

**PENERAPAN *DATA MINING* UNTUK PREDIKSI STOK  
BARANG BERDASARKAN PENJUALAN BAHAN  
PERTANIAN MENGGUNAKAN  
ALGORITMA *C 5.0***

**SKRIPSI**

**SITI MAWADDAH  
NIM. 0702173167**



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA  
MEDAN  
2022 M / 1444 H**

**PENERAPAN *DATA MINING* UNTUK PREDIKSI STOK  
BARANG BERDASARKAN PENJUALAN BAHAN  
PERTANIAN MENGGUNAKAN  
ALGORITMA C 5.0**

**SKRIPSI**

*Diajukan untuk memenuhi syarat mencapai gelar sarjana komputer*

**SITI MAWADDAH  
NIM. 0702173167**



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA  
MEDAN  
2022 M / 1444 H**

## PERSETUJUAN SKRIPSI

Hal : Surat Persetujuan Skripsi

Lamp : -

Kepada Yth.

Dekan Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara Medan

*Assalamu'alaikum Wr. Wb.*

Setelah membaca, meneliti, memberikan petunjuk, dan mengoreksi serta mengadakan perbaikan, maka kami selaku pembimbing berpendapat bahwa skripsi saudara:

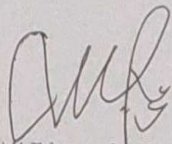
Nama : Siti Mawaddah  
Nomor Induk Mahasiswa : 0702173167  
Program Studi : Sistem Informasi  
Judul : Penerapan *Data Mining* Untuk Prediksi Stok Barang  
Berdasarkan Penjualan Bahan Pertanian  
Menggunakan Algoritma *C 5.0* pada CV. MITRA  
KARYA SEJATI

Dapat disetujui untuk segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Medan, 11 maret 2022 M

Komisi Pembimbing:

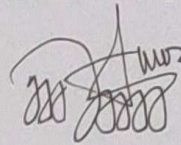
Dosen Pembimbing I



Ali Ikhwan, M.Kom

NIB.1100000109

Dosen Pembimbing II



Muhammad Dedi Irawan, M.Kom

NIP. 199001312019031019

## SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini:

Nama : Siti Mawaddah  
Nomor Induk Mahasiswa : 0702173167  
Program Studi : Sistem Informasi  
Judul : Penerapan *Data Mining* Untuk Prediksi Stok Barang  
Berdasarkan Penjualan Bahan Pertanian  
Menggunakan Algoritma *C 5.0* pada CV. MITRA  
KARYA SEJATI

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya. Apabila dikemudian hari ditemukan plagiat dalam skripsi ini maka saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi lainnya sesuai peraturan yang berlaku.

Medan, 11 Maret 2022



Siti Mawaddah

NIM. 0702173167



**KEMENTERIAN AGAMA REPUBLIK INDONESIA**  
**UNIVERSITAS ISLAM NEGERI SUMATERA UTARA**  
**MEDAN FAKULTAS SAINS DAN TEKNOLOGI**  
Jl. Lap. Golf, Desa Durian Jangak, Kec. Pancur Batu 20353  
Telp. (+6261) 4536090, Fax. (+6261) 6615683  
Url: <http://saintek.uinsu.ac.id>, E-mail: [saintek@uinsu.ac.id](mailto:saintek@uinsu.ac.id)

**PENGESAHAN SKRIPSI**

Nomor : B.173/ST/ST.V.2/PP.01.1/08/2022

Judul : Penerapan *Data Mining* Untuk Prediksi Stok Barang  
Berdasarkan Penjualan Bahan Pertanian Menggunakan  
Algoritma *C 5.0* pada CV. MITRA KARYA SEJATI  
Nama : Siti Mawaddah  
Nomor Induk Mahasiswa : 0702173167  
Program Studi : Sistem Informasi  
Fakultas : Sains dan Teknologi

Telah dipertahankan di hadapan Dewan Penguji Skripsi Program Studi Sistem Informasi  
Fakultas Sains dan Teknologi UIN Sumatera Utara Medan Dan Dinyatakan **LULUS**.  
Pada Hari/Tanggal : Jumat, 18 Maret 2022  
Ruang : Ruang Meeting FST Kampus IV Tuntungan

**TIM UJIAN MUNAQASYAH**

**NETU**


(Samsudin, S.T., M.Kom)  
NIP. 197612272011011002

Dewan Penguji,

Penguji I

Aninda Muliani Harahap, M.Kom  
NIP. 198611292019032009


Penguji III

  
Ali Ikhwan, M.Kom.  
NIB. 1100000109

Penguji II

  
Samsudin, S.T., M.Kom.  
NIP. 197612272011011002

Penguji IV

  
(Muhammad Dedi Irawan, M.Kom)  
NIP. 199001312019031019

Mengesahkan  
Dekan Fakultas Sains dan Teknologi  
UIN Sumatera Utara Medan



(Muhammad Dedi Irawan, MA)  
NIP. 199001312019031019



## ABSTRAK

CV. Mitra Karya Sejati merupakan sebuah usaha dibidang perdagangan bahan pertanian seperti pupuk multiguna, pestisida, serta benih yang berdiri pada tahun 1998 oleh bapak Tardi Sutanto. CV. Mitra Karya Sejati sering mengalami kekurangan dikarenakan dengan meningkatnya permintaan barang dari konsumen dan beberapa barang juga ada yang tidak terjual, sehingga ada kesulitan untuk prediksi stok barang pada gudang penyimpanan berdasarkan data penjualan yang selama ini data penjualan hanya dimanfaatkan sebagai data untuk melaporkan hasil penjualan setiap bulannya kemudian tidak akan digunakan lagi. Perusahaan tersebut memerlukan sistem prediksi penjualan untuk membantu perencanaan stok barang kedepannya. Implementasi *C 5.0* pada *Data Mining* bekerja untuk memprediksi stok barang yang berdasarkan penjualan bahan pertanian. Aplikasi *Data Mining* dirancang dengan menggunakan fitur dan perangkat bahasa coding PHP dan *Database MySQLI* sehingga mengurangi kerugian dalam perencanaan stok barang. Melalui sistem yang diusulkan, dapat membantu perusahaan memprediksi stok barang yang akan disediakan untuk meminimalisir kerugian.

**Kata Kunci:** *C 5.0*, Prediksi, *Data Mining*

## ABSTRACT

Founded in 1998 by Mr. Tardi Sutanto, CV. Mitra Karya Sejati is a company specializing in the distribution of agricultural inputs like multifunctional fertilizers, herbicides, and seeds. Due to rising consumer demand and unsold inventory, CV. Mitra Karya Sejati frequently faces stock-outs, making it difficult to predict warehouse inventory levels using sales data, which is currently only used for reporting monthly sales results. Henceforth it shall never be re-employed. In order to better plan for future stock items, the organization requires a sales forecast system. Companies can better anticipate stock needs based on sales of agricultural materials with the use of C 5.0 on Data Mining. To cut down on waste in stock management, a PHP and MySQL-based Data Mining application was developed. The proposed approach can aid businesses in anticipating the stock of items supplied in order to reduce losses.

**Keywords: C 5.0, Prediction, Data Mining**

## KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillahirabbil'alamin, segala puji dan syukur kehadiran Allah SWT, karena atas segala rahmat dan karunia-Nya, penulis dapat dimudahkan dan dilancarkan dalam menyelesaikan skripsi yang berjudul “Penerapan *Data mining* Untuk Prediksi Stok Barang Berdasarkan Penjualan Bahan Pertanian Menggunakan Algoritma *C 5.0* pada CV. MITRA KARYA SEJATI”. Penulis menyadari bahwa skripsi ini masih memiliki kekurangan yang nantinya akan diperbaiki berdasarkan kritik dan saran yang diberikan. Ucapan terima kasih penulis sampaikan kepada pihak-pihak yang telah banyak membantu dan memberi dukungan dalam menyelesaikan proposal skripsi ini yaitu kepada:

1. Ibu Prof. Dr. Nurhayati, M.Ag., selaku Rektor Universitas Islam Negeri Sumatera Utara
2. Bapak Dr. Mhd. Syahnan, M.A selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara.
3. Bapak Samsudin, S.T., M.Kom selaku Ketua Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara.
4. Bapak Suendri, M.Kom selaku Sekretaris Program Studi Sistem Informasi Fakultas Sains dan Teknologi pada Universitas Islam Negeri Sumatera Utara
5. Bapak Ali Ikhwan, M.Kom selaku dosen pembimbing akademik serta dosen pembimbing I penulis yang telah membantu dalam memberikan arahan dan masukan kepada penulis selama penulisan skripsi ini.
6. Bapak Muhammad Dedi Irawan, M.Kom selaku dosen pembimbing II penulis, yang telah membantu dalam memberikan arahan dan masukan kepada penulis selama penulisan skripsi ini.
7. Seluruh dosen Program Studi Sistem Informasi yang telah banyak memberikan arahan dan masukan kepada penulis.
8. Ibunda saya Hanimah dan Ayah saya Rismatahi Tanjung, serta ke empat abang laki-laki saya, terimakasih atas segala dukungan, kasih sayang, dan doa yang selalu tercurah kepada saya.



9. Teman baik Samsul Bahri Siagian yang telah menemani serta memberikan dukungan dan motivasi kepada penulis, sehingga penulis tetap semangat dalam menyelesaikan skripsi ini.

Semoga Allah SWT selalu memberikan nikmat dan karunia kepada kita semua, dan semoga skripsi ini nantinya dapat bermanfaat kepada Universitas Islam Negeri Sumatera Utara, CV.Mitra Karya Sejati, pembaca, serta bagi banyak masyarakat luas. Akhir kata penulis mengucapkan terima kasih.

Medan, 11 Maret 2022

Penulis,

Siti Mawaddah  
NIM. 0702173167



## DAFTAR ISI

<b>PERSETUJUAN SKRIPSI</b> .....	<b>i</b>
<b>SURAT PERNYATAAN KEASLIAN SKRIPSI</b> .....	<b>ii</b>
<b>PEGESAHAN SKRIPSI</b> .....	<b>iii</b>
<b>ABSTRAK</b> .....	<b>iv</b>
<b>ABSTRACT</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>vi</b>
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Penerapan .....	6
2.2 <i>Data mining</i> .....	6
2.3 <i>Knowledge Discovery Database (KDD)</i> .....	7
2.4 Teknik <i>Data mining</i> .....	8
2.4.1 Klasifikasi.....	8
2.5 Prediksi.....	9
2.5.1 <i>Decision Tree</i> .....	9
2.6 Algoritma <i>C 5.0</i> .....	10
2.7 Stok Barang .....	12
2.7.1 Tujuan Penjualan .....	12
2.8 Produk.....	13
2.8.1 Pupuk.....	13
2.8.2 Herbisida.....	14
2.8.3 Insektisida.....	14
2.9 CV. Mitra Karya Sejati.....	15
2.10 Data.....	15
2.11 Basis Data.....	16

2.11.1	Peran Basis Data .....	16
2.12	PHP .....	17
2.13	MySQL .....	19
2.14	Unified Modelling Language (UML) .....	22
2.15	Website .....	30
2.16	Penelitian Terdahulu .....	30

### **BAB III METODE PENELITIAN**

3.1	Tempat dan Waktu Penelitian .....	33
3.1.1	Tempat Penelitian .....	33
3.1.2	Waktu Penelitian .....	33
3.2	Kebutuhan Sistem .....	36
3.2.1	Perangkat Keras .....	36
3.2.2	Perangkat Lunak .....	36
3.3	Cara Kerja .....	36
3.3.1	Metode Pengumpulan Data .....	36
3.4	Jenis Data .....	37
3.5	Metode Pengembangan Sistem .....	38
3.6	Algoritma C 5.0 .....	40
3.7	Kerangka Berfikir .....	41
3.7.1	Deskripsi Kerangka Berfikir .....	41

### **BAB IV HASIL DAN PEMBAHASAN**

4.1	Requirements Definition .....	43
4.1.1	Sejarah Mitra Karya Sejati .....	43
4.1.2	Visi dan Misi Perusahaan .....	48
4.1.3	Struktur Organisasi .....	49
4.1.4	Pembagian Tugas dan Tanggung Jawab .....	50
4.1.5	Analisis Sistem .....	52
4.1.6	Tahapan Metode C 5.0 .....	54
4.2	Design .....	64
4.2.1	Desain Proses .....	64
4.2.2	Desain Database .....	72
4.2.3	Desain Interface .....	76
4.3	Implementasi Aplikasi .....	81
4.4	Hasil Pengujian Sistem .....	89
4.4.1	Pengujian Validitas .....	92

**BAB V PENUTUP**

5.1 Kesimpulan..... 98

5.2 Saran..... 98

**DAFTAR PUSTAKA..... 99**

**LAMPIRAN.....101**



## DAFTAR TABEL

TABEL	JUDUL TABEL	HALAMAN
2. 1	Simbol-Simbol <i>Use Case Diagram</i> .....	23
2. 2	Contoh <i>Use Case Diagram</i> .....	24
2. 3	Simbol-simbol <i>Class Diagram</i> .....	24
2. 4	Simbol <i>Activity Diagram</i> .....	26
2. 5	Simbol-Simbol <i>Sequence Diagram</i> .....	28
2. 6	Penelitian Terdahulu .....	30
3. 1	Waktu Penelitian .....	34
4. 1	Kategori Pupuk.....	44
4. 2	Kategori Racun Akar.....	44
4. 3	Kategori Racun Bakar .....	46
4. 4	Kategori Insektisida.....	47
4. 6	Kategori Herbisida .....	48
4. 7	<i>Data mining</i> .....	55
4. 8	<i>Gain Ratio</i> tertinggi.....	62
4. 9	<i>Confusion Matrix</i> .....	64
4. 10	Tabel Admin.....	72
4. 11	Tabel Atribut .....	72
4. 12	Tabel Nilai.....	74
4. 13	Tabel <i>Dataset</i> .....	74
4. 14	Tabel Evaluasi.....	74
4. 15	Atribut Nama Pupuk.....	92
4. 16	Atribut Total.....	93
4. 17	Atribut Satuan .....	93
4. 18	Atribut Jenis Penjualan.....	93
4. 19	Atribut Bulan.....	93
4. 20	Atribut Tahun .....	94



## DAFTAR GAMBAR

GAMBAR	JUDUL GAMBAR	HALAMAN
2. 1	Proses pada <i>Knowledge Discovery Database</i> .....	8
2. 2	Logo Mitra Karya Sejati.....	15
2. 3	Logo Bahasa Pemograman PHP .....	17
2. 4	Logo <i>MySQL</i> .....	20
2. 5	Contoh <i>Class Diagram</i> .....	25
2. 6	Contoh <i>Activity Diagram</i> .....	27
2. 7	Contoh <i>Sequence Diagram</i> .....	29
3. 1	CV. Mitra Karya Sejati	33
3. 2	Tahapan <i>Waterfall</i> .....	38
3. 3	Tahapan Metode Algoritma <i>C 5.0</i> .....	40
3. 4	Kerangka Berfikir.....	41
4. 1	Struktur Organisasi CV. Mitra Karya Sejati .....	49
4. 2	Diagram Aliran Sistem Lama.....	53
4. 3	Diagram Aliran Sistem Usulan .....	54
4. 4	Pohon Keputusan.....	63
4. 5	<i>Usecase Diagram Data mining</i> Prediksi Stok Barang .....	65
4. 6	<i>Class Diagram</i> .....	66
4. 7	<i>Activity Diagram Login Admin</i> .....	67
4. 8	<i>Activity Diagram Atribut</i> .....	68
4. 9	<i>Activity Diagram Nilai Atribut</i> .....	69
4. 10	<i>Activity Diagram Dataset</i> .....	70
4. 11	<i>Sequence Diagram Login</i> .....	71
4. 12	<i>Sequence Diagram Atribut</i> .....	71
4. 13	<i>Sequence Diagram Nilai Atribut</i> .....	72
4. 14	Halaman Utama.....	76
4. 15	Halaman <i>Login</i> .....	77
4. 16	Halaman <i>Dashboard</i> .....	77
4. 17	Halaman Atribut.....	78

4. 18	Halaman Nilai Atribut.....	78
4. 19	Halaman <i>Dataset</i> .....	79
4. 20	Halaman Perhitungan .....	80
4. 21	Halaman Pohon Keputusan .....	80
4. 22	Tampilan Utama .....	81
4. 23	Tampilan <i>Login</i> .....	82
4. 24	Tampilan <i>Dashboard</i> .....	82
4. 25	Tampilan Atribut.....	83
4. 26	Tampilan Tambah Atribut.....	83
4. 27	Tampilan Ubah Atribut .....	84
4. 28	Tampilan Tambah Nilai Atribut.....	84
4. 29	Tampilan Ubah Alternatif .....	85
4. 30	Tampilan Nilai Atribut.....	85
4. 31	Tampilan Tambah <i>Dataset</i> .....	86
4. 32	Tampilan Ubah <i>Dataset</i> .....	87
4. 33	Tampilan <i>Dataset</i> .....	88
4. 34	Tampilan Hasil Perhitungan.....	88
4. 35	Tampilan Pohon Keputusan .....	89
4. 36	Hasil Akurasi.....	89
4. 37	Atribut Pupuk.....	95
4. 38	Atribut Total.....	95
4. 39	Atribut Satuan dan Penjualan.....	96
4. 40	Atribut Bulan.....	96
4. 41	Atribut Tahun.....	97

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Penelitian ini dilatar belakangi oleh ayat suci Al-Qur'an pada surat Yusuf ayat 47-49:

قَالَ تَزْرَعُونَ سَبْعَ سِنِينَ دَأْبًا فَمَا حَصَدْتُمْ فَذُرُوهُ فِي سُبُلٍ إِلَّا قَلِيلًا مِّمَّا تَأْكُلُونَ ٤٤

ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ سَنَةٌ يُكَلِّنُ بِهَا النَّاسُ كِبَاحَهُمْ فَسِرْقَاةٌ فِيهَا وَبُهَادٌ فِيهَا وَكُلٌّ فِيهَا خَائِبُونَ ٤٥

ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ عَامٌ فِيهِ يُغَاثُ النَّاسُ وَفِيهِ يُغْصَرُونَ ٤٦

menunjukkan bahwa Allah memerintahkan manusia untuk menanam dan menuai hasil dari tanaman serta menyimpan sebagiannya untuk masa depan. Dalam konteks perusahaan, persediaan menjadi aset yang penting karena merupakan kekayaan yang dimilikinya. Hal ini karena ada tiga alasan penting mengapa persediaan dibutuhkan dalam perusahaan atau organisasi, yaitu:

1. Ketidakpastiaan permintaan.
2. Ketidakpastian dari pasokan atau supplier.
3. Ketidakpastian tenggang waktu pemesanan

Dalam era perkembangan teknologi yang pesat saat ini, masyarakat Indonesia dapat dengan mudah mengakses berbagai macam informasi. Hampir seluruh lapisan masyarakat, mulai dari anak-anak, remaja, hingga dewasa, terlibat dalam penggunaan teknologi setiap hari. Teknologi kini dianggap mampu menggantikan tenaga manusia. Oleh karena itu, penyampaian informasi yang efektif dan efisien menjadi semakin diperlukan mengingat pesatnya perkembangan teknologi modern. Penyampaian informasi pasti telah beralih dari sistem kuno entri data manual ke sistem komputerisasi yang lebih maju sejak lama.

CV. Mitra Karya Sejati terletak di Jl. Menggala Sakti Km.24, Kec. Tanah Putih, Kab. Rokan Hilir, Riau yang merupakan perusahaan bergerak dibidang

Perdagangan yang menyediakan berbagai macam alat dan bahan pertanian dengan harga terjangkau. CV. Mitra Karya Sejati saat ini juga sudah menggunakan sistem informasi penjualan untuk mendukung kinerja pihak perusahaan pada layanan informasi penjualan. Ketersediaan produk pertanian dengan harga yang terjangkau menjadi tantangan tersendiri untuk menjaga kesetabilan stok barang pada gudang penyimpanan. Meski terlihat mudah, mengelola persediaan bukanlah hal yang dapat diremehkan, karena jika persediaan terlalu banyak risiko kerusakan barang akan meningkat, begitupun sebaliknya jika persediaan terlalu sedikit risiko kekurangan persediaan akan semakin besar dan dapat menunda keuntungan serta dapat mengecewakan konsumen.

Dilihat dari beberapa bulan belakangan ini penjualan bahan pertanian mengalami kenaikan sehingga ketersediaan barang pada gudang penyimpanan sering mengalami kekurangan dikarenakan dengan meningkatnya permintaan barang dari konsumen dan beberapa barang juga ada yang tidak terjual. Pada bulan Juni hingga September 2021, perusahaan mengalami kekurangan stok pupuk Meroke dan kerusakan pada pupuk CU Daun yang menyebabkan kerugian. Hal ini terjadi karena terlalu banyaknya pemesanan pada pupuk tertentu dan sulitnya melakukan prediksi stok barang pada gudang penyimpanan. Prediksi ini sangat berguna untuk menentukan berapa banyak produk yang akan disediakan pada bulan selanjutnya. Permasalahan yang sering dihadapi oleh perusahaan adalah bagaimana tingkat akurasi dalam memprediksi atau meramalkan stok barang dimasa mendatang berdasarkan data penjualan sebelumnya. Prediksi tersebut sangat berpengaruh untuk menentukan produk mana yang laris dan tidak laris. Penelitian ini diambil dari data penjualan pada CV. Mitra Karya Sejati yang selama ini data penjualan hanya dimanfaatkan sebagai data untuk melaporkan hasil penjualan setiap bulannya kemudian tidak akan digunakan lagi.

*Data mining* merupakan teknik pengolahan data yang digunakan untuk menggali informasi yang bermanfaat dari data yang besar. Informasi yang didapatkan dari teknik data mining tersebut dapat digunakan untuk memprediksi, pengklasifikasian, dan memperkirakan. Salah satu metode algoritma klasifikasi Data mining adalah algoritma C 5.0 yang diterapkan pada teknik *Decision Tree*. Algoritma C 5.0 sendiri merupakan penyempurnaan dari algoritma ID3 dan C 4.5.

Sebelum melakukan studi penelitian, penulis mencari beberapa referensi dari penelitian terbaru yang telah dibuat oleh Dwi Marisa Efendi dan Pakarti Riswanto dalam karya ilmiah yang berjudul “Penerapan *Data mining* Untuk Prediksi Penjualan Pupuk Menggunakan Metode Algoritma Apriori” yang terbit pada jurnal JIK (Jurnal Informasi dan Komputer) Vol. 9 no. 1, Tahun 2021. Penelitian ini menghasilkan perhitungan dengan menggunakan metode algoritma apriori dalam menentukan banyaknya jumlah pupuk yang terjual pada toko Fadilah Semuli Jaya (Ninla Elmawati Falabiba et al., 2021). Pada penelitian (Yusuf, 2018) telah melakukan penelitian tentang perbandingan performansi algoritma *Decision Tree C 5.0*, *CART* dan *CHAID* dengan studi kasus prediksi status resiko kredit di Bank *X* mendapatkan hasil bahwa rata-rata tingkat akurasi untuk metode algoritma *C 5.0* sebesar 87,72%, *CART* sebesar 87,27% dan *CHAID* sebesar 87,15%.

Berdasarkan fenomena dan permasalahan di atas, Penelitian ini bertujuan untuk mempermudah bagian penyedia stok barang pada CV. Mitra Karya Sejati dalam melakukan perencanaan stok barang serta mempermudah pihak perusahaan untuk mengetahui tentang produk pupuk yang harus disediakan dan banyak diminati oleh konsumen. Maka penulis mengangkat penelitian dengan judul **“PENERAPAN *DATA MINING* UNTUK PREDIKSI STOK BARANG BERDASARKAN PENJUALAN BAHAN PERTANIAN MENGGUNAKAN ALGORITMA *C 5.0*”**.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana menerapkan algoritma *C 5.0* untuk prediksi stok barang berdasarkan penjualan bahan pertanian pada CV. Mitra Karya Sejati?
2. Bagaimana membangun aplikasi *data mining* prediksi stok barang pada CV. Mitra Karya Sejati?



### 1.3 Batasan Masalah

Pada penelitian ini dilakukan batasan masalah agar lebih fokus pada masalah yang diteliti, adapun batasan masalah sebagai berikut:

1. Berbasis web
2. Aplikasi *Data mining* dibangun menggunakan bahasa coding PHP Native dan *database MySQL*.
3. Menggunakan metode algoritma *C 5.0*
4. Penelitian ini hanya membahas tentang prediksi stok barang dari data transaksi penjualan tahun 2018-2021.
5. Pada penelitian ini penulis menggunakan dataset penjualan yang dikelola ada 5 kategori, yaitu : kategori pupuk, kategori racun bakar, kategori racun akar, herbisida dan insektisida.
6. Pada perhitungan manual diambil dari kategori pupuk.

### 1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma *C 5.0* untuk prediksi stok barang berdasarkan penjualan pada CV. Mitra Karya Sejati.
2. Membangun aplikasi *data mining* prediksi stok barang pada CV. Mitra Karya Sejati.

### 1.5 Manfaat Penelitian

Manfaat yang diharapkan pada penelitian ini adalah sebagai berikut:

1. Bagi Peneliti
  - 1) Mengembangkan ilmu pengetahuan dan kemampuan yang telah diberikan oleh prodi sistem informasi serta dapat memanfaatkan referensi sebelumnya.
  - 2) Menjadikan peneliti paham tentang bagaimana pengembangan *Data mining*.
  - 3) Dengan adanya penelitian ini untuk melengkapi syarat kelulusan studi sistem informasi setara S1.

2. Bagi Program Studi

- 1) Mengetahui kemampuan mahasiswa dalam menerapkan ilmunya selama masa perkuliahan.
- 2) Memberikan referensi bagi peneliti selanjutnya.

3. Bagi Perusahaan

- 1) Membantu pihak perusahaan mengetahui peningkatan penjualan perbulannya.
- 2) Membantu pihak perusahaan untuk perencanaan stok barang pada gudang penyimpanan.
- 3) Meminimalisir kerugian akibat salah penyetokan produk.



## BAB II TINJAUAN PUSTAKA

### 2.1 Penerapan

Menurut (Ali Ikhwan, 2019), penerapan adalah proses menentukan ide, konsep kebijakan, atau inovasi menjadi tindakan praktis yang dapat membawa dampak positif pada perubahan pengetahuan, keterampilan, nilai, dan sikap.

Aktivitas, tindakan, tindakan, atau mekanisme sistem semuanya terkait dengan implementasi. Implementasi lebih dari sekedar mengambil tindakan; itu adalah upaya yang disengaja dan terencana dengan tujuan akhir yang spesifik. Implementasi adalah kebutuhan untuk kegiatan yang saling berinteraksi untuk mencapai tujuan dan tindakan, dan itu membutuhkan jaringan tata kelola bikameral yang kuat (Triase et al., 2020).

Dapat diartikan penerapan adalah suatu kegiatan terencana sehingga memberi dampak yang lebih baik untuk mencapai tujuan yang telah dirumuskan.

### 2.2 *Data mining*

Analisis otomatis kumpulan data besar atau rumit dengan tujuan mengidentifikasi tren atau kecenderungan signifikan yang biasanya diabaikan dalam kehidupan sehari-hari dikenal sebagai *data mining*. *Data mining* studi tentang memeriksa kumpulan data untuk menemukan asosiasi yang tidak terduga dan meringkas data dalam bentuk yang jelas bagi pemilik data (Ali Ikhwan et al., 2020).

Memproses volume data yang sangat besar untuk mendapatkan pengetahuan atau informasi yang membutuhkan suatu metode dikenal sebagai *data mining*. *Data mining* harus berperan dalam proses mengubah data mentah menjadi informasi, pola, dan keahlian yang berharga. *Data mining* umumnya terdiri dari lima peran: perkiraan, prediksi, klasifikasi, pengelompokan, dan asosiasi (Suendri et al., 2020).

Data merupakan gabungan fakta yang tersimpan, *Mining* merupakan suatu proses penambangan (Ali Ikhwan et al., 2021). *Data mining* adalah proses mencari pola yang belum ditemukan dalam sekumpulan data yang disimpan

dalam database, gudang data, atau jenis media penyimpanan informasi lainnya. *Data mining* berupaya mengungkap pengetahuan yang masih tersembunyi dalam kumpulan data. Untuk menemukan, menyelidiki, atau menambang pengetahuan dari fakta atau informasi sendiri.

### 2.3 *Knowledge Discovery Database (KDD)*

Teknik *data mining* disebut dengan nama *Knowledge Discovery Database (KDD)*. *Database Discovery Knowledge (KDD)* adalah istilah penerapan metode ilmiah pada *data mining* (Irawan, 2019). Dalam konteks ini, data mining merupakan langkah dalam proses *Knowledge Discovery Database (KDD)*. Prosedur *Knowledge Discovery Database (KDD)* dipecah menjadi tahapan berikut:

1. Informasi Seleksi (*selection*)

Penting untuk mengumpulkan data (seleksi) dari sekumpulan data operasional sebelum memulai tahap *Knowledge Discovery Database (KDD)* dari ekstraksi informasi. Basis data operasional disimpan dalam file yang berbeda dari data pemilihan *data mining*.

2. Persiapan Data (Pembersihan/*Preprocessing*)

Sebagai bagian dari prosedur pra-pemrosesan, duplikasi data dihilangkan, data yang tidak konsisten diperiksa, dan masalah data, seperti kesalahan ketik, diperbaiki. Proses “memperkaya” data yang sudah ada dengan lebih banyak data atau informasi, seperti data atau informasi eksternal, yang signifikan dan diperlukan untuk KDD juga dilakukan.

3. Transformasi

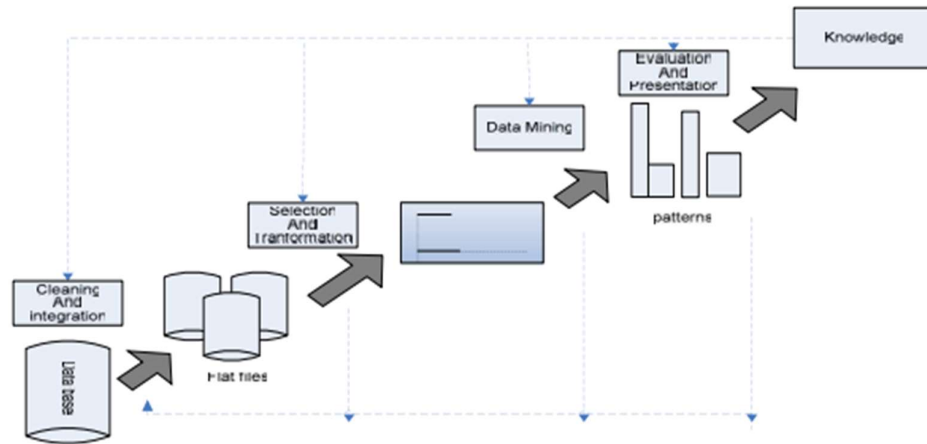
Data yang sudah siap untuk proses *data mining* atau belum memiliki entitas tertentu berupa data yang valid diubah pada langkah ini.

4. Ekstraksi informasi dari data

Dalam proses ini, teknik atau algoritma pencarian informasi diterapkan.

5. Interpretasi / Penilaian

6. Fase terakhir ini melibatkan pembuatan output yang mudah dipahami dan didasarkan pada pola pengetahuan yang ditemukan selama proses penambangan data (A Ikhwan et al., 2020).



**Gambar 2.1** Proses pada *Knowledge Discovery Database* (Ali Ikhwan, 2019)

## 2.4 Teknik *Data mining*

Berikut ini adalah beberapa metode dan sifat *data mining*:

1. Clustering adalah proses membagi kumpulan data menjadi beberapa sub-jaringan atau grup sehingga setiap anggota grup memiliki kumpulan karakteristik yang sama dengan tingkat kesamaan yang rendah disebut “*unsupervised learning*” juga.
2. Dengan mengadopsi model ketergantungan linier atau nonlinier, regresi adalah proses pendugaan nilai suatu variabel kontinu berdasarkan nilai variabel lain.
3. Klasifikasi, yang sering disebut dengan “*supervised learning*”, adalah proses penempatan record data baru ke dalam salah satu kategori (kelas) yang telah ditetapkan.
4. Kaidah Asosiasi (*association rule*) mengidentifikasi sekelompok atribut yang biasanya terjadi bersama (*co-occur*) dan menghasilkan sejumlah aturan dari kelompok ini (Nababan et al., 2018).

### 2.4.1 Klasifikasi

Teknik *data mining* yang disebut klasifikasi adalah suatu proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan karakteristik yang dimilikinya. Salah satu teknik klasifikasi yang populer adalah *Decision Tree* (Triase et al., 2020).



Teknik ini digunakan pada data yang bersifat terawasi atau *supervised*. Tujuan dari klasifikasi adalah untuk mengklasifikasikan data atau objek baru ke dalam kelas tertentu berdasarkan karakteristik yang mirip dengan data yang sudah diketahui kelasnya (Syaliman et al., 2017).

Dari pengertian diatas, klasifikasi merupakan teknik dalam *Data mining* yang digunakan untuk mengelompokkan data berdasarkan keterikatan data terhadap data sampel, sehingga dapat disusun secara sistematis dalam kelompok atau golongan sesuai dengan ketentuan yang ditetapkan dan dapat diterapkan dalam berbagai aspek kehidupan.

## **2.5 Prediksi**

Prediksi melibatkan usaha untuk memperkirakan kejadian di masa depan dengan memanfaatkan informasi yang relevan dari masa lalu melalui metode ilmiah. Tujuannya adalah untuk mendapatkan informasi tentang kejadian yang paling mungkin terjadi di masa depan (Ningsih et al., 2021).

Model pohon keputusan merupakan salah satu cara yang sangat berguna untuk melakukan prediksi. Dengan memanfaatkan model pohon yang dibangun dari data historis, dapat dengan mudah memprediksi hasil untuk catatan di masa depan.

### **2.5.1 Decision Tree**

Pendekatan *data mining* yang dikenal dengan metode pohon keputusan membagi kumpulan data yang sangat besar menjadi pohon keputusan yang mencerminkan aturan. Pendekatan ini dapat digunakan untuk menyelidiki data dan mengungkap hubungan yang belum ditemukan antara berbagai faktor input potensial dan variabel target. Data pohon keputusan diberikan sebagai tabel dengan catatan dan karakteristik. Karakteristik yang digunakan sebagai kriteria untuk membuat pohon dideskripsikan melalui atributnya (Triase et al., 2020).

Metode pohon keputusan digunakan secara umum untuk membangun sistem klasifikasi dan mengembangkan algoritma prediksi untuk variabel target. Metode ini mengklasifikasikan populasi ke dalam segmen seperti cabang yang membangun pohon terbalik dengan simpul akar, simpul internal, dan simpul daun.

Algoritma ini adalah non-parametrik dan dapat menangani kumpulan data yang besar dan rumit tanpa memaksa struktur parametrik yang rumit. Saat sampel yang cukup besar, data studi dapat dibagi menjadi set data pelatihan dan validasi. Set data pelatihan digunakan untuk membangun model pohon keputusan, dan set data validasi digunakan untuk menentukan ukuran pohon yang sesuai untuk mencapai model akhir yang optimal (Ali Ikhwan et al., 2021).

## 2.6 Algoritma C 5.0

Algoritma C 5.0 merupakan salah satu teknik Data mining yang khususnya digunakan dalam pembangunan pohon keputusan. Metode sebelumnya, ID3 dan C 4.5, yang dibuat oleh J. Ross Quinlan pada tahun 1987, diperbaiki dengan algoritma C 5.0. *Information Gain* digunakan pada algoritma C 5.0 untuk memilih karakteristik mana yang harus diproses. Induk untuk simpul berikut akan ditentukan oleh properti dengan Penguatan Informasi terbesar (Darmawan et al., 2018).

Proses pembangunan *tree* pada Algoritma C 5.0 serupa dengan Algoritma C 4.5 dan melibatkan perhitungan kemunculan kejadian, *Entropy*, dan *Information Gain*. Namun, Algoritma C 5.0 memiliki perbedaan dalam hal perhitungan *Gain Ratio*, yang dilakukan setelah perhitungan *Information Gain* dan *Entropy*. *Gain Ratio* dihitung dengan menggunakan nilai *Information Gain* dan *Entropy* yang telah diperoleh sebelumnya. Hal ini memungkinkan Algoritma C 5.0 untuk memberikan hasil yang lebih akurat dan efisien dalam membangun *tree* (Putri et al., 2016).

Berikut adalah rumus *Entropy* dan *Gain*:

### 1) *Entropy*

$$Entropy(S) = \sum_{i=1}^m p_i \log_2(p_i) \dots \dots \dots (Rumus 2.1)$$

Keterangan:

S : himpunan

A : atribut

n : jumlah partisi S

$p_i$  : proporsi dari  $S_i$  terhadap S.

### 2) *Gain*

$$Gain(S,A) = Entropy(S) \sum_{j=1}^Y \frac{D_j}{D} * Entropy(S_j) \dots \dots \dots (Rumus 2.2)$$

*Gain Ratio* termasuk bagian dari perhitungan algoritma *C 5.0* dimana perhitungan ini yang membuat pembeda antara *C 4.5* dengan *C 5.0*. *Gain Ratio* juga bagian dari *Decision Tree* yang mana merupakan tahapan untuk menentukan nilai atribut yang yang lebih tepat, dengan begitu nilai atribut yang di hasilkan akan lebih akurat, dengan adanya *Gain Ratio* bisa mengurangi bias atribut yang memiliki banyak cabang. *Gain Ratio* memiliki sifat yaitu sebagai berikut (Adli Abdillah, 2021):

1. Bernilai besar bila data menyebar rata
2. Bernilai kecil jika semua data masuk dalam satu cabang

Dengan begitu *Gain Ratio* memiliki rumus yaitu sebagai berikut:

$$Gain Ratio = \frac{Gain(S,A)}{Split Info(S,A)} \dots \dots \dots (Rumus 2.3)$$

Akan tetapi sebelum melakukan penentuan nilai atribut dari *Gain Ratio*, harus melakukan perhitungan *split information*, *split information* adalah bagian dari formula yang digunakan untuk memilih sejumlah atribut, dengan rumusnya sebagai berikut:

$$SplitInfo(S,A) = \sum_i^c = 1 \frac{S_i}{S} \text{Log}_2 \frac{S_i}{S} \dots \dots \dots (Rumus 2.4)$$

S = Jumlah Data Sampel

$S_i$  = Jumlah masing-masing pada setiap atribut.

Selanjutnya menghitung *Gain Ratio* berdasarkan formula rumus yang ada di atas.

Dibawah ini merupakan strategi pengembangan *Decision Tree* dengan menggunakan algoritma *C 5.0*:

1. Penggambaran awal pohon adalah simpul tanpa daun tunggal yang mewakili kumpulan data pelatihan.
2. Kedua, jika setiap sampel memiliki kelas yang sama, maka simpul tersebut menjadi daun dan diberi label dari kelas tersebut.

3. Jika variabel tersebut tidak ditemukan, algoritme akan menggunakan pengukuran berbasis entropi (Penguatan Informasi) untuk memilih variabel prediktor yang akan mempartisi data ke dalam kelas terpisah. Variabel ini akan digunakan sebagai variabel tes atau keputusan pada node tersebut.
4. Setiap nilai yang ditentukan oleh variabel uji akan meningkatkan kabin yang sesuai. Daging tersebut kemudian dibagi menurut kriteria cabang.
5. Proses ini diulang tanpa batas waktu hingga Pohon Keputusan yang lengkap telah dibuat.
6. Partisi rekursif berakhir hanya ketika salah satu dari dua kondisi terpenuhi, yaitu:
  - 1) Semua catatan pada node tertentu milik kelas yang sama.
  - 2) Tidak ada karakteristik tetap dari catatan yang menghalangi pembagian lebih lanjut. Dalam hal ini, aturan suara mayoritas berlaku. Node yang sesuai diubah menjadi node daun dan diberi label dengan kelas mayoritas dalam kumpulan data yang tersedia.

## **2.7 Stok Barang**

Penting bagi setiap perusahaan untuk memperhatikan aspek persediaan, karena aktivitas persediaan memiliki biaya yang signifikan. Untuk memudahkan pengelolaan persediaan, dibutuhkan sistem pendukung yang dapat membantu dalam proses pengelolaan. Namun, pada kenyataannya, pengelolaan persediaan masih dilakukan secara manual (Prapitasari et al., 2017).

Persediaan merupakan barang-barang yang dimiliki oleh perusahaan, baik yang diperoleh dari pembelian maupun hasil produksi sendiri, yang bertujuan untuk dijual kepada konsumen.

Dapat disimpulkan stok barang adalah bagian yang perlu diperhatikan oleh setiap perusahaan untuk kestabilan persediaan barang yang dimiliki untuk dijual kembali kepada konsumen.

### **2.7.1 Tujuan Penjualan**

Kunci kesuksesan perusahaan dalam mencari keuntungan terletak pada kemampuan perusahaan dalam menjual produknya. Jika perusahaan tidak mampu

menjual produknya, maka perusahaan akan mengalami kerugian. Tujuan umum penjualan dalam perusahaan meliputi:

1. Tujuan yang dirancang untuk meningkatkan volume penjualan total atau meningkatkan penjualan produk-produk yang lebih menguntungkan.
2. Didesain dengan tujuan mempertahankan posisi penjualan yang kuat melalui kunjungan penjualan rutin untuk berbagi detail tentang item baru.
3. Membantu bisnis berkembang.

Jika penjualan berjalan sesuai rencana, tujuan ini dapat tercapai. Meskipun penjualan tidak selalu berjalan sesuai rencana, lingkungan pemasaran berdampak signifikan terhadap laba dan rugi perusahaan. Lingkungan memiliki dampak yang signifikan terhadap bagaimana bisnis berkembang.

## **2.8 Produk**

Produk memegang peranan penting dalam kesuksesan perusahaan karena produk adalah dasar dari usaha perusahaan. Untuk berhasil dalam pemasaran, produk harus sesuai dengan kebutuhan dan keinginan pelanggan. Oleh karena itu, perusahaan harus mengorientasikan pembuatan produk pada keinginan pasar dan selera konsumen. Saat ini, semua produsen memahami pentingnya kualitas produk yang unggul untuk memenuhi harapan pelanggan. Para petinggi perusahaan juga menyadari keterhubungan langsung antara kualitas produk dengan keputusan pembelian konsumen, yang pada akhirnya akan meningkatkan minat beli konsumen (Pati, 2016).

### **2.8.1 Pupuk**

Pupuk adalah zat yang mengandung unsur hara di dalamnya bagi tanaman untuk membantu pertumbuhan dan perkembangannya. Pupuk tersedia dalam bentuk cair dan padat. Pupuk dibagi menjadi dua kategori berdasarkan proses pembuatannya, yaitu:

1. Pupuk Organik

Untuk memberikan unsur hara, khususnya N dan C organik, serta untuk meningkatkan aspek fisik, kimia, dan biologi tanah, pupuk organik

banyak dihasilkan dari tumbuhan atau hewan. Karena ada berbagai macam kualitas komponen dasar yang digunakan untuk membuat pupuk, kualitas produk jadi juga bervariasi. Berbeda dengan pupuk anorganik, pupuk organik berfungsi sebagai suplemen. Ketika pupuk kimia buatan digunakan dalam rasio yang tidak seimbang dengan pupuk organik, struktur tanah dapat rusak dan aktivitas biologis tanah berkurang (Siregar et al., 2019).

## 2. Pupuk Anorganik

Pupuk anorganik adalah pupuk yang dihasilkan melalui proses rekayasa kimia, fisika, dan biologi sebagai hasil usaha atau pabrik yang membuat pupuk (Dewanto et al., 2017).

### 2.8.2 Herbisida

Herbisida adalah zat kimia yang dapat menghambat atau bahkan membunuh pertumbuhan tanaman. Herbisida dikategorikan menjadi dua kategori berdasarkan cara kerjanya: herbisida kontak dan herbisida sistemik. Herbisida kontak bekerja pada bagian yang terkena herbisida saja. Sedangkan herbisida sistemik adalah herbisida yang menuju ke jaringan tumbuhan.

Penting untuk memperhatikan bahan aktif dalam herbisida tersebut agar dapat membasmi gulma yang mengganggu tanaman, seperti dengan menggunakan herbisida dengan bahan kimia glifosat dan paraquat, yang dapat mengendalikan segala jenis gulma dan bertindak cepat membakar jaringan (Sembiring et al., 2019).

### 2.8.3 Insektisida

Insektisida adalah senyawa kimia beracun yang digunakan untuk membunuh serangga. Salah satu hama yang sangat merugikan pada tanaman sayuran di Indonesia adalah ulat grayak (*Spodoptera litura*). Ulat grayak dapat menyebabkan penurunan produktivitas hingga kegagalan panen karena merusak daun, buah, dan sayuran dengan membuatnya sobek, terpotong-potong, dan berlubang. Namun, penggunaan insektisida pada tingkat petani seringkali masih menggunakan senyawa kimia sintetis yang dapat merusak organisme non-target,



meningkatkan resistensi hama, menimbulkan resurgensi hama, serta meninggalkan efek residu pada tanaman dan lingkungan (Azwana et al., 2019).

## 2.9 CV. Mitra Karya Sejati

CV. Mitra Karya Sejati terletak di Jl. Menggala Sakti Km.24, Kec. Tanah Putih, Kab. Rokan Hilir, Riau yang merupakan distributor pupuk yang bergerak dibidang perdagangan salah satunya menyediakan berbagai macam bahan pertanian dengan harga terjangkau yang dibutuhkan masyarakat.

Penduduk Kabupaten Rokan Hilir, sebagian besar tinggal di daerah pedesaan, karena itu sesuai potensi daerah yang agraris maka mata pencaharian penduduk Kabupaten Rokan Hilir sebagian besar bekerja dibidang pertanian, baik sebagai buruh tani atau petani penggarap, maka CV Mitra Karya Sejati hadir untuk memenuhi kebutuhan masyarakat. Pasar utama untuk produk ini adalah bisnis pertanian, lembaga pemerintah, dan toko khusus. Usaha ini didirikan oleh Bapak Tardi Sutanto pada tahun 1998.

CV. Mitra Karya Sejati saat ini juga sudah menggunakan sistem informasi penjualan untuk mendukung kinerja pihak perusahaan pada layanan informasi penjualan.



**Gambar 2. 2** Logo Mitra Karya Sejati

([www.cvmks.com](http://www.cvmks.com))

## 2.10 Data

Data adalah informasi atau keterangan yang diperoleh melalui pengamatan atau sumber-sumber tertentu (Nababan et al., 2018). Data bisa berupa fakta mengenai suatu objek, baik berupa angka maupun kata-kata. Data adalah