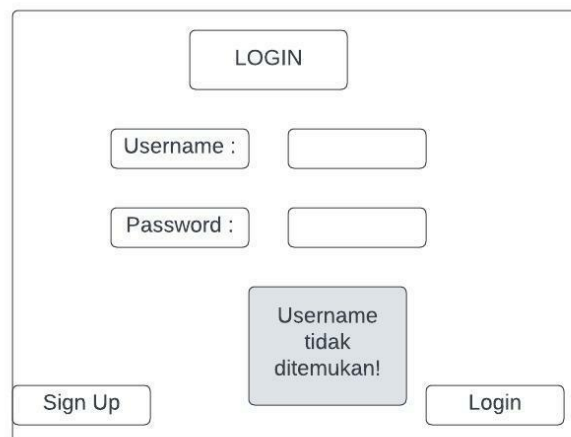


A screenshot of a login page. At the top center is a button labeled "LOGIN". Below it are two input fields: "Username :" and "Password :". In the center, there is a grey box with the text "Login Gagal!". At the bottom left is a button labeled "Sign Up", and at the bottom right is a button labeled "Login".

Gambar 3. 10 Perancangan Halaman Login Gagal

### 3.5.1.2 Perancangan Halaman Login Tidak Ditemukan

Dibawah ini adalah gambaran dari perancangan halaman login tidak ditemukan, yang dimana akan muncul ketika sudah memasukan username dan password tetapi belum terdaftar sehingga tidak ada di database:



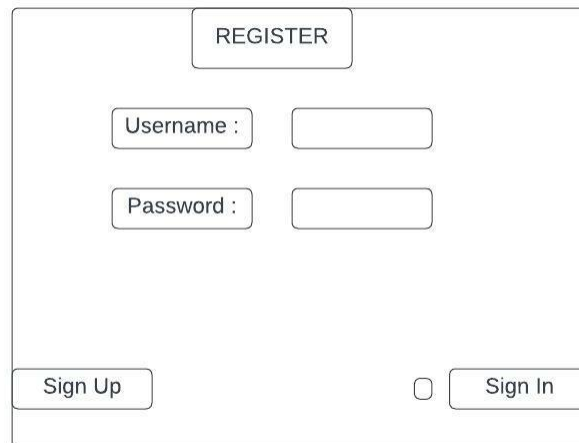
A screenshot of a login page. At the top center is a button labeled "LOGIN". Below it are two input fields: "Username :" and "Password :". In the center, there is a grey box with the text "Username tidak ditemukan!". At the bottom left is a button labeled "Sign Up", and at the bottom right is a button labeled "Login".

Gambar 3. 11 Perancangan Halaman Login Tidak Ditemukan



### 3.5.2 Perancangan Halaman Register

Dibawah ini adalah gambaran dari perancangan halaman register yang dimana akan muncul ketika ingin membuat akun untuk menjalankan aplikasi:

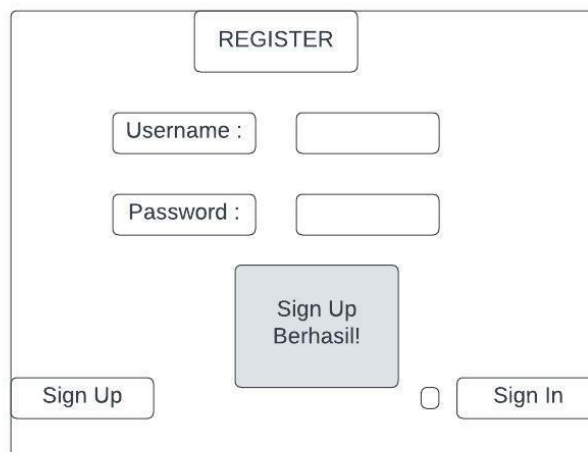


A wireframe of a registration form. At the top center is a button labeled "REGISTER". Below it are two input fields: "Username :" followed by a text box, and "Password :" followed by a text box. At the bottom left is a button labeled "Sign Up", and at the bottom right is a radio button followed by a button labeled "Sign In".

Gambar 3. 12 Perancangan Halaman Register

#### 3.5.2.1 Perancangan Halaman Register Berhasil

Dibawah ini adalah gambaran dari perancangan halaman register berhasil yang dimana akan muncul ketika berhasil membuat akun untuk menjalankan aplikasi:

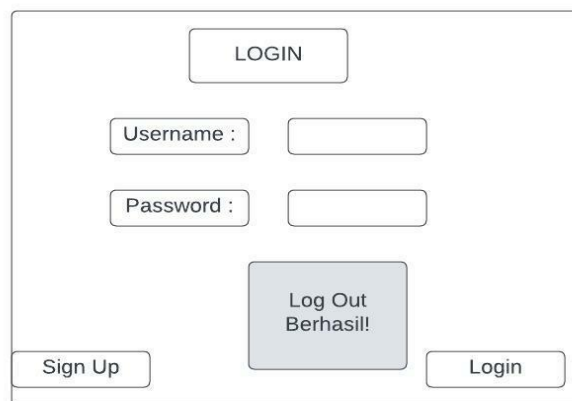


A wireframe of a registration form showing a successful registration message. At the top center is a button labeled "REGISTER". Below it are two input fields: "Username :" followed by a text box, and "Password :" followed by a text box. In the center, there is a grey box with the text "Sign Up Berhasil!". At the bottom left is a button labeled "Sign Up", and at the bottom right is a radio button followed by a button labeled "Sign In".

Gambar 3. 13 Perancangan Halaman Register Berhasil

### 3.5.3 Perancangan Halaman Log Out

Dibawah ini adalah gambaran dari perancangan halaman log out yang dimana akan muncul ketika ingin keluar dari aplikasi atau balik ke halaman log in:



The screenshot shows a web interface with the following elements:

- A "LOGIN" button at the top center.
- Two input fields: "Username :" and "Password :", each followed by a text input box.
- A central grey button with the text "Log Out Berhasil!".
- Two buttons at the bottom: "Sign Up" on the left and "Login" on the right.

Gambar 3. 14 Perancangan Halaman Log Out

### 3.5.4 Perancangan Halaman Menu

Dibawah ini adalah gambaran dari perancangan halaman menu yang dimana akan muncul ketika sudah berhasil login, yang berisi input data, tabel data dan c4.5:



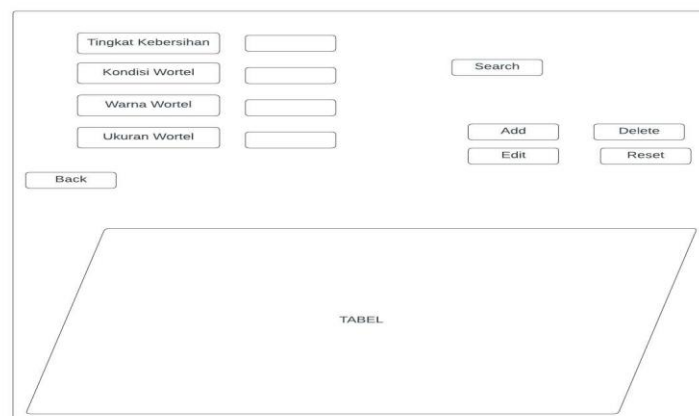
The screenshot shows a web interface with the following elements:

- A "SELAMAT DATANG" button at the top center.
- Two rows of buttons:
  - Row 1: "Input Data" and "Press"
  - Row 2: "Test Klasifikasi" and "Press"
- A "Log Out" button at the bottom right.

Gambar 3. 15 Perancangan Halaman Menu

### 3.5.5 Perancangan Halaman Input Data

Dibawah ini adalah gambaran dari perancangan halaman input data yang dimana akan muncul ketika memencet tombol press di sebelah input data pada halaman menu:



The image shows a web form for data entry. It features four input fields on the left, each with a label: 'Tingkat Kebersihan', 'Kondisi Wortel', 'Warna Wortel', and 'Ukuran Wortel'. To the right of these fields is a 'Search' button. Below the search button are four action buttons: 'Add', 'Delete', 'Edit', and 'Reset'. A 'Back' button is located at the bottom left of the form. The main content area is a large rectangle labeled 'TABEL'.

Gambar 3. 16 Perancangan Halaman Input Data

### 3.5.6 Perancangan Halaman Test Klasifikasi

Dibawah ini adalah gambaran dari perancangan halaman test klasifikasi yang dimana akan muncul ketika memencet tombol press di sebelah test klasifikasi pada halaman menu:

The image shows a web form for classification testing. It contains four input fields with labels: 'Tingkat Kebersihan', 'Kondisi Wortel', 'Warna Wortel', and 'Ukuran Wortel'. Below these fields are two buttons: 'Back' and 'Press'.

Gambar 3. 17 Perancangan Halaman Test Klasifikasi

### 3.5.7 Perancangan Database

Untuk menampung semua proses yang ada didalam aplikasi yang dibuat, dibutuhkan rancangan database untuk memaksimalkan proses tersebut. Yang dimana terdapat 2 tabel yang berbeda dalam sebuah database, untuk 2 proses yang berbeda. Diantaranya adalah:

#### 3.5.7.1 Tabel reg

Pada tabel ini, akan menampung semua username dan password yang dibutuhkan untuk masuk kedalam halaman menu.

Tabel 3. 4 Perancangan Tabel reg

Nama Field	Tipe Field	Peran
user	Varchar(50)	-
pass	Varchar(50)	-

Keterangan:

user = berisi username yang digunakan untuk masuk ke halaman menu.

pass = berisi password yang digunakan untuk masuk ke halaman menu.

### 3.5.7.2 Tabel input

Pada tabel ini, akan menampung semua hasil input data yang berisi nomor, kondisi wortel, tingkat kebersihan, warna wortel dan ukuran wortel:

Tabel 3. 5 Perancangan Tabel input

Nama Field	Tipe Field	Peran
nomor	Varchar(500)	<i>Primary key</i>
kondisi_wortel	Varchar(25)	-
tingkat_kebersihan	Varchar(25)	-
warna_wortel	Varchar(25)	-
ukuran_wortel	Varchar(25)	-

Keterangan:

nomor = berisi nomor urut dari setiap data yang di input.

kondisi\_wortel = berisi kondisi dari setiap wortel yang ada.

tingkat\_kebersihan = berisi tingkat kebersihan dari wortel yang ada.

warna\_wortel = berisi warna dari wortel yang ada.

ukuran\_wortel = berisi ukuran dari wortel yang ada.

