



**PENERAPAN TEKNIK DATA MINING PADA DATABASE
PELANGGAN DENGAN MENGGUNAKAN ALGORITMA DECISION
TREE UNTUK MEMPREDIKSI TINGKAT KEPUASAN DI PT JALA
LINTAS MEDIA BRANCH BANTEN**

SKRIPSI

Oleh:

Andi Yunus Zai

20211000045

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS BUDDHI DHARMA

TANGERANG

2025



**PENERAPAN TEKNIK DATA MINING PADA DATABASE
PELANGGAN DENGAN MENGGUNAKAN ALGORITMA DECISION
TREE UNTUK MEMPREDIKSI TINGKAT KEPUASAN DI PT JALA
LINTAS MEDIA BRANCH BANTEN**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program

Studi Teknik Informatika Fakultas Sains dan Teknologi

Universitas Buddhi Dharma

Jenjang Pendidikan Strata 1

Oleh :

ANDI YUNUS ZAI

20211000045

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS BUDDHI DHARMA

TANGERANG

2025

LEMBAR PERSEMBAHAN

*“Berilah orang bijak nasehat, maka ia akan menjadi lebih bijak, ajarilah orang benar,
maka pengetahuannya akan bertambah”*

Amsal 9:9

Dengan mengucap puji syukur kepada Tuhan Yang Maha Esa, SKRIPSI ini
kupersembahkan untuk:

1. Bapak (Tongoni Zai) dan Ibu (Aslina Mendrofa) tercinta yang telah membesarkan aku dan selalu membimbing, mendukung, memotivasi, memberi apa yang terbaik bagiku serta selalu mendoakan aku untuk meraih kesuksesanku.
2. Kakak dan adik-adikku yang telah memberikan dukungan semangat serta dorongan yang senantiasa diberikan.
3. Teman-teman kelompok belajar yang selalu berjuang bersama.
4. Rekan-rekan dari PT Jala Lintas Media *Branch* Banten yang selalu memberikan dukungan dan semangat.

UNIVERSITAS BUDDHI DHARMA

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini.

NIM : 20211000045
Nama : Andi Yunus Zai
Jenjang Studi : Strata 1
Program Studi : Teknik Informatika
Peminatan : Data Science

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapat gelar akademik Sarjana atau kelengkapan studi, baik di Universitas Buddhi Dharma maupun di Perguruan Tinggi lainnya.
2. Skripsi ini saya buat sendiri tanpa bantuan dari pihak lain, kecuali arahan dosen pembimbing.
3. Dalam Skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dan dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan daftar pustaka.
4. Dalam Skripsi ini tidak terdapat pemalsuan (kebohongan), seperti buku, artikel, jurnal, data sekunder, pengolahan data, dan pemalsuan tanda tangan dosen atau Ketua Program Studi Universitas Buddhi Dharma yang dibuktikan dengan keasliannya.
5. Lembar pernyataan ini saya buat dengan sesungguhnya, tanpa paksaan dan apabila dikemudian hari atau pada waktu lainnya terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, saya bersedia menerima sanksi akademik berupa pencabutan gelar akademik yang telah saya peroleh karena Skripsi ini serta sanksi lainnya sesuai dengan peraturan dan norma yang berlaku.

Tangerang, 04-08-2025

Yang membuat pernyataan,



Andi Yunus Zai
20211000045

UNIVERSITAS BUDDHI DHARMA

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini.

NIM : 20211000045
Nama : Andi Yunus Zai
Jenjang Studi : Strata 1
Program Studi : Teknik Inforamtika
Peminatan : Data Science

Dengan ini menyetujui untuk memberikan ijin kepada pihak Universitas Buddhi Dharma, Hak Bebas Royalti Non – Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah kami yang berjudul: **“PENERAPAN DATA MINING DENGAN DECISION TREE UNTUK PREDIKSI KEPUASAN PELANGGAN DI PT JALA LINTAS MEDIA”**, beserta alat yang diperlukan (apabila ada). Dengan Hak Bebas Royalti Non – Eksklusif ini pihak Universitas Buddhi Dharma berhak menyimpan, mengalih-media atau format-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya, dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis pertama atau pencipta karya ilmiah tersebut. Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Buddhi Dharma, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 04-08-2025

Yang membuat pernyataan,



Andi Yunus Zai

20211000045

UNIVERSITAS BUDDHI DHARMA

LEMBAR PENGESAHAN PEMBIMBING

**PENERAPAN TEKNIK DATA MINING PADA DATABASE
PELANGGAN DENGAN MENGGUNAKAN ALGORITMA DECISION
TREE UNTUK MEMPREDIKSI TINGKAT KEPUASAN DI PT JALA
LINTAS MEDIA BRANCH BANTEN**

Dibuat Oleh:

NIM : 20211000045

Nama : Andi Yunus Zai

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian
Komprehensif

Program Studi Teknik Informatika

Peminatan Data Science

Tahun Akademik 2024/2025

Disahkan oleh,

Tangerang, 04-08-2025

Pembimbing,

Desiyanna Lasut, S.Kom., M.Kom

NUPTK. 1534764665230253

UNIVERSITAS BUDDHI DHARMA

LEMBAR PENGESAHAN SKRIPSI

**PENERAPAN TEKNIK DATA MINING PADA DATABASE
PELANGGAN DENGAN MENGGUNAKAN ALGORITMA DECISION
TREE UNTUK MEMPREDIKSI TINGKAT KEPUASAN DI PT JALA
LINTAS MEDIA BRANCH BANTEN**

Dibuat Oleh:

NIM : 20211000045

Nama : Andi Yunus Zai

Telah di setujui untuk di pertahankan dihadapan tim penguji

Komprehensif

Program studi teknik informatika

Peminatan Data Science

Tahun akademik 2024/2025

Disahkan oleh,

Tangerang, 04-08-2025

Dekan,



Dr. Yakub, M.Kom., M.M.

NUPTK. 1836747648130172

Ketua Program Study,



Hartana Wijaya, M.Kom

NUPTK. 3844759660131192

UNIVERSITAS BUDDHI DHARMA
LEMBAR PENGESAHAN TIM PENGUJI

Nama : Andi Yunus Zai

NIM : 20211000045

Fakultas : Sains dan Teknologi

Judul Skripsi : Penerapan Teknik Data Mining Pada Database Pelanggan Dengan
Menggunakan Algoritma Decision Tree Untuk Memprediksi Tingkat
Kepuasan di PT Jala Lintas Media Branch Banten

Dinyatakan LULUS setelah mempertahankan di depan Tim Penguji pada hari Senin, 04-
08-2025.

Nama Penguji:

Tanda Tangan:

Ketua Sidang : Rudy Arijanto, S.Kom., M.Kom
NUPTK. 5047749651200033

Penguji I : Ramona Dyah Safitri, S.Si, M.Si
NUPTK. 8652771672230312

Penguji II : Desiyanna Lasut, S.Kom., M.Kom
NUPTK. 1534764665230253

Mengetahui,

Dekan Fakultas Sains Dan Teknologi



Dr. Yakub, M.Kom., M.M.

NUPTK. 1836747648130172

KATA PENGANTAR

Dengan mengucapkan Puji Syukur kepada Tuhan Yang Maha Esa, yang telah memberikan Rahmat dan karunia-Nya kepada penulis sehingga dapat menyusun dan menyelesaikan Skripsi ini dengan judul **Penerapan Teknik Data Mining Pada Database Pelanggan Dengan Menggunakan Algoritma Decision Tree Untuk Memprediksi Tingkat Kepuasan Di Pt Jala Lintas Media Branch Banten**. Tujuan utama dari pembuatan Skripsi ini adalah sebagai salah satu syarat kelengkapan dalam menyelesaikan program pendidikan Strata 1 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. Ibu Dr. Limajatini, S.E., M.M., B.K.P. sebagai Rektor Universitas Buddhi Dharma.
2. Bapak Dr. Yakub, M.Kom., M.M. Dekan Fakultas Sains dan Teknologi.
3. Bapak Hartana Wijaya, M.Kom sebagai Ketua Program Studi Teknik Informatika.
4. Ibu Desiyanna Lasut, 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 disebutkan satu-persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini masih belum sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa yang akan datang. Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Tangerang, 04-08-2025

Penulis

PENERAPAN TEKNIK DATA MINING PADA DATABASE PELANGGAN DENGAN MENGGUNAKAN ALGORITMA DECISION TREE UNTUK MEMPREDIKSI TINGKAT KEPUASAN DI PT JALA LINTAS MEDIA BRANCH BANTEN

149 Halaman + xxi / 32 tabel / 53 gambar / 3 lampiran

ABSTRAK

Dalam era digitalisasi yang berkembang pesat, kemampuan perusahaan untuk memahami dan menganalisis data pelanggan menjadi salah satu faktor kunci dalam meningkatkan tingkat kepuasan pelanggan. Penelitian ini bertujuan untuk menerapkan teknik data mining dengan menggunakan algoritma *Decision Tree* untuk memprediksi tingkat kepuasan pelanggan di PT Jala Lintas Media *Branch* Banten berdasarkan data jumlah transaksi dan pendapatan bulanan. Permasalahan yang dihadapi adalah sulitnya perusahaan dalam menentukan pola kepuasan pelanggan secara otomatis, yang menyebabkan keputusan strategis sering kali diambil tanpa dasar yang kuat. Pendekatan yang digunakan dalam penelitian ini melibatkan pengumpulan data pelanggan, pengolahan data, penerapan algoritma *Decision Tree* untuk pembentukan model prediksi, dan evaluasi model berdasarkan akurasi yang diperoleh. Penelitian ini juga memanfaatkan sistem berbasis web untuk mempermudah proses pengolahan data dan visualisasi hasil prediksi. Solusi yang ditawarkan adalah model prediksi yang dapat digunakan perusahaan untuk mengklasifikasikan pelanggan menjadi dua kategori, yaitu puas dan tidak puas, berdasarkan parameter tertentu. Hasil dari penelitian menunjukkan bahwa model *Decision Tree* yang dikembangkan memiliki tingkat akurasi yang baik dalam memprediksi tingkat kepuasan pelanggan. Dengan implementasi sistem ini, perusahaan dapat membuat keputusan yang lebih tepat dan strategis untuk meningkatkan kepuasan pelanggan. Penelitian ini memberikan kontribusi dalam bentuk integrasi antara data *mining* dan teknologi berbasis web untuk mendukung pengambilan keputusan perusahaan secara lebih efisien.

Kata kunci: *Database, Datamining, Decision Tree, Kepuasan pelanggan, Teknik Informatika*

***APPLICATION OF DATA MINING TECHNIQUES ON CUSTOMER DATABASE
USING DECISION TREE ALGORITHM TO PREDICTE SATISFACTION LEVEL AT
PT JALA LINTAS MEDIA BANTEN BRANCH***

149 Pages + xxi / 32 tables / 53 images / 3 attachments

ABSTRACT

In the era of rapid digitalization, the ability of companies to understand and analyze customer data is one of the key factors in increasing customer satisfaction levels. This study aims to apply data mining techniques using the Decision Tree algorithm to predict customer satisfaction levels at PT Jala Lintas Media Branch Banten based on data on the number of transactions and monthly income. The problem faced is the difficulty of companies in determining customer satisfaction patterns automatically, which often causes strategic decisions to be taken without a strong basis. The approach used in this study involves collecting customer data, processing data, applying the Decision Tree algorithm to form a prediction model, and evaluating the model based on the accuracy obtained. This study also utilizes a web-based system to facilitate the process of data processing and visualization of prediction results. The solution offered is a prediction model that can be used by companies to classify customers into two categories, namely satisfied and dissatisfied, based on certain parameters. The results of the study indicate that the Decision Tree model developed has a good level of accuracy in predicting customer satisfaction levels. With the implementation of this system, companies can make more precise and strategic decisions to increase customer satisfaction. This study contributes in the form of integration between data mining and web-based technology to support more efficient corporate decision making.

Key Word: *Customer Satisfaction, Database, Datamining, Decision Tree, Teknik Informatika*

DAFTAR ISI

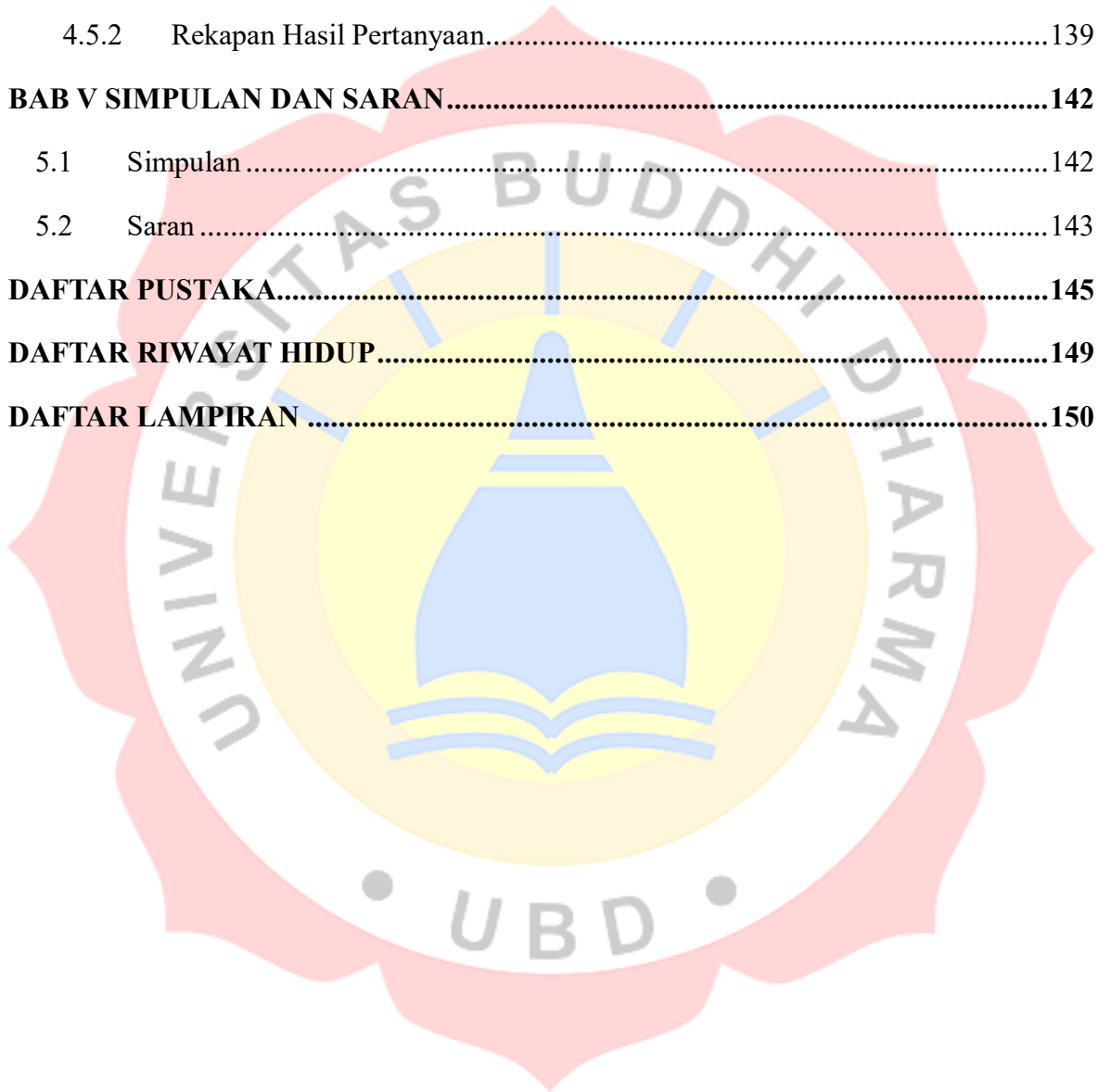
LEMBAR PERSEMBAHAN.....	iii
LEMBAR PERNYATAAN KEASLIAN SKRIPSI	iv
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
LEMBAR PENGESAHAN PEMBIMBING.....	vi
LEMBAR PENGESAHAN SKRIPSI	vii
LEMBAR PENGESAHAN TIM PENGUJI	viii
KATA PENGANTAR.....	ix
ABSTRAK.....	x
<i>ABSTRACT</i>	xi
DAFTAR ISI	xii
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR.....	xviii
DAFTAR LAMPIRAN	xx
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.2.1 Rumusan Masalah.....	4
1.3 Ruang Lingkup Penelitian.....	4
1.4 Tujuan dan Manfaat Penelitian.....	6
1.4.1 Tujuan	6
1.4.2 Manfaat	6
1.5 Sistematika Penulisan	7
BAB II LANDASAN TEORI.....	9
2.1 Teori Umum	9

2.1.1	Konsep Pengembangan.....	9
2.2	Konsep Informasi	10
2.2.1	Data.....	13
2.2.2	<i>Database</i>	14
2.3	Teori Khusus	16
2.3.1	Pelayanan Dan Jasa.....	16
2.3.2	Kepuasan Pelanggan.....	17
2.3.3	Loyalitas Pelanggan.....	17
2.3.4	<i>Data Mining</i>	18
2.3.5	Algoritma	22
2.3.6	<i>Decision Tree</i>	24
2.3.7	<i>MySQL</i>	27
2.3.8	<i>Python</i>	28
2.3.9	Aplikasi <i>Web</i>	28
2.4	Tinjauan Studi	29
2.4.1	Penerapan Algoritma <i>Decision Tree</i> C4.5 Untuk Prediksi Kelayakan Calon Pendorong Darah Dengan Klasifikasi <i>Data Mining</i>	29
2.4.2	Implementasi Algoritma <i>Decision Tree</i> dan <i>Naïve Bayes</i> Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks	32
2.4.3	Analisis Komparasi Algoritma Klasifikasi atau <i>Mining</i> dalam Klasifikasi Website <i>Phishing</i>	34
2.4.4	Implementasi Algoritma <i>Decision Tree</i> dalam Klasifikasi Kompetensi Siswa.	36
2.4.5	Implementasi <i>Data Mining</i> untuk Menentukan Minat Siswa dalam Menentukan Jurusan Pada Perguruan Tinggi	38
2.4.6	Penerapan <i>Data Mining Clustering</i> terhadap Perekonomian pada Kelurahan Wonorejo Kecamatan Marpoyan Damai Provinsi Riau Menggunakan Metode Algoritma <i>K-Means</i>	40

2.4.7	Penerapan Data <i>Mining</i> untuk Penentuan Paket Penjualan Produk Pecah Belah Menggunakan Algoritma <i>Apriori</i> (Studi Kasus: Toko Sumber Cahaya)	43
2.4.8	Analisis Perancangan Sistem untuk Kepuasan Pelanggan pada UD. Shinta <i>Elektronik</i> Dengan Menggunakan Metode Algoritma C4.5	44
2.4.9	Aplikasi <i>Klasifikasi</i> Kepribadian Manusia Menggunakan Algoritma <i>Decision Tree</i> (C4.5) Berbasis <i>Web</i>	46
2.4.10	Penerapan Data <i>Mining</i> Berbasis <i>Klastering</i> untuk Penilaian Status Gizi Balita di Puskesmas	48
2.4.11	<i>Machine Learning Algorithms – A Review</i>	50
2.4.12	<i>Classification Based on Decision Tree Algorithm for Machine Learning</i>	52
2.4.13	<i>Predictive Analytics in Business Analytics: Decision Tree</i>	54
2.4.14	<i>Implementation Educational Data Mining for Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5</i>	56
2.4.15	<i>Educational Data Mining Using Cluster Analysis and Decision Tree Technique: A Case Study</i>	59
BAB III METODOLOGI PENELITIAN.....		64
3.1	Tinjauan Umum Perusahaan	64
3.1.1	Profil dan Sejarah Perusahaan	64
3.1.2	Struktur organisasi, tugas dan wewenang setiap organisasi perusahaan	65
3.2	Identifikasi Kebutuhan.....	66
3.3	Metode Yang Digunakan.....	67
3.3.1	Metode Pengumpulan Data	67
3.3.2	Metode Data Mining.....	68
3.3.3	Metode Analisis Data	68
3.3.4	Evaluasi Model Prediksi	69
3.4	Algoritma Yang digunakan	70
3.4.1	Pohon Keputusan Algoritma Proses	70

3.4.2	Keunggulan Decision Tree	71
3.4.3	Implementasi	71
3.5	Kerangka Pemikiran	72
3.6	Perancangan <i>UML</i>	76
3.6.1	<i>Use Case Diagram</i>	76
3.6.2	<i>Use-case Scenario</i>	76
3.6.3	<i>Activity Diagram</i>	80
3.7	Perancangan Layar	83
3.7.1	Perancangan Menu.....	83
3.7.2	Perancangan <i>Database</i>	86
3.8	Jadwal Penelitian.....	89
BAB IV HASIL DAN PEMBAHASAN.....		90
4.1	Spesifikasi <i>Hardware</i> dan <i>Software</i>	90
4.1.1	Spesifikasi <i>Hardware</i> (Perangkat Keras).....	90
4.1.2	Spesifikasi <i>Software</i> (Perangkat Lunak).....	90
4.2	Tampilan Program	92
4.2.1	Halaman <i>Login</i>	92
4.2.2	Halaman <i>Register</i>	93
4.2.3	Halaman <i>Dashboard</i>	94
4.2.4	Halaman <i>Upload Dataset</i> Pelanggan.....	94
4.2.5	Halaman Latih Model	95
4.2.6	Halaman <i>Upload Data</i> Tanpa <i>Label</i> /Kepuasan	96
4.2.7	Halaman Prediksi Kepuasan Pelanggan.....	98
4.2.8	Halama Lihat Hasil Prediksi	99
4.2.9	Halaman <i>Statistik</i>	99
4.2.10	Halaman Lihat Hasil Analisis	100
4.2.11	Halaman <i>Logout</i>	107

4.3	Penerapan Metode dan Algoritma	108
4.3.1	Algoritman <i>Decision Tree</i> C4.5	108
4.4	<i>Blackbox Testing</i>	118
4.5	<i>User Acceptance Testing (UAT)</i>	121
4.5.1	Proses pengelolaan data kuesioner	124
4.5.2	Rekapan Hasil Pertanyaan.....	139
BAB V SIMPULAN DAN SARAN.....		142
5.1	Simpulan	142
5.2	Saran	143
DAFTAR PUSTAKA.....		145
DAFTAR RIWAYAT HIDUP		149
DAFTAR LAMPIRAN		150



DAFTAR TABEL

Tabel 3. 1 Kerangka Pemikiran.....	72
Tabel 3. 2 <i>Use-case Scenario login</i>	76
Tabel 3. 3 <i>Use-case Scenario upload dataset</i>	77
Tabel 3. 4 <i>Use-case Scenario</i> tambahkan kolom kepuasan.....	78
Tabel 3. 5 <i>Use-case Scenario</i> latih model	78
Tabel 3. 6 <i>Use-case Scenario</i> prediksi kepuasan.....	79
Tabel 3. 7 <i>Use-case Scenario logout</i>	80
Tabel 3. 8 Tabel <i>database</i> Pelanggan	86
Tabel 3. 9 Tabel <i>Dataset</i>	87
Tabel 3. 10 Tabel Prediksi	88
Tabel 3. 11 Tabel <i>User</i>	88
Tabel 4. 1 Tabel <i>Blackbox Testing</i>	119
Tabel 4. 2 Hasil kuesioner dari seluruh pertanyaan	122
Tabel 4. 3 Skor Ideal	124
Tabel 4. 4 Hasil Jawaban Pertanyaan Ke 1.....	124
Tabel 4. 5 Hasil Jawaban Pertanyaan ke 2	125
Tabel 4. 6 Hasil Jawaban Pertanyaan Ke 3.....	126
Tabel 4. 7 Hasil Jawaban Pertanyaan Ke 4.....	127
Tabel 4. 8 Hasil Jawaban Pertanyaan Ke 5.....	128
Tabel 4. 9 Hasil Jawaban Pertanyaan Ke 6.....	128
Tabel 4. 10 Hasil Jawaban Pertanyaan Ke 7.....	129
Tabel 4. 11 Hasil Jawaban Pertanyaan Ke 8.....	130
Tabel 4. 12 Hasil Jawaban Pertanyaan Ke 9.....	131
Tabel 4. 13 Hasil Jawaban Pertanyaan Ke 10.....	132
Tabel 4. 14 Hasil Jawaban Pertanyaan Ke 11	133
Tabel 4. 15 Hasil Jawaban Pertanyaan Ke 12.....	134
Tabel 4. 16 Hasil Jawaban Pertanyaan Ke 13.....	135
Tabel 4. 17 Hasil Jawaban Pertanyaan Ke 14.....	136
Tabel 4. 18 Hasil Jawaban Pertanyaan Ke 15.....	137
Tabel 4. 19 Hasil Jawaban Pertanyaan Ke 16.....	138
Tabel 4. 20 Hasil Rekapitulasi Pertanyaan Kuesioner	140

DAFTAR GAMBAR

Gambar 2. 1 Tahapan proses data mining	21
Gambar 2. 2 Contoh <i>Desision Tree</i> (Pohon Keputusan)	25
Gambar 3. 1 Struktur organisasi, tugas dan wewenang	65
Gambar 3. 2 Struktur organisasi, tugas dan wewenang	65
Gambar 3. 3 <i>Use-case Diagram</i>	76
Gambar 3. 4 <i>Activity Diagram Login</i>	80
Gambar 3. 5 <i>Activity Diagram Upload Dataset</i>	81
Gambar 3. 6 <i>Activity Diagram Latih Model</i>	81
Gambar 3. 7 <i>Activity Diagram Prediksi Kepuasan</i>	82
Gambar 3. 8 <i>Activity Diagram Latih Model</i>	82
Gambar 3. 9 Perancangan Menu <i>Login</i>	83
Gambar 3. 10 Perancangan Menu <i>Dashboard</i>	83
Gambar 3. 11 Perancangan Menu <i>Upload Dataset</i>	84
Gambar 3. 12 Perancangna Menu Latih Model.....	84
Gambar 3. 13 Perancangan Menu Prediksi Kepuasan	85
Gambar 3. 14 Perancangan Menu Lihat <i>Statistik</i>	85
Gambar 3. 15 Perancangan Menu <i>Logout</i>	86
Gambar 3. 16 Jadwal Penelitian	89
Gambar 4. 1 tampilan <i>Login</i>	93
Gambar 4. 2 Tampilan <i>Register</i>	93
Gambar 4. 3 Tampilan <i>Dashboard</i>	94
Gambar 4. 4 Tampilan <i>Upload Dataset</i> Pelanggan.....	95
Gambar 4. 5 Hasil <i>Upload Dataset</i> Pelanggan.....	95
Gambar 4. 6 Halaman Latih Model	96
Gambar 4. 7 Hasil Evaluasi Model.....	96
Gambar 4. 8 Halaman <i>Upload Data Tanpa Label</i>	97
Gambar 4. 9 Hasil <i>Upload Data Tanpa Label</i>	97
Gambar 4. 10 Tampilan Halaman Prediksi Kepuasan.....	98
Gambar 4. 11 Hasil Prediksi Kepuasan di <i>Database</i>	98
Gambar 4. 12 Tampilan Hasil Prediksi Kepuasan	99
Gambar 4. 13 Tampilan <i>Statistik</i> Pelanggan.....	99

Gambar 4. 14 Tampilan Halaman Analisis Evaluasi Data Uji dan Hasil Prediksi	100
Gambar 4. 15 tampilan Hasil Pohon Keputusan.....	104
Gambar 4. 16 Tampilan Halaman Konfirmasi Keluar	107
Gambar 4. 17 Tampilan Berhasil Keluar	107
Gambar 4. 18 Sempel Data Trining Hitung Manual	108
Gambar 4. 19 Diagram Hasil Pertanyaan ke 1	125
Gambar 4. 20 Diagram Hasil Pertanyaan Ke 2.....	126
Gambar 4. 21 Diagram Hasil Pertanyaan Ke 3.....	127
Gambar 4. 22 Diagram Hasil Pertanyaan Ke 4.....	127
Gambar 4. 23 Diagram Hasil Pertanyaan Ke 5.....	128
Gambar 4. 24 Diagram Hasil Pertanyaan Ke 6.....	129
Gambar 4. 25 Diagram Hasil Pertanyaan Ke 7.....	130
Gambar 4. 26 Diagram Hasil Pertanyaan Ke 8.....	131
Gambar 4. 27 Diagram Hasil Pertanyaan Ke 9.....	132
Gambar 4. 28 Diagram Hasil Pertanyaan Ke 10.....	133
Gambar 4. 29 Diagram Hasil Pertanyaan Ke 11.....	134
Gambar 4. 30 Diagram Hasil Pertanyaan Ke 12.....	135
Gambar 4. 31 Diagram Hasil Pertanyaan Ke 13.....	136
Gambar 4. 32 Diagram Hasil Pertanyaan Ke 14.....	137
Gambar 4. 33 Diagram Hasil Pertanyaan Ke 15.....	138
Gambar 4. 34 Diagram Hasil Pertanyaan Ke 16.....	139
Gambar 4. 35 Diagram Hasil Rekapitulasi.....	141

DAFTAR LAMPIRAN

Lampiran 1: Kartu Bimbingan Skripsi	150
Lampiran 2: Surat Keterangan Penelitian Skripsi.....	151
Lampiran 3: Pertanyaan Kuesioner.....	152



BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam dunia persaingan bisnis yang semakin ketat, kepuasan pelanggan memainkan peran yang sangat penting dalam menjaga loyalitas serta meningkatkan reputasi perusahaan. Bagi perusahaan yang bergerak di bidang layanan, seperti PT Jala Lintas Media *Branch* Banten, pemahaman yang mendalam tentang bagaimana pelanggan menilai kualitas layanan mereka merupakan kebutuhan yang sangat krusial. Kepuasan pelanggan tidak hanya menjadi tolak ukur keberhasilan perusahaan dalam memberikan layanan terbaik, tetapi juga menjadi dasar pengambilan keputusan strategis yang dapat meningkatkan kualitas layanan dan daya saing.

Kepuasan pelanggan berkaitan erat dengan penciptaan nilai pelanggan. Ketika pelanggan merasa puas, hubungan antara perusahaan dan pelanggannya menjadi lebih harmonis, menciptakan kepuasan yang lebih tinggi dan menghasilkan rekomendasi dari mulut ke mulut yang bermanfaat. Jika pelanggan merasa bahwa suatu produk atau layanan memenuhi atau melampaui harapan mereka, mereka akan lebih puas dan cenderung menggunakan produk atau layanan tersebut lagi (Sambodo Rio Sasongko, 2021).

Salah satu tujuan utama dan terpenting dalam industri jasa, khususnya untuk perusahaan, adalah memastikan bahwa pelanggan mereka merasa puas untuk membina hubungan yang langgeng. *Fitsimmons* mengklaim bahwa perbandingan antara harapan dan persepsi pelanggan terhadap layanan dapat mengungkapkan apakah pelanggan puas dengan layanan tersebut atau tidak. Jumlah konsumen yang menggunakan layanan dari waktu ke waktu menghasilkan kepuasan ini. Umpan balik positif dari pelanggan menumbuhkan

kepercayaan pelanggan, yang mendorong mereka untuk terus menggunakan layanan tersebut dan merekomendasikan lokasi yang sama kepada orang lain. Pelanggan yang memiliki pandangan positif terhadap perusahaan akan berjanji untuk menggunakan layanannya lagi dan merekomendasikannya kepada orang lain. Evaluasi pelanggan terhadap kinerja yang sesuai dengan kesenjangan antara apa yang diharapkan dan apa yang diterima dikenal sebagai kepuasan pelanggan. Persepsi pelanggan terhadap layanan, sistem dan dukungan layanan, pengalaman pelanggan menggunakan layanan, faktor situasional, dan faktor pribadi semuanya merupakan faktor yang memengaruhi kepuasan pelanggan (Azis, 2020).

Namun, PT Jala Lintas Media *Branch* Banten menghadapi kendala dalam memahami kekuatan dan kelemahan mereka. Mereka belum memiliki sistem yang efektif untuk mengidentifikasi kekurangan dalam pelayanan dan menemukan kelebihan yang dapat diandalkan sebagai nilai jual. Akibatnya, perusahaan kesulitan mengetahui bagaimana pelanggan mereka sebenarnya menilai kualitas layanan yang diberikan. Tanpa pemahaman yang jelas terkait *feedback* pelanggan, perusahaan mengalami kebingungan dalam merumuskan strategi peningkatan kualitas layanan yang tepat.

Permasalahan ini tidak hanya membatasi kemampuan PT Jala Lintas Media dalam memahami kebutuhan dan harapan pelanggan, tetapi juga menghalangi perusahaan untuk merespons perubahan tren dan keinginan pasar dengan cepat. Kondisi ini membuat mereka berada dalam situasi yang kurang optimal dalam menjaga hubungan jangka panjang dengan pelanggan, yang pada akhirnya dapat mempengaruhi pertumbuhan dan *stabilitas* perusahaan.

Untuk mengatasi permasalahan tersebut, penelitian ini bertujuan untuk menerapkan teknik *data mining*, khususnya algoritma *Decision Tree*, guna menganalisis dan memprediksi tingkat kepuasan pelanggan. Dengan menggunakan data historis pelanggan yang ada, algoritma ini akan membantu mengidentifikasi faktor-faktor yang paling berpengaruh terhadap kepuasan pelanggan, sehingga perusahaan dapat lebih memahami kebutuhan pelanggan dan melakukan perbaikan layanan secara tepat.

Berdasarkan latar belakang tersebut, penulis membuat penelitian dengan judul **”Penerapan Teknik Data Mining pada Database Pelanggan Dengan Menggunakan Algoritma Decision Tree untuk Memprediksi Tingkat Kepuasan di PT Jala Lintas Media Branch Banten”**. Penerapan yang akan di buat di harapkan memiliki kelebihan baik dalam kepuasan pelanggan maupun bagi pelayanan perusahaan.

1.2 Identifikasi Masalah

Berdasarkan permasalahan di latar belakang yang telah tertulis, kepuasan pelanggan sangat berpengaruh terhadap kelemahan dan kekurangan mereka dalam melayani setiap pelanggan, sehingga mereka perlu memiliki sistem yang efektif untuk mengidentifikasi kekurangan mereka dalam melayani pelanggan. Dengan demikian yang menjadi permasalahan adalah bagaimana penerapan algoritman *Decision Tree* dalam menganalisis dan memprediksi tingkat kepuasan pelanggan di PT Jala Lintas Media Branch Banten.

1.2.1 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan algoritma *Decision Tree* dalam menganalisis data pelanggan di PT Jala Lintas Media *Branch* Banten?
2. Faktor-faktor apa saja yang paling berpengaruh terhadap tingkat kepuasan pelanggan di PT Jala Lintas Media *Branch* Banten berdasarkan hasil analisis *Decision Tree*?
3. Seberapa akurat model *Decision Tree* dalam memprediksi kepuasan pelanggan berdasarkan data historis yang tersedia?
4. Bagaimana hasil dari penerapan algoritma *Decision Tree* dapat membantu PT Jala Lintas Media *Branch* Banten dalam meningkatkan layanan kepada pelanggan?

1.3 Ruang Lingkup Penelitian

1. Penelitian ini dilakukan pada *database* pelanggan PT Jala Lintas Media *Branch* Banten. Data yang digunakan meliputi informasi pelanggan, riwayat interaksi, dan tingkat kepuasan pelanggan yang telah dikumpulkan dari berbagai sumber, seperti *survei* dan data *operasional*.
2. Fokus dari penelitian ini adalah penerapan algoritma *Decision Tree* untuk memprediksi tingkat kepuasan pelanggan. Algoritma ini akan digunakan untuk mengidentifikasi faktor-faktor penting yang memengaruhi kepuasan, serta membantu memprediksi apakah seorang pelanggan akan merasa puas atau tidak berdasarkan data *historis* yang tersedia.

3. Data yang dianalisis hanya mencakup data pelanggan yang sudah ada dalam *database* perusahaan dan yang memiliki catatan lengkap mengenai riwayat *interaksi* dengan layanan. Data pelanggan yang tidak lengkap atau tidak relevan akan dibersihkan (*data cleansing*) sebelum analisis dilakukan.
4. Metode yang digunakan dalam penelitian ini adalah teknik *data mining* dengan menggunakan algoritma *Decision Tree*. Algoritma ini dipilih karena mampu menghasilkan model klasifikasi yang dapat diinterpretasikan secara mudah dan jelas oleh perusahaan.
5. Proses pengolahan data mencakup pengumpulan data, *preprocessing* data (seperti pembersihan dan normalisasi), pembagian data menjadi set pelatihan dan pengujian, serta evaluasi hasil dengan menggunakan metrik kinerja seperti *accuracy*, *precision*, *recall*, dan *F1-score*.
6. Penelitian ini tidak mencakup pengembangan aplikasi operasional secara penuh, namun akan fokus pada pengembangan model prediksi menggunakan algoritma *Decision Tree* dan penyediaan laporan hasil analisis yang dapat digunakan oleh perusahaan dalam pengambilan keputusan strategis terkait layanan pelanggan. Penelitian ini dilakukan selama periode dari 15 Maret 2025 sampai 28 Mei 2025 tentunya di PT Jala Lintas Media Branch Banten, dengan data yang diambil dari sistem manajemen pelanggan perusahaan.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan

Penelitian ini bertujuan untuk:

1. Menerapkan algoritma **Decision Tree** pada database pelanggan PT Jala Lintas Media Branch Banten guna memprediksi tingkat kepuasan pelanggan.
2. Mengidentifikasi faktor-faktor yang paling berpengaruh terhadap kepuasan pelanggan berdasarkan analisis data pelanggan.
3. Mengevaluasi akurasi model **Decision Tree** dalam memprediksi kepuasan pelanggan, serta memberikan *insight* yang dapat digunakan oleh perusahaan untuk meningkatkan kualitas layanan.
4. Memberikan rekomendasi kepada PT Jala Lintas Media Branch Banten berdasarkan hasil analisis, untuk membantu perusahaan meningkatkan pengalaman pelanggan dan mempertahankan loyalitas mereka.

1.4.2 Manfaat

Penelitian ini diharapkan dapat memberikan beberapa manfaat, baik bagi perusahaan maupun dalam pengembangan ilmu pengetahuan, sebagai **berikut**:

1. Bagi PT Jala Lintas Media Branch Banten:
 - Mendapatkan pemahaman yang cukup mendalam mengenai unsur-unsur yang memengaruhi kepuasan pelanggan.

- Menyediakan alat bantu berupa model prediksi yang dapat digunakan oleh perusahaan untuk memonitor dan meningkatkan kualitas layanan berdasarkan dataset pelanggan.
- Mempermudah perusahaan dalam menyusun strategi pelayanan yang lebih efektif, sehingga dapat meningkatkan kepuasan dan loyalitas pelanggan.

2. Bagi Dunia Akademik:

- Menambah wawasan mengenai penerapan teknik *data mining*, khususnya algoritma *Decision Tree*, dalam memprediksi kepuasan pelanggan.
- Menyajikan studi kasus penerapan algoritma *Decision Tree* yang dapat dijadikan referensi untuk penelitian serupa di masa mendatang.

3. Bagi Peneliti Lain:

- a. Memberikan referensi dan contoh penerapan metode *Decision Tree* yang dapat digunakan sebagai acuan dalam penelitian di bidang yang sama, terutama yang berkaitan dengan prediksi dan analisis data pelanggan.

1.5 Sistematika Penulisan

Sistematika penulisan ini disusun untuk memberikan gambaran umum mengenai alur dan isi dari skripsi yang akan dibahas. Dalam skripsi ini adapun sistematika penulisan adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang masalah yang menjadi alasan utama dilakukannya penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian, serta sistematika penulisan ini sendiri.

BAB II LANDASAN TEORI

Bab ini berisi kajian teori, baik teori umum, teori khusus, ataupun teori analisis yang relevan dengan penelitian dan pengembangan sistem, meliputi konsep dasar, pengembangan sistem mengenai data mining, *database* algoritma **Decision Tree**, kepuasan pelanggan, dan juga teori-teori lain yang mendukung pembahasan dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini membahas metode yang digunakan dalam penelitian untuk mencapai tujuan yang telah dirumuskan. Bab ini meliputi pendekatan penelitian, teknik pengumpulan data, metode pengolahan data, serta teknik analisis yang diterapkan dalam penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil implementasi sistem prediksi, hasil pengujian dan pelatihan model, evaluasi dan akurasi, serta visualisasi pohon keputusan yang di hasilkan. Bab ini juga membahas faktor – faktor yang paling memengaruhi tingkat kepuasan pelanggan berdasarkan analisis yang dilakukan

BAB V SIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian dan juga demikian dengan saran – sarang yang dapat digunakan untuk pengembangan sistem atau penelitian selanjutnya. Simpulan dan saran disusun berdasarkan hasil pengujian dan hasil analisis yang telah dilakukan pada bab – bab sebelumnya.

BAB II

LANDASAN TEORI

2.1 Teori Umum

2.1.1 Konsep Pengembangan

1. Definisi Sistem Pengembangan

Menurut Kebenaran Wau, pengembangan sistem sangat penting, salah satunya pengembangan sebuah sistem berbasis *website*, sehingga membuat kemudahan dalam mengontrol semua persediaan misalnya barang, pemesanan, juga pencatatan barang masuk atau keluar, dan tentunya diharapkan membuat pekerjaan lebih mudah dan lancar (Wau, 2022).

Pengembangan sistem juga merupakan proses menciptakan sistem baru atau mengubah sistem yang sudah ada untuk memenuhi kebutuhan tertentu.

Proses ini melibatkan sejumlah proses, seperti analisis persyaratan, desain sistem, implementasi, dan pemeliharaan.

2. Tujuan Pengembangan Sistem

Pengembangan sistem adalah proses perencanaan, pembangunan, pengujian, dan pemeliharaan sistem atau perangkat lunak untuk memenuhi tujuan tertentu. Proses ini melibatkan sejumlah langkah, dimulai dengan analisis kebutuhan pengguna dan berlanjut hingga desain sistem, implementasi, pengujian, dan pemeliharaan. Tujuannya adalah untuk menciptakan sistem yang efektif dan efisien yang dapat membantu atau meningkatkan operasi perusahaan atau aktivitas yang dilakukan oleh individu atau organisasi.

2.2 Konsep Informasi

1. Definisi Dari Informasi

Secara sederhana, informasi adalah data yang telah diproses atau dimodifikasi sehingga dapat dipahami dan bermanfaat bagi penerimanya. Informasi merupakan bagian penting dari proses pengambilan keputusan baik di tingkat individu maupun organisasi. Ketepatan waktu, relevansi, kebenaran, dan kejelasan merupakan beberapa karakteristik mendasar dari informasi berkualitas tinggi.

Teori *kuantum* dan teori informasi mendefinisikan bahwa, informasi merupakan komponen penting untuk memahami perkembangan dan nilai-nilai sosial. Informasi tidak hanya digunakan untuk menyampaikan fakta atau data; informasi juga merupakan komponen penting dari proses pemberian makna dan penyesuaian sistem nilai masyarakat. Terutama mengingat munculnya fenomena seperti krisis lingkungan, kecerdasan buatan (AI), dan kemajuan *transdisipliner* lainnya, informasi dianggap sebagai alat untuk memengaruhi dan memperkirakan perubahan sosial yang besar (Adolph, 2024).

2. Tujuan Dari Informasi

Tujuan informasi adalah untuk memungkinkan pengambilan keputusan, penyelesaian masalah, dan pelaksanaan berbagai tugas dengan menawarkan data dan pengetahuan yang relevan. Data mentah diproses untuk membuat informasi yang jika diperiksa dan diatur secara metodis, menghasilkan nilai tambah dalam bentuk pengetahuan atau wawasan yang relevan. Peran ini sangat penting dalam sejumlah *domain*, termasuk teknologi informasi, industri, pemerintahan, dan pendidikan.

Menurut (Adolph, 2024), informasi berfungsi untuk memberdayakan ilmuwan sosial terutama ekonom untuk menawarkan perkiraan yang lebih akurat dan panduan yang relevan saat masyarakat bergerak menuju keberlanjutan sosial. Fungsi informasi juga melibatkan pelanggaran hukum ekonomi yang sudah lama berlaku dan penurunan asumsi dengan mengambil sudut pandang fisika *kuantum*. Hal ini bertujuan untuk mendorong nilai-nilai sosial agar berkembang dan beradaptasi guna mengatasi masalah-masalah baru yang tidak dapat ditangani oleh sistem nilai *konvensional*.

Oleh karena itu, informasi memegang peranan penting dalam memfasilitasi pelaksanaan beberapa tugas, terutama dalam lingkungan yang semakin dinamis dan *kompetitif*. Bagi siapapun yang ingin membuat keputusan yang lebih tepat dan efisien, informasi yang tepat waktu, relevan, dan dapat dipercaya merupakan sumber daya yang sangat berharga.

3. Siklus dari Informasi

Menurut (Farell et al., 2018), siklus informasi merupakan proses di mana data diproses melalui model tertentu untuk diubah menjadi informasi. Informasi tersebut kemudian diterima oleh penerima yang menggunakannya untuk membuat keputusan dan melakukan tindakan. Tindakan tersebut menghasilkan data baru yang akan ditangkap kembali sebagai input dan diproses kembali melalui model yang sama, sehingga membentuk siklus yang berkelanjutan. Siklus ini disebut sebagai "*information cycle*" atau dalam beberapa referensi dikenal sebagai siklus pengolahan data (*data processing cycle*). Dalam konteks yang lebih luas, siklus informasi tidak hanya berhenti pada pemrosesan data menjadi informasi, tetapi

juga mencakup langkah-langkah seperti pengumpulan, distribusi, penggunaan, penyimpanan, dan penghapusan informasi.

Siklus informasi dari berbagai sumber awalnya berbagai bentuk misalnya media, termasuk pembuatan, pendistribusian, penggunaan, penyimpanan, dan pembuangan, disebut sebagai siklus informasi. Biasanya, siklus ini memiliki beberapa tahap utama:

1. **Pembuatan:** Orang atau organisasi membuat informasi. Temuan penelitian, studi, artikel, berita, dan data lain yang belum dirilis mungkin merupakan contohnya. Pada titik ini, informasi biasanya hanya dapat diakses dalam keadaan belum diproses.
2. **Pengumpulan dan Pemrosesan:** Setelah informasi dibuat, informasi tersebut dikumpulkan dan dianalisis untuk tujuan tertentu. Informasi sekarang dapat diperiksa, diteliti, dan dimodifikasi untuk memenuhi persyaratan.
3. **Distribusi:** Informasi yang diproses disebarkan ke khalayak yang lebih luas melalui berbagai media, termasuk jurnal cetak dan ilmiah, internet, media sosial, *konferensi*, dan publikasi lainnya.
4. **Penggunaan:** Setelah informasi dapat diakses secara luas, informasi tersebut dapat digunakan untuk berbagai alasan, termasuk pendidikan, pengambilan keputusan, dan penelitian masa depan oleh para peneliti, pembuat kebijakan, dan masyarakat umum.
5. **Pemeliharaan dan Penyimpanan:** Informasi yang digunakan biasanya disimpan untuk akses selanjutnya di lokasi yang ditentukan, seperti basis data, perpustakaan, arsip digital, atau *repository*.

6. Pembuangan: Untuk menjaga agar data yang direkam tetap relevan, informasi yang tidak lagi diperlukan atau relevan dapat dihapus atau diubah.

E Triandini, S Jayanatha, A Indrawan, G W Putra, dan B Iswara menyatakan bahwa, teknik pengembangan ini akan digunakan untuk menentukan perangkat operasi sistem informasi, yang dapat berupa platform *desktop*, *mobile*, atau berbasis *web*. Karena platform atau perangkat tempat sistem berjalan dapat digunakan untuk mengukur seberapa mudah sistem diakses oleh pengguna, pemilihan sistem operasi merupakan langkah penting dalam proses tersebut. Mengingat pentingnya *metodologi* pengembangan sistem informasi dan pemilihan platform tempat sistem berjalan, data dari penelitian sistem informasi sebelumnya dikumpulkan untuk penelitian ini guna menentukan *platform* sistem informasi dan metode pengembangannya (Triandini et al., 2019).

2.2.1 Data

Kumpulan informasi yang dikumpulkan dan disusun dengan tujuan tertentu disebut data. Angka, teks, gambar, dan format lain yang dapat diproses dan diperiksa semuanya dapat dianggap sebagai jenis data. Data pada dasarnya adalah bahan mentah yang belum diproses yang akan diproses untuk menghasilkan informasi yang bermakna.

Bagian Analisis Data menjelaskan gagasan mendasar tentang data dalam penulisan ini.

Ada dua kategori utama data yang digunakan dalam penelitian yaitu, data *kuantitatif* dan data *kualitatif* (Sutriani dan Octaviani, 2019).

1. Data *kuantitatif*: Angka-angka seperti populasi, usia, dan statistik lainnya adalah contoh data *kuantitatif*. Teknik statistik seperti menentukan mean,

simpangan baku, atau *korelasi* antar variabel dapat digunakan untuk memeriksa data ini.

2. Data *Kualitatif*: Informasi yang tidak dapat diukur, seperti penggunaan bahasa, jenis kelamin, atau warna kulit. Konstruksi gagasan berdasarkan data lapangan biasanya dikaitkan dengan analisis data *kualitatif*, yang memerlukan klasifikasi dan interpretasi tanpa memerlukan perhitungan numerik yang ketat.

Penyederhanaan dan pengklasifikasian data untuk interpretasi baik dalam bentuk tabel atau frekuensi untuk penelitian *kuantitatif*, atau kata-kata dan deskripsi untuk penelitian *kualitatif*, yang merupakan tujuan analisis data. Metode pemrosesan data adalah pembeda utama antara keduanya: induktif (dari data ke teori) dalam penelitian *kualitatif* dan deduktif (dari teori ke data) dalam penelitian *kuantitatif* (Sutriani dan Octaviani, 2019).

2.2.2 *Database*

Database adalah kumpulan informasi yang telah diatur secara metodis untuk memudahkan pengelolaan, pencarian, dan pengambilan informasi. Data dapat disimpan, dimodifikasi, dan dihapus oleh pengguna menggunakan sistem manajemen basis data (*DBMS*), yang biasanya digunakan untuk mengelola basis data (Riyan Dirgantara et al., 2023).

1. Konsep Dasar *Databas*

- a. Struktur Data: Basis data relasional menyimpan data sebagai tabel, basis data berbasis objek menyimpan data sebagai objek, dan basis data *NoSQL* menyimpan data sebagai dokumen. Setiap baris dalam

tabel mewakili entitas atau rekaman data yang berbeda. Tabel terdiri dari baris dan kolom.

- b. Relasional: Pada database relasional memiliki data dalam tabel-tabel yang saling terkait. Hubungan antara table yang diatur melalui kunci (key), seperti *primary key* dan *Foreign key*. Ini digunakan untuk menjaga integritas data dan memudahkan manajemen hubungan antar entitas.
- c. Normalisasi: Normalisasi adalah proses mengatur data dalam tabel untuk menghindari duplikasi dan menghilangkan ketergantungan yang tidak perlu, serta memastikan bahwa data disimpan secara efektif dan konsisten.
- d. Transaksi: Dalam *database*, transaksi adalah sekelompok aktivitas yang dijalankan secara bersamaan. Setiap transaksi harus memenuhi karakteristik *ACID* (*Atomicity, Consistency, Isolation, Durability*) untuk menjamin bahwa data akan tetap konsisten bahkan jika terjadi kegagalan sistem. Karakteristik ini membantu menjaga konsistensi data saat terjadi kegagalan sistem.
- e. Keamanan dan Pengelolaan Akses: Untuk membatasi akses ke data, basis data juga dilengkapi mekanisme keamanan yang memungkinkan hak pengguna.

2. Jenis – jenis *Database*

- a. *Database Relasional* : Jenis ini menggunakan kunci relasional untuk menangani hubungan antara data dan menyimpan data sebagai tabel. Contohnya termasuk *Microsoft SQL Server, PostgreSQL, dan MySQL*.

- b. *Database* Terdistribusi : Data yang tersebar di beberapa server atau lokasi dapat disimpan menggunakan database terdistribusi. Ini menurunkan latensi dan menjamin ketersediaan data yang tinggi (Riyan Dirgantara et al., 2023).

2.3 Teori Khusus

2.3.1 Pelayanan Dan Jasa

Layanan dan kualitas layanan merupakan dua komponen penting yang saling terkait tetapi berbeda dalam konteks manajemen perusahaan. Kualitas layanan mengacu pada bagaimana layanan diberikan kepada pelanggan dan seberapa baik layanan tersebut memenuhi atau melampaui harapan mereka. Di antara elemen yang dinilai dalam konteks ini adalah ketepatan, kecepatan, keramahan, dan kemampuan penyedia layanan untuk mengatasi masalah *klien*. Persepsi konsumen terhadap hubungan perusahaan dengan pelanggannya sering digunakan untuk mengukur kualitas layanan yang diberikan. Semakin baik pengalaman, semakin puas pelanggan (Diza et al., 2021).

Kedua elemen ini penting bagi perusahaan di sektor jasa, seperti PT Jala Lintas Media *Branch* Banten, di mana kepuasan dan loyalitas klien sebagian besar ditentukan oleh kombinasi kualitas layanan yang diberikan dan kualitas layanan yang mereka dapatkan.

Menurut Diza dkk (Diza et al., 2021), Istilah "kualitas layanan" mengacu pada serangkaian langkah yang diambil penyedia layanan untuk memenuhi kebutuhan dan harapan kliennya. Kualitas layanan didefinisikan sebagai tindakan atau aktivitas yang pada dasarnya tidak berwujud dan tidak memberikan kepemilikan kepada entitas lain. Kualitas layanan dapat dinilai

dengan membandingkan harapan pelanggan dengan kesan mereka terhadap layanan yang mereka terima. Jika pelanggan tidak puas dengan layanan yang mereka terima, mereka akan merasa tidak senang. Namun, jika layanannya melebihi harapan, konsumen akan lebih puas.

2.3.2 Kepuasan Pelanggan

Menurut Sambodo Rio Sasongko, Pengembangan nilai pelanggan *berkorelasi* dengan kepuasan pelanggan. Karena menumbuhkan kepuasan pelanggan akan memberikan manfaat bagi bisnis, khususnya hubungan antara bisnis dan kliennya menjadi harmonis, memberikan landasan yang kuat bagi kepuasan pelanggan, dan menghasilkan rekomendasi positif dari mulut ke mulut yang menguntungkan bisnis dengan mendorong pelanggan untuk membeli atau menggunakan layanan bisnis (Sambodo Rio Sasongko, 2021).

Menurut Ali dkk, Kepuasan pelanggan merupakan ukuran yang menentukan seberapa senang pelanggan dengan produk, layanan, dan kemampuan perusahaan. Kepuasan ini dapat dinilai melalui survei dan penilaian yang memungkinkan perusahaan mengetahui cara terbaik untuk meningkatkan atau mengubah produk serta layanan mereka. Kepuasan pelanggan memiliki dampak signifikan terhadap perusahaan dan produknya. Semakin puas pelanggan terhadap kualitas produk dan layanan, semakin besar potensi peningkatan produk dan keuntungan yang dihasilkan (Ali et al., 2021).

2.3.3 Loyalitas Pelanggan

Menurut Sambodo Rio Sasongko, Loyalitas pelanggan adalah komitmen mendalam dari pelanggan untuk terus membeli atau mendukung produk atau

jasa yang disukai di masa depan, meskipun ada pengaruh situasi atau usaha pemasaran yang mungkin memungkinkan mereka untuk beralih. Loyalitas ini biasanya terbentuk melalui kepuasan pelanggan terhadap merek atau layanan yang diterima. Untuk membangun dan memelihara loyalitas, perusahaan harus memperlakukan pelanggan dengan hormat, menjaga hubungan dekat, mengukur kepuasan pelanggan secara berkala, menciptakan biaya peralihan, dan memberikan ekstra nilai berupa hadiah atau penghargaan (Sambodo Rio Sasongko, 2021).

Loyalitas konsumen merujuk pada kesetiaan pelanggan terhadap merek atau perusahaan tertentu, yang terlihat dari keputusan untuk melakukan pembelian berulang secara konsisten. Loyalitas ini tidak hanya terbentuk berdasarkan kualitas produk atau layanan, tetapi juga dari pengalaman positif, hubungan emosional, serta kepercayaan yang dibangun antara pelanggan dan perusahaan. Konsumen yang loyal cenderung tidak mudah terpengaruh oleh harga kompetitif atau promosi dari perusahaan lain karena mereka merasa puas dengan produk, layanan, dan nilai yang mereka peroleh.

2.3.4 Data Mining

Menurut Juandi Dan Hartana, *Data mining* adalah suatu proses penemuan pola atau informasi yang tersembunyi dari kumpulan data besar dengan menggunakan teknik seperti *statistik*, kecerdasan buatan, dan pembelajaran mesin. Tujuan utamanya adalah untuk mengekstrak pengetahuan yang bermanfaat dari data yang tersedia, yang sebelumnya tidak diketahui. Dalam penerapannya, *data mining* dapat digunakan untuk berbagai keperluan seperti analisis pasar, deteksi penipuan, dan prediksi tren masa depan. Proses ini

melibatkan beberapa tahapan seperti pengumpulan data, pembersihan, transformasi, hingga interpretasi hasil (Setiawan dan Wijaya, 2023).

Melalui analisis data dan ekstraksi informasi, teknik ini memungkinkan organisasi untuk membuat keputusan yang tepat. Untuk menemukan pola atau tren dalam data, data *mining* pada dasarnya menerapkan teknik dari berbagai disiplin ilmu, termasuk basis data, statistik, pembelajaran mesin, dan kecerdasan buatan. Dimulai dengan pengumpulan data, pembersihan data, transformasi data, dan analisis, proses data *mining* biasanya terdiri dari banyak komponen utama. Menemukan asosiasi penting, juga pola atau korelasi antara *variabel*, dan yang dapat digunakan untuk memperkirakan perilaku masa depan adalah salah satu tujuan utama data *mining*.

1. Tahapan Data Mining

Pengolahan data terdiri dari beberapa tahapan, dimulai dari pengumpulan data dan berakhir pada interpretasi hasilnya:

- a. Pengumpulan Data: Data untuk proses *mining* dapat diperoleh dari berbagai sumber, seperti basis data, gudang data, atau file log.
- b. Pembersihan Data: Pada langkah ini, data yang kotor, hilang, atau tidak relevan dihapus atau diperbaiki.
- c. Pemilihan Data: Menemukan data yang perlu dianalisis.
- d. Transformasi Data: Untuk penambangan algoritma, diformat data dengan tepat dengan menormalisasi atau mengubah tipe data.
- e. Penggunaan Algoritma *Mining*: Algoritma yang tepat digunakan untuk menemukan pola, hubungan, atau anomali dalam data.
- f. Evaluasi dan Interpretasi: Informasi yang ditemukan dievaluasi untuk memastikan apakah *valid* dan berguna. Selanjutnya, temuan

ini berlaku untuk digunakan dalam membuat keputusan ilmiah atau bisnis.

2. Klasifikasi Data *Mining*

Teknologi Klasifikasi dalam data *mining* adalah proses menempatkan data barang ke dalam kategori yang telah ditetapkan. Untuk tugas ini, algoritma seperti *Decision Tree*, *Naive Bayes*, dan *SVM* digunakan. Mengelompokkan data berdasarkan kesamaan yang ditemukan dalam kumpulan data tertentu disebut *clustering*. Metode ini sering digunakan untuk menemukan pola tersembunyi atau segmen pelanggan dalam data yang tidak diketahui sebelumnya.

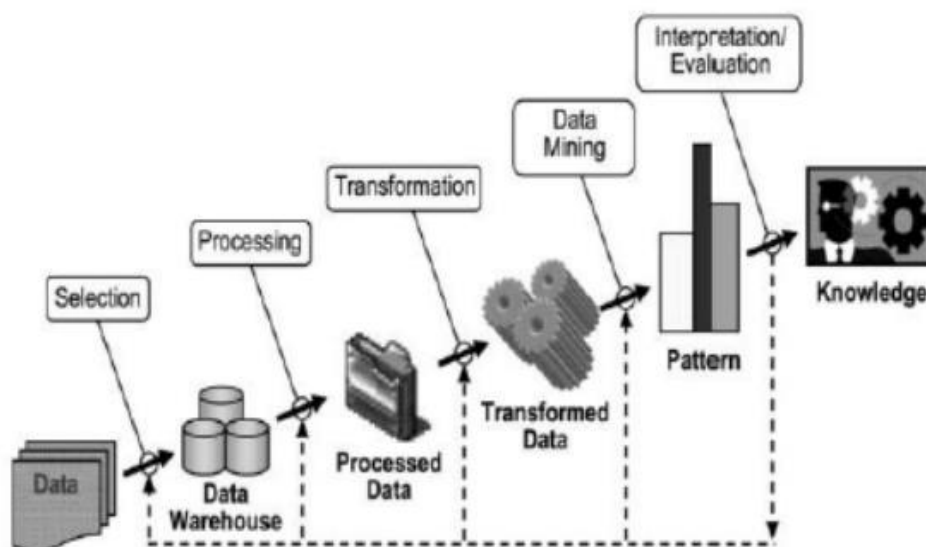
- 1) *Learning Association Rule*: Prosedur untuk menemukan aturan yang berkaitan antara variabel dalam data. Metode regresi, juga dikenal sebagai regresi, digunakan untuk memprediksi bagaimana variabel independen dan variabel dependen berinteraksi satu sama lain.
- 2) Deteksi Anomali: Digunakan untuk menemukan data yang tidak sesuai dengan pola atau norma umum.

Menurut Dwi Winanti, Elinda Revita, dan Efri Yandani, Data mining adalah teknik untuk menemukan pola dan informasi tersembunyi dalam kumpulan data besar menggunakan algoritma canggih dari berbagai bidang seperti *statistik* dan pembelajaran mesin. Dengan demikian, data mining memungkinkan perusahaan atau organisasi untuk menemukan wawasan penting yang sebelumnya tersembunyi dalam data mereka, yang kemudian bisa digunakan untuk pengambilan keputusan yang lebih baik.

Proses data *mining* merupakan bagian dari *Knowledge Discovery in Databases* (KDD), yang terdiri dari beberapa tahapan utama:

1. *Data Selection*: Proses ini melibatkan pemilihan dataset yang relevan untuk dianalisis.
2. *Preprocessing* dan *Cleaning Data*: Membersihkan data dari inkonsistensi dan duplikasi serta memperbaiki kesalahan data.
3. Transformasi: Mengubah data ke dalam format yang lebih tepat untuk proses *mining*.
4. *Data Mining*: Mencari pola atau informasi menarik menggunakan metode atau algoritma tertentu.
5. Interpretasi/Evaluasi: Mengevaluasi pola yang ditemukan untuk memastikan relevansi dan akurasi hasilnya.

Proses pengumpulan informasi dari kumpulan data menggunakan metode dan algoritma yang terkait dengan sistem manajemen basis data, pembelajaran mesin, dan statistik dikenal sebagai penambangan data. Kumpulan data yang besar dapat memiliki informasi tersembunyi yang berharga yang diekstraksi menggunakan penambangan data (Winarti et al., 2021).



Gambar 2. 1 Tahapan proses data *mining*

Sumber: (Winarti et al., 2021)

Menurut Egye Santoso dan Susanto Hariyanto, Data *mining* adalah proses dalam ilmu komputer yang memanfaatkan teknik statistik, pembelajaran mesin, visualisasi data, pengenalan pola, dan *database* untuk mengekstraksi informasi berharga dari kumpulan data besar. Tujuan utamanya adalah menemukan pola atau hubungan yang tersembunyi dalam data yang kemudian diolah menjadi informasi yang bermakna dan bermanfaat.

Beberapa metode utama dalam data mining meliputi:

1. Klasifikasi (*Classification*): Teknik ini memetakan data ke dalam beberapa kelas yang telah didefinisikan sebelumnya.
2. Asosiasi (*Association*): Mengidentifikasi hubungan antara item atau kejadian tertentu dalam kumpulan data.
3. *Clustering*: Proses pengelompokan data berdasarkan kesamaan atribut, yang sering disebut sebagai *unsupervised learning*. Dalam *clustering*, data dibagi menjadi beberapa kelompok atau *cluster* berdasarkan kemiripan sehingga kesamaan dalam satu kelompok adalah maksimum, sedangkan kesamaan antar kelompok berbeda adalah minimum.
4. Deskripsi: Menyajikan pola atau kecenderungan dari data yang berguna sebagai informasi tambahan.
5. Prediksi (*Forecasting*): Menggunakan data historis untuk memperkirakan nilai di masa depan, seperti memprediksi perubahan nilai saham (Santoso dan Hariyanto, 2022).

2.3.5 Algoritma

Menurut Pinkan Indriani Daulay dan Yahfizham, Algoritma adalah serangkaian tindakan *metodis* yang dimaksudkan untuk menggunakan logika terstruktur

guna memecahkan suatu tugas atau masalah. Algoritma diimplementasikan sebagai instruksi berurutan dalam pemrograman komputer untuk mengatasi berbagai masalah pemrograman. Algoritma sering digunakan dalam ilmu komputer untuk penggabungan, *hashing*, pencarian *biner* dan *linear*, dan *kompresi* data. Masing-masing metode ini memiliki tujuan yang berbeda, seperti menggunakan pencarian *biner* untuk menemukan elemen dalam data terstruktur atau strategi *kompresi* untuk mengurangi ukuran data. Algoritma pada dasarnya memungkinkan prosedur logis yang memungkinkan komputer memecahkan masalah rumit dengan cepat dan efektif (Pinkan Indriani Daulay dan Yahfizham, 2023).

Dalam teknik informatika, algoritma digunakan untuk berbagai fungsi, seperti pengurutan data (*sorting*), pencarian (*searching*), hingga klasifikasi data. Misalnya, dalam konteks data mining, algoritma membantu dalam mengidentifikasi pola atau hubungan tertentu dalam data untuk menghasilkan prediksi atau rekomendasi. Algoritma juga dapat diimplementasikan melalui bahasa pemrograman agar dapat berinteraksi langsung dengan perangkat lunak atau sistem lain.

Untuk tugas yang memerlukan prediksi, seperti pada penelitian dengan algoritma *Decision Tree*, algoritma ini memungkinkan pemetaan keputusan yang optimal berdasarkan variabel-variabel yang relevan, sehingga pengguna dapat mendapatkan hasil prediksi yang sesuai dengan data yang dimiliki.

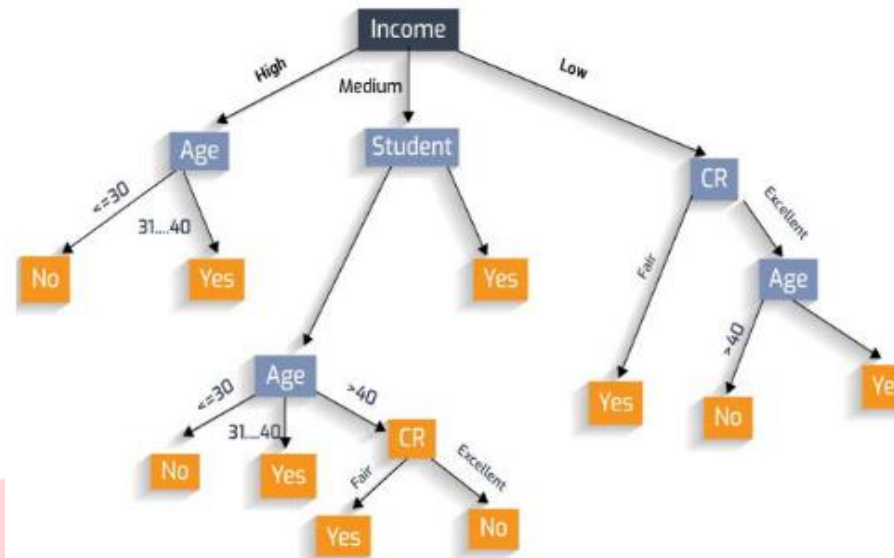
Menurut Bata Mahesh, algoritma memiliki karakteristik unik yang dapat disesuaikan dengan jenis data dan tujuan klasifikasi atau prediksi tertentu, tergantung pada apakah data tersebut bersifat terstruktur atau tidak, *linear* atau *non-linear*, serta skala datasetnya. Contohnya seperti algoritma *Decision Tree*,

algoritma ini menggunakan struktur pohon untuk membuat keputusan berdasarkan atribut data. Setiap *node* dalam pohon mewakili atribut tertentu yang harus diklasifikasikan, sementara setiap cabang menunjukkan nilai yang dapat diambil oleh atribut tersebut. Proses ini berlanjut hingga mencapai hasil yang diinginkan atau mencapai klasifikasi tertentu. Algoritma ini diilustrasikan dalam bentuk *pseudo-code* untuk membangun pohon keputusan yang dapat dipahami dengan mudah (Mahesh, 2020).

2.3.6 *Decision Tree*

Menurut Bahzad Taha Jijo¹, dan Adnan Mohsin Abdulazeez *Decision Tree* adalah metode klasifikasi berbasis pohon yang sangat populer dalam bidang ilmu data dan *machine learning*. Secara struktural, algoritma ini memanfaatkan cabang-cabang pohon untuk menyajikan aturan keputusan yang dipakai dalam menentukan kategori data berdasarkan atribut tertentu. Setiap *node* pada pohon melambangkan *atribut*, dan setiap cabang menggambarkan keputusan atau kondisi yang mengarahkan data ke kategori tertentu.

Prinsip kerja utama *Decision Tree* adalah membuat pemisahan atau "*split*" pada data berdasarkan atribut yang paling relevan, yang dinilai dengan *entropi* atau *gain* informasi. Algoritma ini cocok untuk digunakan pada tugas klasifikasi, di mana model dibangun untuk memprediksi kelas atau kategori data dengan akurasi yang tinggi. Salah satu kelebihan adalah hasil *klasifikasi Decision Tree* mudah dipahami, karena setiap cabang menunjukkan proses pengambilan keputusan yang logis dan terstruktur (Taha Jijo dan Abdulazeez Mohsin, 2021).



Gambar 2. 2 Contoh *Desision Tree* (Pohon Keputusan)

Sumber: (Taha Jijo dan Abdulazeez Mohsin, 2021)

Cara kerja pohon keputusan adalah dengan terus-menerus membagi informasi menurut karakteristik yang menawarkan pemisahan terbaik. Dengan menggunakan teknik ini, pohon keputusan menghasilkan model prediktif yang memungkinkan kategorisasi data baru menggunakan pola yang dipelajari dari data pelatihan sebelumnya. Manfaat utamanya adalah keserbagunaannya dalam menangani data kategoris dan numerik, serta kemudahan interpretasinya.

Dimulai dengan pengumpulan dan pembersihan data dan diakhiri dengan pembuatan model pohon keputusan, metode Pohon Keputusan memiliki beberapa tahap.

Langkah-langkah dan rumus yang digunakan adalah sebagai berikut:

1. Pengumpulan dan Pembersihan Data: Kumpulkan informasi yang relevan dengan masalah atau tujuan yang ingin dicapai. Selanjutnya, bersihkan data dengan menghilangkan informasi yang tidak lengkap, *outlier*, atau tidak perlu.

2. Pemilihan Atribut Utama: Pada langkah ini, perolehan informasi atau indeks *Gini* untuk setiap atribut ditentukan untuk memilih atribut utama. *Node* akar akan memiliki karakteristik dengan nilai tertinggi.
3. Pembentukan Cabang dan *Node*: Data akan dibagi menjadi banyak cabang berdasarkan nilai atribut utama yang dipilih. Setelah itu, setiap cabang menjadi simpul baru, yang jika data memungkinkan, dapat dibagi menjadi cabang berikutnya.
4. Pemangkasan (*Pruning*): Setelah pohon selesai dibuat, pemangkasan dilakukan untuk mengurangi kompleksitas pohon. Pemangkasan dapat membantu mengurangi risiko *overfitting* dengan menghilangkan cabang yang tidak terlalu penting atau yang memiliki nilai prediksi rendah.
5. Evaluasi dan Validasi Model: Untuk menjamin akurasi prediksi, model diverifikasi menggunakan data uji setelah pembuatan pohon keputusan. Tujuan evaluasi adalah untuk menentukan seberapa baik model memprediksi data.

Rumus pada *Decision Tree*:

1. *Entropy*

Entropy (S): Entropy adalah ukuran ketidakpastian atau ketidakaturan dalam sebuah kumpulan data. Digunakan untuk menentukan seberapa baik data dapat dipisahkan berdasarkan kategori tertentu. Entropy akan rendah jika data dapat diklasifikasikan dengan jelas dalam kategori tertentu dan tinggi jika data lebih acak atau tidak pasti.

Menurut Putri dkk (T. A. Q. Putri et al., 2023) rumus dari *entropy* dan *gain* adalah:

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

Keterangan:

- 3) S Himpunan data atau subset data kasus
- 4) N Jumlah partisi S
- 5) Pi Proporsi atau probabilitas dari kategori atau kelas

2. Gain

Information Gain: Information *gain* mengukur seberapa banyak informasi yang didapatkan dari pemilihan atribut tertentu untuk membagi data. Nilai information gain yang lebih tinggi menunjukkan bahwa atribut tersebut lebih baik dalam mengurangi ketidakpastian. Rumusnya:

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

Keterangan:

- a. S Himpunan kasus
- b. A Atribut
- c. N Jumlah partisi atribut A
- d. |Si| Jumlah kasus pada partisi ke-i
- e. |S| Jumlah kasus dalam S

2.3.7 MySQL

Menuurt Ivan Suster dan Tamara ranisavljevic, *MySQL* adalah sistem manajemen basis data relasional yang populer, dikenal karena kecepatan, ketangguhan, dan skalabilitasnya. Dikembangkan oleh *oracle corporation*, *MySQL* bersifat *open-source* dan tersedia secara gratis di bawah lisensi *GNU General Public License*. Banyak perusahaan besar seperti Facebook, Google, dan Twitter, serta bisnis kecil dan pengembang independen, mengandalkan

MySQL untuk mengelola data mereka menggunakan *MySQL* (Šušter dan Ranisavljević, 2023).

2.3.8 *Python*

Python menurut Fedor V. Ryzhkov, Yuliya E. Ryzhkova, dan Michail N. Elinson, *Python* adalah bahasa pemrograman yang semakin populer di berbagai bidang, terutama dalam sains, karena *fleksibilitas*, kemudahan penggunaannya, dan ekosistem pustaka yang kaya. *Python* memiliki pustaka seperti *NumPy*, *Pandas*, dan *Matplotlib* yang berguna untuk pemrosesan data besar dan visualisasi dalam. Pustaka ini sangat membantu dalam menganalisis data, memahami hasil perhitungan, dan memvisualisasikan data yang kompleks (Ryzhkov et al., 2023). *Python* sering digunakan dalam penambangan data, *machine learning*, analisis data, dan pengembangan aplikasi. Karena sintaksnya yang sederhana, yang membuat pembuatan kode lebih efektif dan langsung, bahasa ini ideal untuk pemula. *Pandas*, *Numpy*, dan *Scikit-Learn* termasuk beberapa pustaka dan kerangka kerja yang bermanfaat untuk penambangan dan analisis data yang disediakan *Python*. Alat-alat ini dapat digunakan untuk memproses dan menganalisis data klien guna menemukan tren dalam tingkat kepuasan pelanggan.

2.3.9 *Aplikasi Web*

Menurut Dinda Puspita Hindriyani dan Rizki Elisa Nalawati, Aplikasi website adalah kumpulan halaman yang menampilkan berbagai informasi dan elemen multimedia seperti gambar, suara, atau video yang saling terhubung melalui *hyperlink*. Aplikasi ini memberikan akses yang lebih mudah bagi pengguna

untuk menemukan dan memanfaatkan informasi yang mereka cari. Web *semantik* yang diterapkan pada aplikasi ini memungkinkan data di dalamnya diproses oleh mesin untuk memahami bahasa manusia secara lebih baik, sehingga pencarian informasi menjadi lebih akurat dan relevan. Melalui aplikasi website, pengguna dapat berinteraksi dengan sistem secara dinamis, termasuk melakukan pencarian, registrasi, hingga transaksi secara online (Hindriyani dan Elisa Nalawati, 2021).

Teknologi internet digunakan oleh aplikasi *web* untuk menawarkan layanan kepada pengguna secara *daring*. Beberapa bahasa pemrograman digunakan dalam pengembangan aplikasi ini untuk menghasilkan antarmuka pengguna (UI), fungsionalitas logis, serta manajemen dan penyimpanan data yang efektif. Aplikasi web menggunakan *back-end* (seperti *Python* dengan *Flask* atau *Django*, *PHP*, atau *Node.js*) untuk pemrosesan data dan logika aplikasi, dan *front-end* (*HTML*, *CSS*, dan *JavaScript*) untuk tampilan antarmuka.

2.4 Tinjauan Studi

Pada bagian ini dilakukan tinjauan studi dalam menunjukan penelitian yang diusulkan dengan sesuai pada bidang pengkajian. Dengan pendukung minimal 10 jurnal nasional ber p-ISSN atau e-ISSN dan minimal 5 jurnal internasional. Dengan maksimal 5 tahun terakhir sejak melakukan penyusunan.

2.4.1 Penerapan Algoritma *Decision Tree* C4.5 Untuk Prediksi Kelayakan Calon Pendoror Darah Dengan Klasifikasi Data *Mining*

1. Jurnal Yuda Irawan

2. Judul

Penerapan Algoritma *Decision Tree* C4.5 Untuk Prediksi Kelayakan Calon Pendoror Darah Dengan Klasifikasi Data *Mining*

3. Jurnal

Jurnal Teknologi Informasi dan Multimedia (JTIM)

4. Volume, Nomor dan Halaman

Vol. 2, No. 4, hal. 181-189

5. Bulan dan Tahun

Februari 2021

6. Penulis

Yuda Irawan

7. Penerbit

STMIK Hang Tuah Pekanbaru

8. Tujuan penelitian

Penelitian ini bertujuan untuk membuat sebuah aplikasi prediksi yang dapat menentukan kelayakan calon pendonor darah menggunakan algoritma *Decision Tree* C4.5, sehingga PMI dapat memiliki sistem prediktif yang lebih akurat dalam menyaring calon pendonor berdasarkan data yang valid.

9. Lokasi dan Subjek Penelitian

Lokasi penelitian adalah di Kabupaten Kampar, dengan subjek penelitian yaitu calon pendonor darah yang data kesehatannya dianalisis untuk kelayakan mendonorkan darah mereka.

10. Metode yang digunakan

Metode klasifikasi menggunakan algoritma *Decision Tree* C4.5 diterapkan untuk membangun model prediktif. Dalam pengembangan aplikasi, penulis menggunakan metodologi rekayasa perangkat lunak *Waterfall* dengan tahapan analisis, desain, pengkodean, pengujian, dan implementasi aplikasi.

11. Perancangan Sistem

Sistem dirancang menggunakan diagram UML untuk memodelkan *use case* dan *activity*, dan diimplementasikan dengan bahasa pemrograman *PHP* serta *MySQL* untuk mengelola *database*. Sistem ini mengumpulkan data variabel seperti usia, berat badan, *hemoglobin*, dan tekanan darah calon pendonor.

12. Hasil Penelitian

Aplikasi data mining yang dikembangkan dapat mengklasifikasikan kelayakan calon pendonor darah berdasarkan variabel kritis kesehatan, dengan hasil menunjukkan bahwa *hemoglobin* memiliki pengaruh tertinggi. Sistem ini memudahkan proses pemilahan calon pendonor yang layak dengan cepat.

13. Kekuatan Penelitian

Penelitian ini memberikan kontribusi signifikan dalam meningkatkan efisiensi proses seleksi pendonor darah. Sistem klasifikasi ini memanfaatkan metode *Decision Tree* untuk menghasilkan akurasi yang lebih tinggi dalam prediksi kelayakan donor berdasarkan data kesehatan yang kritis.

14. Kelemahan Penelitian

Penelitian hanya menggunakan satu algoritma klasifikasi, yaitu *Decision Tree* C4.5, sehingga belum terdapat perbandingan performa dengan metode

lainnya yang mungkin juga mampu menghasilkan prediksi dengan hasil akurasi yang lebih tinggi atau waktu pemrosesan yang lebih singkat.

15. Kesimpulan

Aplikasi yang dikembangkan mampu mengklasifikasikan calon pendonor darah secara efisien berdasarkan *variabel* kesehatan utama, khususnya *hemoglobin* sebagai faktor dominan. Penelitian ini menunjukkan bahwa metode *Decision Tree* C4.5 dapat diterapkan secara efektif dalam klasifikasi donor (Irawan, 2021).

2.4.2 Implementasi Algoritma *Decision Tree* dan *Naïve Bayes* Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan *Starbucks*

1. Jurnal Tiyas Asih Qurnia Putri, Agung Triayudi, Rima Tamara Aldisa

2. Judul

Implementasi Algoritma *Decision Tree* dan *Naïve Bayes* Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan *Starbucks*

3. Jurnal

Journal of Information System Research (JOSH)

4. Volume, Nomor dan Halaman

Vol. 4, No. 2, hal. 641-649

5. Bulan dan Tahun

Januari 2023

6. Penulis

Tiyas Asih Qurnia Putri, Agung Triayudi, Rima Tamara Aldisa

7. Penerbit

Universitas Nasional, Jakarta

8. Tujuan penelitian

Penelitian ini bertujuan untuk mengklasifikasikan kepuasan pelanggan *Starbucks* menjadi tiga kategori (puas, cukup puas, dan tidak puas) dengan menggunakan algoritma *Decision Tree* dan *Naïve Bayes*, serta untuk mengetahui algoritma terbaik dalam klasifikasi ini.

9. Lokasi dan Subjek Penelitian

Lokasi penelitian berbasis data dari *Twitter*, dengan subjek penelitian yaitu kepuasan pelanggan *Starbucks* yang dianalisis dari komentar pengguna *Twitter* terkait *Starbucks*.

10. Metode Yang Digunakan

Penelitian menggunakan algoritma *Decision Tree* dan *Naïve Bayes* untuk klasifikasi *sentimen*, serta proses *Agile Software Development* untuk pengembangan sistem yang berulang (*iteratif*) melalui tahapan pengujian dan evaluasi data *Twitter* dengan *Python* dan *Streamlit*.

11. Perancangan Sistem

Sistem dirancang dengan tahapan *pre-processing* seperti *case folding*, *cleaning*, *tokenizing*, *stopword removal*, *stemming*, dan *labelling* untuk mengklasifikasikan *tweet* menjadi *sentimen* positif, negatif, dan netral. Data dikumpulkan dengan metode *crawling* melalui API.

12. Hasil Penelitian

Hasil klasifikasi menunjukkan bahwa 19,2% dari *tweet* pelanggan bersentimen positif, 70,3% netral, dan 10,4% negatif. Algoritma *Decision*

Tree menghasilkan akurasi tertinggi, yaitu 83%, dibandingkan dengan *Naïve Bayes* yang menghasilkan akurasi 74%.

13. Kekuatan Penelitian

Penelitian ini berhasil mengekstraksi sentimen pelanggan secara efisien dan memberikan model klasifikasi yang kuat untuk membantu bisnis dalam memahami kepuasan pelanggan dengan lebih baik, khususnya dalam konteks kinerja algoritma *Decision Tree* yang lebih akurat.

14. Kelemahan Penelitian

Penelitian ini terbatas pada data dari *Twitter* yang mungkin tidak mencakup semua jenis pelanggan, dan model hanya mengevaluasi dua algoritma klasifikasi sehingga belum ada perbandingan dengan metode lain seperti *SVM* atau *k-NN* yang mungkin memberikan hasil berbeda.

15. Kesimpulan

Penelitian ini menunjukkan bahwa algoritma *Decision Tree* menghasilkan akurasi lebih tinggi dalam mengklasifikasikan sentimen pelanggan *Starbucks* dibandingkan *Naïve Bayes*, sehingga algoritma ini lebih disarankan untuk memprediksi kepuasan pelanggan (T. A. Q. Putri et al., 2023).

2.4.3 Analisis Komparasi Algoritma Klasifikasi Data *Mining* dalam Klasifikasi *Website Phishing*

1. Jurnal Nabila Bianca Putri dan Arie Wahyu Wijayanto

2. Judul

Analisis Komparasi Algoritma Klasifikasi Data *Mining* dalam Klasifikasi *Website Phishing*

3. Jurnal

Komputika: Jurnal Sistem Komputer

4. Volume, Nomor dan Halaman

Vol. 11, No. 1, hlm. 59 - 66

5. Bulan dan Tahun

April 2022

6. Penulis

Nabila Bianca Putri, Arie Wahyu Wijayanto

7. Penerbit

Universitas Komputer Indonesia (UNIKOM)

8. Tujuan Penelitian

Membandingkan akurasi, presisi, dan sensitivitas dari beberapa algoritma klasifikasi data *mining* untuk menemukan algoritma terbaik dalam mengklasifikasi *website phishing*.

9. Lokasi dan Subjek Penelitian

Dataset terdiri dari 1.353 data *website* (702 data situs *non-phishing*, 103 data situs mencurigakan, dan 548 data situs *phishing*)

10. Metode Yang Digunakan

Metode komparatif dengan algoritma *Naïve Bayes*, *Random Forest*, *Decision Tree*, dan *Support Vector Machine (SVM)*

11. Perancangan Sistem

Pengumpulan dataset Pengujian dan evaluasi model menggunakan validasi silang dan *confusion matrix* untuk akurasi, presisi, dan *sensitivitas*

12. Hasil Penelitian

Algoritma *Random Forest* memiliki akurasi tertinggi (90,77%) dibandingkan algoritma lainnya, diikuti oleh *SVM*, *Decision Tree*, dan *Naïve Bayes*

13. Kekuatan Penelitian

Menggunakan dataset yang cukup besar (1.353 data) Menggunakan beberapa algoritma umum dalam klasifikasi data mining untuk perbandingan

14. Kelemahan Penelitian

Belum menguji algoritma lain yang mungkin memberikan hasil lebih baik
Rentan terhadap hasil yang terbatas pada dataset tertentu

15. Kesimpulan

Algoritma *Random Forest* lebih baik dalam klasifikasi *website phishing* dengan akurasi, presisi, dan *sensitivitas* tertinggi dibandingkan algoritma lainnya. Model ini dinilai cocok untuk klasifikasi *website phishing*, dan disarankan untuk penelitian lanjutan dengan algoritma berbeda guna memperkaya hasil (N. B. Putri dan Wijayanto, 2022).

2.4.4 Implementasi Algoritma *Decision Tree* dalam Klasifikasi Kompetensi Siswa

1. **Jurnal Hilman Rifa'i dkk**

2. **Judul**

Implementasi Algoritma *Decision Tree* dalam Klasifikasi Kompetensi Siswa

3. **Jurnal**

KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer

4. **Volume, Nomor Dan Halaman**

Vol. 06, No. 01, hlm. 15-20

5. **Bulan Dan Tahun**

Februari 2020

6. **Penulis**

Hilman Rifa'i, Ryan Hamonangan, Dian Ade Kurnia, Kaslani, Mulyawan

7. Penerbit

STMIK IKMI Cirebon

8. Tujuan Penelitian

Menerapkan algoritma *Decision Tree* untuk mengklasifikasi kompetensi siswa dan menghasilkan rekomendasi yang mengidentifikasi siswa yang kompeten dan tidak kompeten.

9. Lokasi dan Subjek Penelitian

Siswa di SMK PGRI Jatiwangi, Majalengka

10. Metode Yang Digunakan

Metode *Knowledge Discovery in Database (KDD)* dan algoritma *Decision Tree* untuk klasifikasi data siswa

11. Perancangan Sistem

Data *collection* dan *cleaning* Data *transformation* ke dalam nilai numerik
Penggunaan algoritma *Decision Tree* untuk klasifikasi data, menggunakan *RapidMiner* untuk penerapan klasifikasi

12. Hasil Penelitian

Hasil akurasi model klasifikasi *Decision Tree* mencapai 76,96%, dengan hasil klasifikasi yang mengidentifikasi 111 siswa kompeten dan 46 siswa tidak kompeten

13. Kekuatan Penelitian

Menggunakan data siswa aktual dari hasil ujian Algoritma *Decision Tree* memberikan hasil yang dapat dipahami dengan baik dan digunakan untuk evaluasi oleh sekolah

14. Kelemahan Penelitian

Terbatas pada klasifikasi berdasarkan data nilai, tidak mencakup aspek lain
Hasil prediksi mungkin terbatas pada parameter yang digunakan, sehingga
perlu diperluas untuk hasil yang lebih akurat

15. Kesimpulan

Algoritma *Decision Tree* efektif dalam mengidentifikasi kompetensi siswa
berdasarkan data nilai mereka. Metode ini dapat dijadikan bahan evaluasi bagi
pihak sekolah dan direkomendasikan untuk analisis kompetensi lebih lanjut
menggunakan algoritma lain (Rifa'i et al., 2022).

2.4.5 Implementasi Data Mining untuk Menentukan Minat Siswa dalam Menentukan Jurusan Pada Perguruan Tinggi

1. Jurnal Saeful Bahri

2. Judul

Implementasi Data Mining untuk Menentukan Minat Siswa dalam
Menentukan Jurusan pada Perguruan Tinggi

3. Jurnal

Jurnal Sistem Informasi (JUSIN)

4. Volume, Nomor dan Halaman

Vol. 3, No. 1, hlm. 23-33

5. Bulan Dan Tahun

Januari - Juni 2022

6. Penulis

Saeful Bahri

7. Penerbit

Institut Teknologi dan Bisnis Ahmad Dahlan, Jakarta

8. Tujuan Penelitian

Untuk mengklasifikasi minat siswa dalam memilih jurusan di perguruan tinggi menggunakan algoritma *Decision Tree*, *Naive Bayes*, dan *K-Nearest Neighbor*, sehingga dapat membantu siswa memilih jurusan sesuai minat dan bakat.

9. Lokasi dan Subjek Penelitian

Beberapa perguruan tinggi di wilayah Tangerang, dengan subjek siswa SMA kelas XII yang mengalami kesulitan memilih jurusan

10. Metode Yang Digunakan

CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang mencakup pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran hasil

11. Perancangan Sistem

Penggunaan algoritma *Decision Tree* (C4.5) untuk klasifikasi Algoritma *Naive Bayes* untuk klasifikasi berdasarkan probabilitas Algoritma *K-Nearest Neighbor* berdasarkan jarak data

12. Hasil Penelitian

Algoritma *Decision Tree* memberikan akurasi tertinggi sebesar 75,38%, diikuti oleh *K-Nearest Neighbor* (71,67%), sedangkan *Naive Bayes* memiliki akurasi lebih rendah sebesar 60,98%. Faktor utama yang mempengaruhi tujuan pemilihan adalah pengaruh dari diri sendiri, teman, atau orang tua

13. Kekuatan Penelitian

Menggunakan beberapa algoritma untuk perbandingan Memiliki hasil klasifikasi yang jelas bagi siswa sesuai minat dan bakat

14. Kelemahan Penelitian

Akurasi *Naive Bayes* yang rendah Tidak mengukur faktor eksternal lain yang mungkin mempengaruhi minat siswa

15. Kesimpulan

Algoritma *Decision Tree* paling efektif dalam membantu siswa memilih mata pelajaran berdasarkan minat. Penelitian ini diharapkan dapat membantu perguruan tinggi dalam memberikan arah sesuai minat siswa (Bahri, 2022).

2.4.6 Penerapan Data Mining Clustering terhadap Perekonomian pada Kelurahan Wonorejo Kecamatan Marpoyan Damai Provinsi Riau Menggunakan Metode Algoritma *K-Means*

1. Jurnal Egye Santoso, Susanto Hariyanto

2. Judul

Penerapan Data *Mining Clustering* terhadap Perekonomian pada Kelurahan Wonorejo Kecamatan Marpoyan Damai Provinsi Riau Menggunakan Metode Algoritma *K-Means*

3. Jurnal

Akselerator

4. Volume, Nomor dan Halaman

Vol. 4, No. 1, hlm. 50-64

5. Bulan dan Tahun

Januari 2021

6. Penulis

Egye Santoso, Susanto Hariyanto

7. Penerbit

Universitas Buddhi Dharma, Tangerang

8. Tujuan Penelitian

Menggunakan metode *clustering K-Means* untuk mengklasifikasi kelompok perekonomian masyarakat di Kelurahan Wonorejo sebagai langkah membantu distribusi bantuan sosial yang lebih tepat sasaran.

9. Lokasi dan Subjek Penelitian

Wilayah Kelurahan Wonorejo, Kecamatan Marpoyan Damai, Kota Pekanbaru; subjek adalah data keluarga terdampak pandemi *COVID-19*

10. Metode Yang Digunakan

Data Mining dengan pendekatan *clustering* menggunakan algoritma *K-Means*, serta metode *Knowledge Discovery Databases (KDD)* yang mencakup seleksi data, *cleaning*, *transformasi*, dan evaluasi hasil *clustering*

11. Perancangan Sistem

Pengumpulan data keluarga terdampak COVID-19 Seleksi dan transformasi atribut seperti pendapatan, jumlah tanggungan, dan pekerjaan Proses *clustering* menggunakan *K-Means* dengan *RapidMiner* sebagai alat bantu

12. Hasil Penelitian

Hasil *clustering* menunjukkan tiga kategori: masyarakat mampu, kurang mampu, dan tidak mampu. Dari 1.583 data, kategori 'tidak mampu' terdiri dari 456 keluarga, 'kurang mampu' 583 keluarga, dan 'mampu' 544 keluarga.

13. Kekuatan Penelitian

Mampu mengelompokkan data dengan cepat membantu penyaluran bantuan sosial yang lebih efektif berdasarkan kategori ekonomi

14. Kelemahan Penelitian

Metode evaluasi *cluster* masih terbatas Perlu penyempurnaan pada logika pemrograman untuk analisis lebih lanjut pada aplikasi

15. Kesimpulan

Metode *K-Means* efektif dalam mengelompokkan kondisi ekonomi masyarakat, yang dapat membantu penyaluran bantuan sosial pada kelompok yang benar-benar membutuhkan (Santoso dan Hariyanto, 2022).

2.4.7 Penerapan Data *Mining* untuk Penentuan Paket Penjualan Produk Pecah Belah Menggunakan Algoritma *Apriori* (Studi Kasus: Toko Sumber Cahaya)

1. Jurnal Juandi Setiawan, Hartana Wijaya

2. Judul

Penerapan Data *Mining* untuk Penentuan Paket Penjualan Produk Pecah Belah Menggunakan Algoritma *Apriori* (Studi Kasus: Toko Sumber Cahaya)

3. Jurnal

AKSELERATOR

4. Volume, Nomor dan Halaman

Vol. 4, No. 2, Hal. 31–41

5. Bulan dan Tahun

Tidak tercantum bulan; Tahun 2023

6. Penulis

Juandi Setiawan, Hartana Wijaya

7. Penerbit

Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

8. Tujuan Penelitian

Menggunakan algoritma *Apriori* untuk mengidentifikasi pola pembelian produk dan membantu keputusan pengelolaan stok di Toko Sumber Cahaya

9. Lokasi dan Subjek Penelitian

Toko Sumber Cahaya, Kota Tangerang; Subjek: produk pecah belah

10. Metode Yang Digunakan

Metode *CRISP-DM* untuk data *mining*

11. Perancangan Sistem

Tahapan *CRISP-DM*: memahami bisnis, memahami data, persiapan data, pemodelan, evaluasi, dan implementasi

12. Hasil Penelitian

Ditemukan pola transaksi yang dapat dimanfaatkan untuk strategi pemasaran dan pengelolaan stok barang

13. Kekuatan Penelitian

Mempermudah toko memahami pola penjualan, membantu dalam pengambilan keputusan persediaan stok dengan data *real-time*

14. Kelemahan Penelitian

Perlu dataset yang lebih relevan dan besar agar hasil lebih akurat

15. Kesimpulan

Algoritma *Apriori* membantu memahami pola transaksi, memberikan informasi yang bermanfaat untuk promosi dan pengelolaan stok di toko (Setiawan dan Wijaya, 2023).

2.4.8 Analisis Perancangan Sistem untuk Kepuasan Pelanggan pada UD. Shinta Elektronik Dengan Menggunakan Metode Algoritma C4.5

1. Jurnal Steven Chris Limong, Edy, Dera Susilawati

2. Judul

Analisis Perancangan Sistem untuk Kepuasan Pelanggan pada UD. Shinta Elektronik dengan Menggunakan Metode Algoritma C4.5

3. Jurnal

Jurnal *ALGOR*

4. Volume, Nomor dan Halaman

Vol. 1, No. 2, Hal. 52-58

5. Bulan dan Tahun

Mei 2020

6. Penulis

Steven Chris Limong, Edy, Dera Susilawati

7. Penerbit

Universitas Buddhi Dharma

8. Tujuan Penelitian

Menilai tingkat kepuasan pelanggan pada UD. Shinta *Electronic* dengan menggunakan algoritma C4.5 untuk mengklasifikasi pelanggan yang puas atau tidak terhadap layanan

9. Lokasi dan Subjek Penelitian

UD. Shinta *Elektronik*, Tangerang; Subjek: Pelanggan dan data penjualan

10. Metode Yang Digunakan

Algoritma C4.5 *Decision Tree* dan aplikasi *mobile*

11. Perancangan Sistem

Sistem berbasis aplikasi *Android* untuk *survei aftersales* dan mengumpulkan data kepuasan, kemudian dianalisis menggunakan *RapidMiner* dengan algoritma C4.5

12. Hasil Penelitian

Dengan algoritma C4.5 dan analisis *RapidMiner*, akurasi hasil survei kepuasan pelanggan mencapai 93,33%

13. Kekuatan Penelitian

Tingkat akurasi yang tinggi (93,33%) Sistem membantu pihak UD. Shinta untuk memahami kebutuhan pelanggan dan meningkatkan layanan

14. Kelemahan Penelitian

Data pelanggan yang digunakan tidak mencakup detail nama pelanggan, sehingga data mungkin kurang lengkap

15. Kesimpulan

Algoritma C4.5 efektif dalam mengevaluasi kepuasan pelanggan UD. Shinta *Electronic*, membantu pengambilan keputusan bisnis yang lebih akurat melalui data analisis yang efisien dan dapat diandalkan (Chris et al., 2020)

2.4.9 Aplikasi Klasifikasi Kepribadian Manusia Menggunakan Algoritma *Decision Tree* (C4.5) Berbasis Web

1. Jurnal Kent Christopher, Desiyanna Lasut, Lianny W. Kusuma

2. Judul

Aplikasi Klasifikasi Kepribadian Manusia Menggunakan Algoritma *Decision Tree* (C4.5) Berbasis Web

3. Jurnal

Jurnal ALGOR

4. Volume, Nomor dan Halaman

Vol. IV, No. I, Hal. 79-86

5. Bulan dan Tahun

Mei 2022

6. Penulis

Kent Christopher, Desiyanna Lasut, Lianny W. Kusuma

7. Penerbit

Universitas Buddhi Dharma

8. Tujuan penelitian

Membuat aplikasi berbasis *web* untuk mengklasifikasi jenis kepribadian manusia dengan Algoritma C4.5, sehingga pengguna dapat mengetahui karakter, kelebihan, dan kekurangan mereka

9. Lokasi dan Subjek Penelitian

Tidak tercantum lokasi spesifik; Subjek: Pengguna aplikasi kepribadian

10. Metode yang digunakan

Algoritma Decision Tree C4.5

11. Perancangan Sistem

Aplikasi berbasis *web* dengan antarmuka pengguna untuk klasifikasi kepribadian melalui pengisian survei, yang memanfaatkan Algoritma C4.5 untuk menganalisis data pengguna

12. Hasil Penelitian

Aplikasi ini membantu pengguna mengetahui tipe kepribadian mereka dengan akurasi *klasifikasi* antara 90-95%

13. Kekuatan Penelitian

Memberikan hasil *klasifikasi* yang cepat dan mudah Akurasi hasil *klasifikasi* yang tinggi (90-95%)

14. Kelemahan Penelitian

Masih perlu pembaruan tampilan agar lebih menarik, perlu dilakukan uji coba dengan algoritma lain untuk perbandingan akurasi

15. Kesimpulan

Algoritma C4.5 berhasil diterapkan untuk *klasifikasi* kepribadian dalam aplikasi ini, membantu pengguna memahami karakter mereka secara akurat dan menyediakan dasar untuk pengembangan aplikasi yang lebih menarik dan akurat di masa depan (Christopher et al., 2022).

2.4.10 Penerapan Data *Mining* Berbasis *Klustering* untuk Penilaian Status Gizi Balita di Puskesmas

1. Jurnal Melissa Indah Fianty dkk

2. Judul

Penerapan Data *Mining* Berbasis *Klustering* untuk Penilaian Status Gizi Balita di Puskesmas

3. Jurnal

Journal of Information Systems and Informatics

4. Volume, Nomor dan Halaman

Vol. 5, No. 4, hlm. 1350-1362

5. Bulan dan Tahun

Desember 2023

6. Penulis

Melissa Indah Fianty, Monika Evelin Johan, Azka Aulia, Mella Margareta Veronica

7. Penerbit

DRPM-UBD

8. Tujuan Penelitian

Menggunakan algoritma *k-means* untuk mengelompokkan data balita berdasarkan status gizi dan mengidentifikasi masalah pertumbuhan yang membutuhkan *intervensi* tepat waktu.

9. Lokasi dan Subjek Penelitian

Puskesmas Kelapa Dua, Tangerang, dengan subjek balita

10. Metode Yang Digunakan

Metode *Knowledge Discovery in Databases (KDD)* dan algoritma *K-means Clustering* untuk *klasifikasi* status gizi balita

11. Perancangan Sistem

Data *cleaning* untuk membersihkan data balita, data *integration* untuk penggabungan data dari berbagai *table*, data *selection* untuk memilih atribut penting, data transformation untuk normalisasi data balita

12. Hasil Penelitian

Ditemukan tiga klaster berdasarkan status gizi balita: *Klaster 0* (kondisi gizi baik), *Klaster 1* (sangat kurus dan pendek), dan *Klaster 2* (pertumbuhan seimbang). Hasil ini memberi wawasan untuk mengatasi masalah gizi pada balita lebih efektif.

13. Kekuatan Penelitian

Menggunakan data yang luas dari berbagai sumber di Puskesmas, pendekatan *klustering* yang *komprehensif* dengan algoritma yang cocok untuk *klasifikasi* data gizi balita

14. Kelemahan Penelitian

Hasil tergantung pada ketersediaan data dari satu Lokasi, perlu data tambahan dari wilayah lain untuk generalisasi hasil

15. Kesimpulan

Penerapan data *minig* dalam *klasifikasi* data balita *efektif* dalam mengidentifikasi kondisi gizi berdasarkan beberapa *indikator antropometri*. Temuan ini membantu Puskesmas dalam intervensi gizi yang tepat bagi balita di wilayah Kelapa Dua (Fianty et al., 2023).

2.4.11 *Machine Learning Algorithms – A Review*

1. **Jurnal Batta Mahesh**

2. **Judul**

Machine Learning Algorithms - A Review

3. **Jurnal**

International Journal of Science and Research (IJSR)

4. **Volume, Nomor dan Halaman**

Vol. 9, No. 1, hlm. 381-386

5. **Bulan dan Tahun**

Januari 2020

6. **Penulis**

Batta Mahesh

7. **Penerbit**

IJSR

8. **Tujuan penelitian**

Membahas peran dan prospek algoritma pembelajaran mesin, termasuk *Decision Tree* dalam *klasifikasi* dan pemrosesan data besar.

9. Lokasi dan Subjek Penelitian

Kajian literatur tentang algoritma data *mining*

10. Metode yang digunakan

Studi literatur, berfokus pada *Decision Tree* sebagai metode *supervised learning* dalam data *mining* untuk *klasifikasi* data berdasarkan aturan dan cabang dalam bentuk pohon yang mudah dipahami.

11. Perancangan Sistem

Decision Tree dirancang untuk membuat cabang-cabang keputusan dari *node* berdasarkan nilai variabel tertentu. Algoritma ini mengelompokkan data ke dalam kategori dengan aturan yang dapat digunakan dalam *klasifikasi* lebih lanjut.

12. Hasil Penelitian

Decision Tree efektif untuk mengidentifikasi pola dalam data dengan tingkat akurasi tinggi dalam situasi *klasifikasi*, tetapi hasilnya bisa terpengaruh oleh variasi data dan struktur pohon.

13. Kekuatan Penelitian

Memberikan gambaran algoritma *Decision Tree* sebagai alat *klasifikasi* data yang *intuitif* dengan keputusan berbasis aturan.

14. Kelemahan Penelitian

Tidak ada eksperimen langsung, dan hasil tergantung pada kondisi data yang digunakan dalam simulasi.

15. Kesimpulan

Decision Tree adalah algoritma yang cocok untuk data *mining* berbasis *klasifikasi*, karena struktur pohon memudahkan dalam interpretasi hasil dan

penggunaan aturan yang mendasar bagi pengambilan keputusan *klasifikasi* (Mahesh, 2020).

2.4.12 Classification Based on Decision Tree Algorithm for Machine Learning

1. Jurnal Bahzad Taha Jijo, Adnan Mohsin Abdulazeez

2. Judul

Classification Based on Decision Tree Algorithm for Machine Learning

3. Jurnal

Journal of Applied Science and Technology Trends

4. Volume, Nomor dan Halaman

Vol. 02, No. 01, hlm. 20 – 28

5. Bulan dan Tahun

Januari 2021

6. Penulis

Bahzad Taha Jijo, Adnan Mohsin Abdulazeez

7. Penerbit

Journal of Applied Science and Technology Trends

8. Tujuan Penelitian

Mengulas pendekatan terbaru penggunaan algoritma *Decision Tree* dalam berbagai aplikasi pembelajaran mesin, termasuk pada analisis penyakit, pengenalan tulisan tangan, dan *klasifikasi* pengguna ponsel cerdas berdasarkan dataset berbeda.

9. Lokasi dan Subjek Penelitian

Tidak ada lokasi spesifik; penelitian literatur pada aplikasi *Decision Tree* di berbagai bidang

10. Metode Yang Digunakan

Studi literatur dengan metode perbandingan untuk algoritma *Decision Tree* (DT), fokus pada *klasifikasi* data dalam *Data Mining*

11. Perancangan Sistem

Menyusun taksonomi algoritma *DT*, mulai dari pemilihan data hingga evaluasi model, dengan pembahasan mengenai *entropi* dan *gain* informasi sebagai indikator pemisahan dalam *klasifikasi*.

12. Hasil Penelitian

Algoritma *DT* menunjukkan efektivitas tinggi dalam berbagai tugas *klasifikasi*, seperti deteksi angka tulisan tangan dan prediksi kondisi medis; juga menghasilkan aturan klasifikasi yang mudah dipahami dan implementasi yang efisien dalam data mining.

13. Kekuatan Penelitian

Memaparkan tinjauan luas tentang *DT*, membahas variasi implementasi pada berbagai dataset; akurasi *DT* umumnya tinggi dalam tugas klasifikasi khusus.

14. Kelemahan Penelitian

Tidak menguji langsung efektivitas *DT* pada kumpulan data baru; hasil berbasis kajian literatur dengan berbagai ketergantungan pada tipe data yang digunakan

15. Kesimpulan

Algoritma *Decision Tree* merupakan metode andal untuk klasifikasi dalam data mining, dengan kemampuan menghasilkan aturan keputusan yang jelas dan aplikatif. *DT* dapat digunakan untuk prediksi dan *klasifikasi* pada data yang bervariasi, efektif pada berbagai konteks (Taha Jijo dan Abdulazeez Mohsin, 2021).

2.4.13 *Predictive Analytics in Business Analytics: Decision Tree*

1. Jurnal Chee Sun Lee dkk

2. Judul

Predictive Analytics in Business Analytics: Decision Tree

3. Jurnal

Advances in Decision Sciences

4. Volume, Nomor dan Halaman

Vol. 26, No. 1, Maret 2022, halaman tidak tercantum

5. Bulan dan Tahun

Maret 2022

6. Penulis

Chee Sun Lee, Peck Yeng Sharon Cheang, Massoud Moslehpour

7. Penerbit

Asia University, Taiwan

8. Tujuan penelitian

Penelitian ini bertujuan mengeksplorasi penggunaan *Decision Tree* dalam analisis prediktif untuk membantu bisnis dalam membuat keputusan yang lebih baik berbasis data. Penelitian ini juga bertujuan memberikan panduan bagi para peneliti dan analis di bidang analitik bisnis dan prediktif.

9. Lokasi dan Subjek Penelitian

Lokasi tidak *spesifik*; subjek meliputi berbagai industri yang menggunakan analisis bisnis berbasis data.

10. Metode yang digunakan

Peninjauan sistematis literatur menggunakan basis data seperti *Web of Science* dan *Scopus*, dengan fokus utama pada *analitik prediktif* dan *Decision Tree* untuk *klasifikasi* data.

11. Perancangan Sistem

Sistem *analitik* yang dirancang memanfaatkan *Decision Tree* sebagai alat *klasifikasi* untuk analisis *prediktif*. Struktur *Decision Tree* melibatkan *node* dan cabang yang memisahkan data ke dalam kategori-kategori yang ditentukan melalui algoritma sederhana, dengan kelebihan mudah dipahami tanpa pengetahuan *statistik* yang kompleks.

12. Hasil Penelitian

Decision Tree merupakan alat *klasifikasi* yang efektif dan dapat diterapkan dalam berbagai sektor seperti *manufaktur*, kesehatan, dan analisis risiko. Hasil menunjukkan bahwa *Decision Tree* lebih *user-friendly* dan mudah diinterpretasikan dibanding algoritma lain, seperti *Naive Bayes* dan *Support Vector Machine*.

13. Kekuatan Penelitian

Penggunaan *Decision Tree* dapat menyederhanakan proses *klasifikasi* data dalam analisis *prediktif* tanpa memerlukan formula statistik yang rumit. Alat ini sangat *fleksibel* dan cocok untuk berbagai jenis data input, serta mampu menghasilkan model yang mudah dipahami dalam pengambilan keputusan.

14. Kelemahan Penelitian

Meskipun *Decision Tree* sangat bermanfaat, algoritma ini rentan terhadap *overfitting* pada dataset tertentu jika tidak dilakukan proses pemangkasan

(*pruning*) yang memadai. Perlu adanya strategi untuk mencegah kompleksitas berlebih yang dapat menurunkan akurasi model pada data baru.

15. Kesimpulan

Decision Tree merupakan metode *prediktif* yang efektif dalam analitik bisnis, terutama untuk klasifikasi data besar. Keunggulannya terletak pada kemudahan penggunaan dan interpretasi, menjadikannya salah satu alat paling populer dalam analitik *prediktif* di berbagai sektor bisnis (Lee et al., 2022).

2.4.14 Implementation Educational Data Mining for Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5

1. Jurnal Lili Dwi Yulianto dkk

2. Judul

Implementation Educational Data Mining for Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5

3. Jurnal

Jurnal Mantik

4. Volume, Nomor dan Halaman

Vol. 4, No. 1, Mei 2020, hal. 441-451

5. Bulan dan Tahun

Mei 2020

6. Penulis

Lili Dwi Yulianto, Agung Triayudi, Ira Diana Sholihati

7. Penerbit

Universitas Nasional

8. Tujuan penelitian

Penelitian ini bertujuan mengembangkan sistem prediksi untuk mengevaluasi kinerja akademik siswa menggunakan metode data *mining* dengan membandingkan algoritma *K-Nearest Neighbor (KNN)* dan *Decision Tree C4.5*, sehingga dapat memberikan panduan untuk meningkatkan kualitas pendidikan dan membantu mengidentifikasi potensi keterlambatan studi pada mahasiswa.

9. Lokasi dan Subjek Penelitian

Lokasi penelitian tidak spesifik; subjek penelitian adalah data kinerja mahasiswa di perguruan tinggi, yang dianalisis untuk mengidentifikasi pola dalam keterlambatan atau keberhasilan akademik

10. Metode Yang Digunakan

Dua metode *klasifikasi* diterapkan, yaitu *Decision Tree C4.5* dan *K-Nearest Neighbor (KNN)*. Algoritma *Decision Tree* digunakan untuk memodelkan struktur hirarki *klasifikasi* data dari akar ke daun, sedangkan *KNN* bekerja dengan mengukur jarak antara data uji dan data yang sudah dikelompokkan. Kedua metode ini diuji menggunakan *Rapid Miner Studio*, dengan analisis tingkat akurasi sebagai parameter utama.

11. Perancangan Sistem

Sistem dirancang untuk melakukan *klasifikasi* data mahasiswa berdasarkan atribut kinerja akademik, seperti nilai IPK, waktu belajar, dan status beasiswa. Proses dimulai dari analisis kebutuhan sistem, pengumpulan data melalui *survei*, normalisasi data untuk mengurangi noise, hingga pembuatan *decision*

tree sebagai model prediksi. Model ini kemudian diimplementasikan menggunakan bahasa pemrograman *PHP* dan *data mining tools* untuk menghasilkan aturan *klasifikasi* yang dapat memprediksi keberhasilan atau keterlambatan kelulusan.

12. Hasil Penelitian

Berdasarkan perbandingan kedua metode, hasil penelitian menunjukkan bahwa *K-Nearest Neighbor* memiliki tingkat akurasi lebih tinggi (59.32%) dibandingkan dengan *Decision Tree C4.5* yang memiliki akurasi 54.80%. *KNN* mampu memberikan hasil prediksi yang lebih optimal dalam *klasifikasi* data kinerja siswa untuk mendukung proses pengambilan keputusan dalam bidang pendidikan.

13. Kekuatan Penelitian

Penggunaan *KNN* dan *Decision Tree* dalam *Educational Data Mining* efektif untuk mengevaluasi kinerja mahasiswa dan membantu menentukan faktor utama yang berpengaruh terhadap kelulusan tepat waktu. Model klasifikasi yang dihasilkan sangat berguna dalam mengidentifikasi siswa yang berisiko keterlambatan, dan sistem ini dapat diintegrasikan ke dalam aplikasi pendidikan untuk pemantauan kinerja akademik yang lebih proaktif.

14. Kelemahan Penelitian

Penelitian ini terbatas pada penggunaan dua metode *klasifikasi*, yakni *Decision Tree* dan *KNN*, tanpa memperhitungkan metode lain yang mungkin memberikan hasil akurasi lebih tinggi. Selain itu, metode *Decision Tree* cenderung rentan terhadap *overfitting* pada data tertentu jika tidak dilakukan pemangkasan cabang yang memadai, yang dapat menurunkan akurasi saat diterapkan pada data baru.

15. Kesimpulan

KNN terbukti menghasilkan akurasi lebih tinggi dibandingkan *Decision Tree* C4.5 dalam prediksi kinerja akademik siswa. *Educational Data Mining* dengan kedua algoritma ini memberikan kontribusi yang signifikan dalam mendukung proses evaluasi akademik di perguruan tinggi, dan diharapkan dapat dikembangkan lebih lanjut untuk meningkatkan mutu pendidikan dengan memantau dan mengidentifikasi potensi keterlambatan akademik secara lebih dini (Yulianto et al., 2020).

2.4.15 *Educational Data Mining Using Cluster Analysis and Decision Tree*

Technique: A Case Study

1. **Snježana Križanić**

2. **Judul**

*Educational Data Mining Using Cluster Analysis and Decision Tree
Technique: A Case Study*

3. **Jurnal**

International Journal of Engineering Business Management

4. **Volume, Nomor dan Halaman**

Volume 12, 2020, halaman 1-9

5. **Bulan dan Tahun**

Bulan tidak disebutkan, Tahun 2020

6. **Penulis**

Snježana Križanić

7. **Penerbit**

SAGE Publications

8. Tujuan penelitian

Penelitian ini bertujuan menganalisis perilaku siswa dalam lingkungan *e-learning* menggunakan teknik *Educational Data Mining* (EDM) dengan pendekatan *cluster analysis* dan *decision tree*, untuk mengeksplorasi variabel yang memengaruhi keberhasilan ujian dalam suatu kursus di institusi pendidikan tinggi di Kroasia.

9. Lokasi dan Subjek Penelitian

Lokasi di Kroasia, subjek penelitian adalah mahasiswa yang terdaftar pada kursus *e-learning* di institusi pendidikan tinggi selama semester tertentu.

10. Metode yang digunakan

Penelitian ini menggunakan metode *cluster analysis* untuk mengelompokkan siswa berdasarkan perilaku akses mereka ke konten *e-learning* dan teknik *decision tree* untuk membuat representasi pohon keputusan yang menunjukkan hubungan antara variabel akses materi dan hasil ujian, dengan analisis frekuensi akses pada berbagai jenis materi kursus.

11. Perancangan Sistem

Sistem pengolahan data dirancang dengan mengunduh log aktivitas siswa dari *e-learning*, kemudian data diproses melalui data *cleaning*, diikuti dengan partisi data untuk *cluster analysis*, dan kemudian dianalisis lebih lanjut dengan *decision tree*. Alur ini bertujuan memvalidasi model perilaku berdasarkan frekuensi akses terhadap berbagai materi kursus.

12. Hasil Penelitian

Berdasarkan *clustering*, siswa terbagi menjadi tiga kelompok dengan pola akses berbeda. Hasil menunjukkan bahwa siswa dengan frekuensi akses tinggi pada materi utama cenderung mendapatkan nilai ujian yang lebih baik. Teknik

decision tree kemudian mengonfirmasi variabel frekuensi akses sebagai penentu utama keberhasilan dalam ujian.

13. Kekuatan Penelitian

Pendekatan kombinasi *cluster analysis* dan *decision tree* memungkinkan analisis perilaku siswa yang lebih mendalam dalam *e-learning*, yang membantu mengidentifikasi materi kursus penting untuk kesuksesan akademik. Model ini menyediakan informasi bagi institusi pendidikan tentang pola akses yang memengaruhi hasil belajar.

14. Kelemahan Penelitian

Penelitian ini terbatas pada satu kursus dan satu generasi siswa di institusi yang sama, sehingga hasilnya mungkin tidak dapat digeneralisasi untuk kursus atau konteks lain. Selain itu, teknik *decision tree* bisa rentan terhadap perubahan data sehingga diperlukan pemeliharaan model secara berkala untuk validitas hasil yang berkelanjutan.

15. Kesimpulan

Cluster analysis dan *decision tree* terbukti efektif dalam mengidentifikasi pola perilaku siswa yang relevan dengan kesuksesan akademik. Penggunaan data dari sistem *e-learning* dapat mendukung perbaikan konten kursus dan metode pengajaran, serta memungkinkan institusi memahami dan mengoptimalkan pengalaman belajar siswa (Križanić, 2020).

Berikut beberapa kesimpulan yang penulis berikan dalam tinjauan 15 jurnal yang sudah dibahas dalam bab 2, sebagai berikut:

1. **Efektivitas dalam Klasifikasi dan Prediksi:** Algoritma *Decision Tree* secara konsisten menunjukkan keunggulan dalam *klasifikasi* data dan prediksi di berbagai

bidang, mulai dari kesehatan, pendidikan, hingga bisnis. Struktur pohon keputusannya memungkinkan algoritma ini untuk membagi data dengan jelas berdasarkan aturan yang *intuitif*, menjadikannya metode yang populer untuk analisis berbasis kategori.

2. **Aplikasi Luas di Berbagai Bidang:** *Decision Tree* digunakan secara luas dalam data *mining* di banyak sektor seperti prediksi kepuasan pelanggan, kelayakan donor darah, penentuan paket penjualan, dan analisis perilaku siswa. Keandalannya dalam *mengidentifikasi* pola dari data besar membuat algoritma ini sangat *aplikatif* dalam banyak konteks, termasuk pendidikan, bisnis, kesehatan, dan analitik sosial.
3. **Kemudahan Interpretasi dan Implementasi:** Keunggulan utama *Decision Tree* adalah kemampuannya memberikan hasil yang mudah dipahami. Struktur pohon dan cabangnya membuat proses pengambilan keputusan lebih transparan, yang memudahkan pengguna dalam memahami hasil tanpa perlu pengetahuan statistik yang mendalam. Hal ini sangat berguna dalam konteks yang memerlukan interaksi pengguna yang luas, seperti aplikasi *web* atau *mobile* untuk *klasifikasi* kepribadian atau kepuasan pelanggan.
4. **Kekuatan dalam Mengelola Data Besar dan Bervariasi:** *Decision Tree* mampu menangani dataset besar dan beragam dengan efisien. Dalam beberapa penelitian, algoritma ini digunakan untuk *mengidentifikasi* faktor-faktor utama yang memengaruhi hasil tertentu, seperti variabel kesehatan untuk kelayakan donor darah atau frekuensi akses terhadap materi pembelajaran dalam *e-learning*. Kemampuan ini memungkinkan *Decision Tree* untuk menghasilkan model yang dapat diandalkan dan adaptif dalam skenario data yang dinamis.
5. **Tantangan *Overfitting* dan Evaluasi yang Terbatas:** Walaupun algoritma ini efektif, beberapa penelitian menunjukkan bahwa *Decision Tree* rentan terhadap

overfitting, terutama ketika bekerja dengan dataset tertentu yang tidak divariasikan atau memiliki cabang pohon yang terlalu kompleks. Oleh karena itu, pemangkasan (*pruning*) dan evaluasi algoritma pada dataset yang lebih beragam sangat disarankan untuk menjaga akurasi model pada data baru.

6. **Penggunaan Beragam Metode Pendukung untuk Validasi:** Sebagian besar penelitian menggunakan metode tambahan, seperti *k-fold cross-validation* atau membandingkan *Decision Tree* dengan algoritma lain seperti *Naïve Bayes* dan *K-Nearest Neighbor (KNN)*, untuk menilai performa. Hasilnya menunjukkan bahwa *Decision Tree* sering kali memiliki tingkat akurasi yang lebih tinggi dalam *klasifikasi*, namun metode tambahan seperti *KNN* kadang menghasilkan akurasi lebih optimal untuk data yang memerlukan analisis berbasis jarak.

BAB III

METODOLOGI PENELITIAN

3.1 Tinjauan Umum Perusahaan

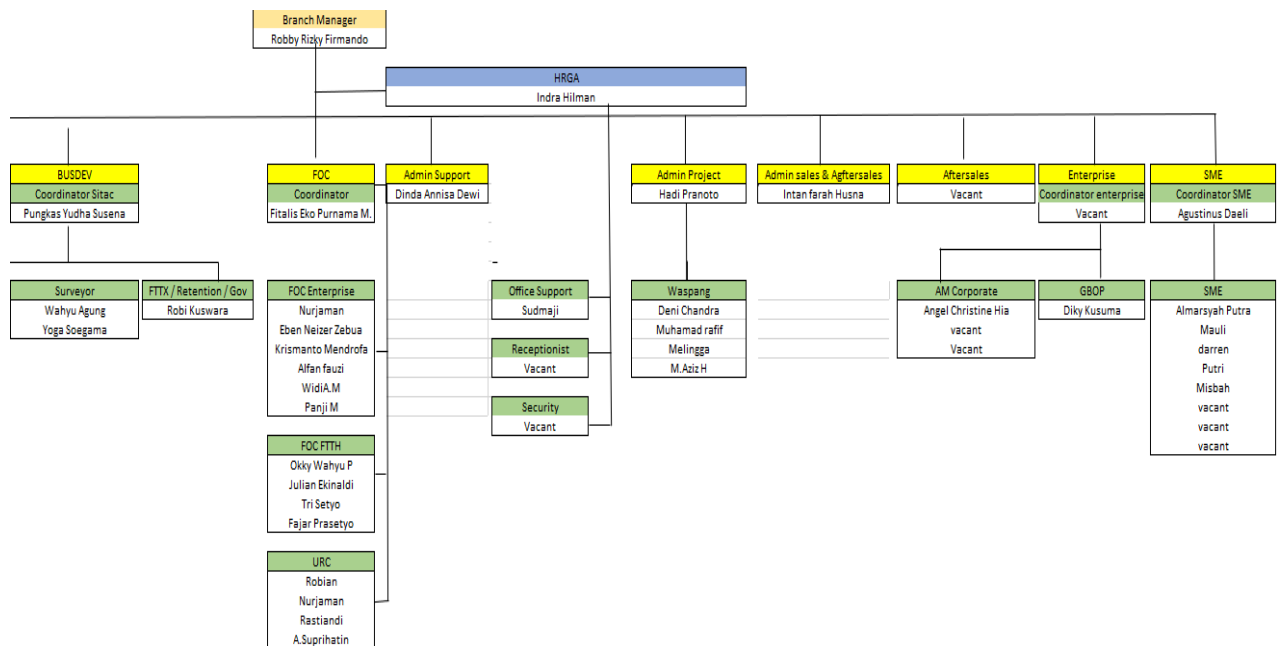
3.1.1 Profil dan Sejarah Perusahaan

PT Jala Lintas Media merupakan perusahaan yang berdiri sejak tahun 2009, Jala Lintas Media menghadirkan layanan Teknologi Informasi dan Komunikasi (ICT) untuk berbagai segmen bisnis dan industri. Jala Lintas media merupakan perusahaan penyedia layanan Internet dan Akses Jaringan untuk segmen individu maupun bisnis, menawarkan layanan terlengkap melalui solusi ICT *one-stop* dan *end-to-end* yang menjangkau seluruh wilayah di Indonesia.

Jala lintas media juga mempunyai beberapa produk yang ditawarkan diantaranya untuk FTTH (*Fiber To The Home*) dia disebut *Benefit*, untuk perusahaan dia mempunya produk yang bernama *Netvolve*, dan *Mahavira* System Integra yang merupakan produk yang fokus pada pengembangan infrastruktur ICT dan *Managed Service*. Untuk kantor pusat Jala Lintas Media sendiri berlokasi di Cibinong, Jawa Barat yang dimana lokasi tersebut adalah awal berdirinya perusahaan dan menjadi pusat bisnis dari pertama Didirikan hingga saat ini.

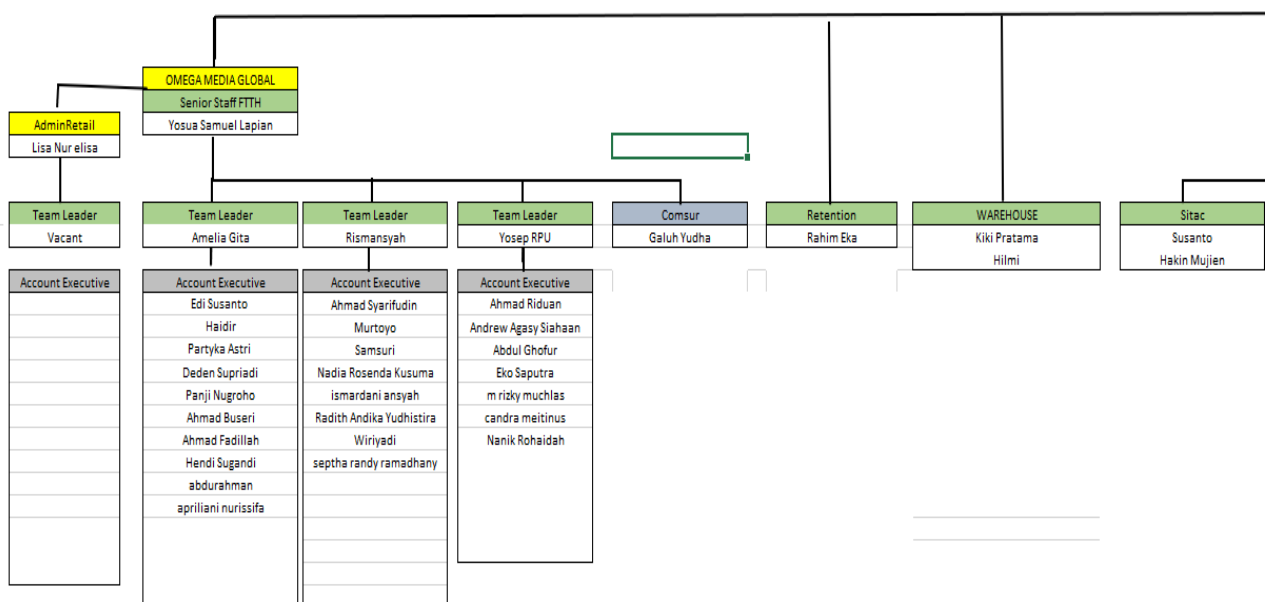
Pada kesempatan ini penulis berkesempatan Kerja Praktek di salah satu Cabang Jala lintas media yang sudah beroperasi sejak tahun 2020.

3.1.2 Struktur organisasi, tugas dan wewenang setiap organisasi perusahaan



Gambar 3. 1 Struktur organisasi, tugas dan wewenang

Sumber (HRD PT Jala lintas media)



Gambar 3. 2 Struktur organisasi, tugas dan wewenang

Sumber (HRD PT Jala lintas media)

3.2 Identifikasi Kebutuhan

Untuk menjamin kepuasan pelanggan, identifikasi tuntutan perusahaan difokuskan pada hal-hal yang berkaitan dengan pemantauan dan peningkatan kualitas layanan. Pemeriksaan terhadap kondisi bisnis mengungkapkan kebutuhan penting berikut:

1. Jumlah Pelanggan yang Tidak Puas

PT Jala Lintas Media Branch Banten menyediakan berbagai pelanggan, dan seiring bertambahnya Pelanggan, semakin sulit untuk mempertahankan tingkat kepuasan yang tinggi. Perusahaan terkadang menerima keluhan tentang gangguan layanan, waktu respons teknis, dan kualitas layanan internet. Hal ini menunjukkan bahwa diperlukan sistem yang dapat mengevaluasi data konsumen secara lebih menyeluruh untuk mengantisipasi dan mengatasi ketidakpuasan sebelum menjadi masalah yang lebih signifikan.

2. Tidak Ada Sistem Prediksi Tingkat Kepuasan Pelanggan yang Akurat

Saat ini, bisnis mungkin menggunakan teknik manual atau terbatas, termasuk survei atau wawancara, untuk melacak kepuasan pelanggan. Namun, karena pendekatan ini bersifat *reaktif* daripada *proaktif*, pendekatan ini memiliki kekurangan. Oleh karena itu, agar manajemen dapat mengambil tindakan pencegahan lebih cepat, organisasi memerlukan alat prediksi berbasis data *mining* yang dapat membantu memperkirakan penurunan tingkat kepuasan pelanggan sebelum masalah besar berkembang.

3. Meningkatkan Pengambilan Keputusan Berdasarkan Data

Untuk meningkatkan layanan dan daya saing, PT Jala Lintas Media harus membuat penilaian yang lebih cerdas dan berbasis data. Bisnis dapat *mengekstrak* informasi yang berguna dari data klien mereka saat ini, termasuk riwayat keluhan, tren penggunaan layanan, dan umpan balik pelanggan, dengan menerapkan algoritma

Pohon Keputusan atau *Decision Tree*. Ini akan membantu bisnis dalam menciptakan inisiatif retensi klien dan kualitas layanan yang lebih baik.

4. Segmentasi Pelanggan Diperlukan untuk Layanan yang Lebih Disesuaikan

Kebutuhan setiap kategori pelanggan dapat bervariasi. Bisnis dapat menggunakan penambahan data untuk membagi klien mereka sesuai dengan selera dan perilaku mereka. Pada akhirnya, segmentasi ini dapat meningkatkan kepuasan pelanggan dengan memungkinkan PT Jala Lintas Media menawarkan layanan yang lebih sesuai dengan persyaratan setiap kelompok klien.

5. Meningkatkan Penanganan Keluhan dan Umpan Balik Pelanggan

Perusahaan akan dapat menindaklanjuti masalah pelanggan dengan lebih cepat dan tepat jika mereka memiliki sistem yang memasukkan umpan balik dan keluhan konsumen ke dalam analisis. Agar setiap masukan klien dapat diubah menjadi data yang dapat diukur dan ditanggapi secara efisien, diperlukan sistem manajemen data yang kuat.

3.3 Metode Yang Digunakan

Pada Metode penelitian yang berjudul **“Penerapan Teknik Data Mining pada Database Pelanggan Menggunakan Algoritma Decision Tree untuk Memprediksi Tingkat Kepuasan di PT Jala Lintas Media *Branch* Banten”**, metode yang digunakan dapat mengungkap beberapa tahapan penting:

3.3.1 Metode Pengumpulan Data

Data yang dibutuhkan untuk penelitian ini dikumpulkan melalui beberapa sumber, antara lain:

1. Data Sekunder: Menggunakan data historis dari database pelanggan PT Jala Lintas Media, seperti data profil pelanggan, riwayat transaksi, dan umpan balik. Data ini diperoleh dari sistem internal perusahaan.

3.3.2 Metode Data Mining

Teknik yang digunakan dalam penelitian ini adalah algoritma *Decision Tree* yang merupakan salah satu metode data mining. *Decision Tree* dipilih karena kemampuannya dalam mengolah data dalam jumlah besar, mengidentifikasi pola, dan menghasilkan model yang mudah dipahami. Langkah-langkah penerapan *Decision Tree* antara lain:

1. *Data Preprocessing*: Pembersihan data untuk menghilangkan data duplikat, data kosong, atau data yang tidak relevan.
2. Pembagian Data: Data akan dibagi menjadi data latih dan data uji, dengan tujuan untuk membangun model prediksi yang andal.
3. Pembentukan Model: Algoritma *Decision Tree* akan digunakan untuk membuat model *prediktif* berdasarkan data latih.
4. Evaluasi Model: Setelah model terbentuk, model diuji menggunakan data uji untuk mengetahui seberapa baik model tersebut memprediksi tingkat kepuasan pelanggan.

3.3.3 Metode Analisis Data

Setelah model *Decision Tree* terbentuk, data prediksi dianalisis untuk mengetahui faktor-faktor apa saja yang paling berpengaruh terhadap kepuasan pelanggan. Hasil analisis ini akan menjadi dasar untuk memberikan saran atau rekomendasi kepada PT Jala Lintas Media dalam meningkatkan kualitas layanannya.

3.3.4 Evaluasi Model Prediksi

Untuk mengetahui *spesifikasi* model, dilakukan pengukuran dengan menggunakan beberapa metrik evaluasi seperti akurasi, presisi, dan *recall*. Metode evaluasi ini akan memberikan gambaran seberapa efektif model tersebut dalam memprediksi tingkat kepuasan pelanggan.

Berikut adalah rumus untuk *Recall*, Presisi, dan Akurasi dalam evaluasi model:

1. Recall

Recall mengukur seberapa baik model dapat menemukan seluruh data positif.
$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}}$$

16. TP: Prediksi positif yang benar.

17. FN: Data positif yang diprediksi salah sebagai negat.

2. Presisi (*Precision*)

Presisi mengukur seberapa akurat prediksi positif model.

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Positive (FP)}}$$

- FP: Data negative yang di prediksi salah sebagai positif.

3. Akurasi (*Accuracy*)

Akurasi mengukur seberapa banyak prediksi model yang benar dibandingkan seluruh data.

$$\text{Accuracy} = \frac{\text{True Positive (TP)} + \text{True Negative (TN)}}{\text{Total Data}}$$

- TN: Prediksi negative yang benar.

Dengan pendekatan *metodologis* ini, penelitian ini diharapkan mampu memberikan wawasan yang kuat mengenai faktor-faktor yang memengaruhi

kepuasan pelanggan di PT Jala Lintas Media dan membantu perusahaan membuat keputusan yang lebih berbasis data.

3.4 Algoritma Yang digunakan

Pada penelitian ini, algoritma yang digunakan untuk memprediksi tingkat kepuasan pelanggan adalah **Decision Tree C4.5**. *Decision Tree* adalah salah satu algoritma dari teknik data *mining* yang digunakan untuk *klasifikasi* dan prediksi. Algoritma ini sangat populer karena memiliki struktur yang menyerupai pohon keputusan, di mana setiap cabang mewakili keputusan berdasarkan atribut tertentu, dan setiap daun mewakili hasil atau *klasifikasi* akhir.

3.4.1 Pohon Keputusan Algoritma Proses

1. Pemilihan Atribut Utama

Algoritma *Decision Tree* akan memilih atribut yang paling *informatif* sebagai akar pohon. Pemilihan atribut utama ini biasanya berdasarkan nilai *Information Gain* atau *Gini Index*, yang mengukur seberapa baik atribut tersebut dalam memisahkan data menjadi beberapa kelompok.

2. Pembuatan Cabang

Setelah atribut utama dipilih, data akan dibagi ke dalam cabang-cabang berdasarkan nilai-nilai dari atribut tersebut. Proses ini diulangi untuk setiap subset data yang dihasilkan sampai semua data telah diklasifikasikan atau tidak ada lagi atribut yang relevan untuk diproses.

3. Pembentukan Daun

Daun di *Decision Tree* adalah hasil dari *klasifikasi* akhir, dalam hal ini adalah tingkat prediksi kepuasan pelanggan, yang bisa berupa kategori seperti "Puas" atau "Tidak Puas". Setiap daun mewakili hasil rangkaian keputusan berdasarkan atribut yang dipilih.

4. Pemangkasan Pohon

Setelah pohon terbentuk, dilakukan evaluasi terhadap kompleksitas pohon. Jika pohon terlalu kompleks (*overfitting*).

3.4.2 Keunggulan *Decision Tree*

1. Mudah dipahami : Model yang dihasilkan berupa keputusan pohon yang sederhana dan mudah diinterpretasikan oleh manusia.
2. Tidak memerlukan normalisasi data : Algoritma ini dapat bekerja dengan baik meskipun data tidak dinormalisasi.
3. menyertakan data kategori dan *numerik* : *Decision Tree* mampu menangani atribut kategori maupun *numerik* dengan baik.

3.4.3 Implementasi

Dalam implementasi penelitian ini, algoritma *Decision Tree* akan digunakan untuk *menganalisis* data pelanggan dari PT Jala Lintas Media dan membangun model *prediktif* berdasarkan atribut-atribut seperti usia, jenis kelamin, jenis layanan yang digunakan, dan sebagainya. Setelah model terbentuk, model tersebut dapat digunakan untuk memprediksi tingkat kepuasan pelanggan baru atau mengidentifikasi pelanggan yang berpotensi tidak puas.

Algoritma ini akan diimplementasikan menggunakan *Python* dengan pustaka data mining seperti *scikit-learn*.

Langkah-langkah implementasinya meliputi:

1. Mengimpor *dataset* pelanggan.
2. Membersihkan dan memproses data.
3. Membangun model Pohon Keputusan.
4. Menguji dan memutar performa model menggunakan data uji.

Dengan menggunakan algoritma *Decision Tree*, penelitian ini diharapkan dapat memberikan prediksi yang akurat terkait kepuasan pelanggan dan membantu perusahaan dalam meningkatkan kualitas layanan.

3.5 Kerangka Pemikiran

Kerangka pemikiran yang di ringkas mulai dari rumusan masalah yang berfungsi sebagai masukan pemikiran, dan bertujuan untuk hasil pemikiran yang juga berfungsi sebagai solusi. Agar kerangka berpikir mudah dipahami, maka kerangka tersebut harus menggunakan tabel yang jelas dan sederhana.

Tabel 3. 1 Kerangka Pemikiran

Bab	Pokok Pembahasan	Kerangka pemikiran
Pendahuluan	Latar belakang Masalah	Menyoroti pentingnya memahami kepuasan pelanggan di tengah persaingan bisnis layanan internet yang semakin ketat. PT Jala Lintas Media Branch Banten perlu

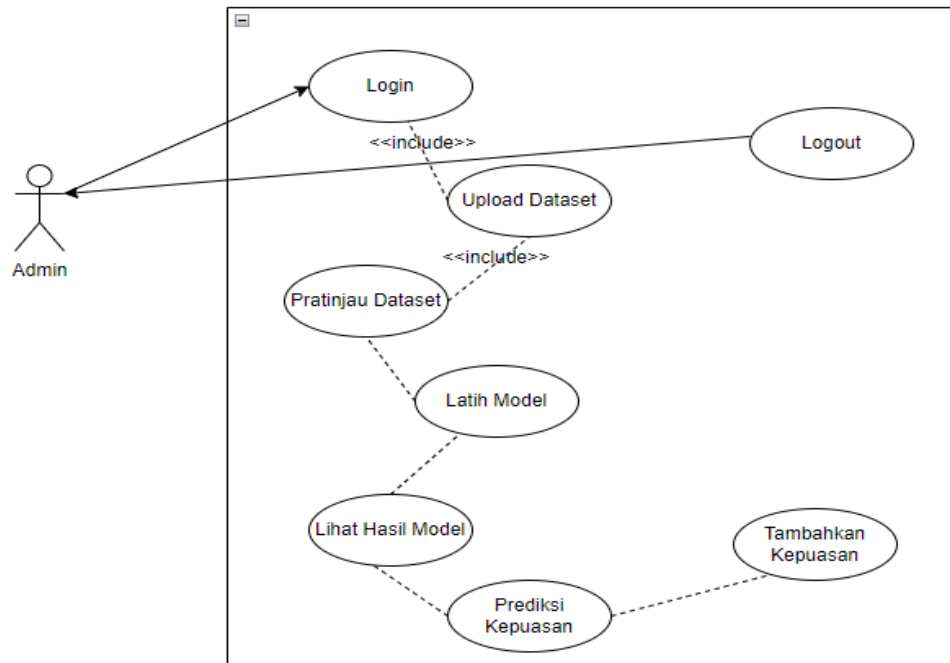
		menerapkan teknologi data <i>mining</i> untuk memprediksi tingkat kepuasan guna meningkatkan layanan.
Pendahuluan	Rumusan Masalah	Bagaimana algoritma <i>Decision Tree</i> dapat digunakan untuk memprediksi kepuasan pelanggan? Apa faktor-faktor utama yang memengaruhi kepuasan pelanggan di perusahaan ini?
Pendahuluan	Tujuan dan Manfaat	Menyediakan metode untuk memprediksi kepuasan pelanggan dengan menggunakan <i>Decision Tree</i> , memberikan wawasan mengenai faktor-faktor utama yang memengaruhi kepuasan, serta meningkatkan kualitas layanan perusahaan.
Landasan Teori	Data Mining	Memahami teknik data <i>mining</i> sebagai alat yang digunakan untuk mengidentifikasi pola dari

		data pelanggan yang dapat memengaruhi kepuasan dan loyalitas mereka.
Landasan Teori	Algoritma <i>Decision Tree</i>	Mempelajari cara kerja <i>Decision Tree</i> dan keunggulannya dalam mengklasifikasikan data kepuasan pelanggan berdasarkan variabel-variabel tertentu, seperti kecepatan layanan dan kualitas dukungan teknis.
Metodologi Penelitian	Desain Sistem	Merancang sistem berbasis web yang mengintegrasikan model <i>Decision Tree</i> untuk mengolah data pelanggan dan menghasilkan prediksi kepuasan.
Metodologi Penelitian	Implementasi Sistem	Mengimplementasikan desain sistem menjadi aplikasi web berbasis <i>Python</i> untuk memudahkan akses dan analisis kepuasan pelanggan secara <i>real-time</i> oleh pihak perusahaan.

Hasil dan Pembahasan	Hasil Implementasi	Memaparkan hasil dari prediksi kepuasan pelanggan yang diperoleh dari model <i>Decision Tree</i> , serta menguraikan faktor-faktor yang memiliki pengaruh signifikan terhadap kepuasan.
Hasil dan Pembahasan	Analisis Kinerja	Menganalisis kinerja model berdasarkan akurasi prediksi dan efektivitasnya dalam memberikan wawasan yang dapat membantu peningkatan kualitas layanan di PT Jala Lintas Media.
Kesimpulan	Ringkasan Hasil	Merangkum hasil penelitian yang menyoroti efektivitas penggunaan <i>Decision Tree</i> untuk memprediksi kepuasan pelanggan dan menyarankan perbaikan layanan yang dapat dilakukan berdasarkan hasil analisis yang diperoleh dari model.

3.6 Perancangan UML

3.6.1 Use Case Diagram



Gambar 3. 3 Use-case Diagram

3.6.2 Use-case Scenario

Tabel 3. 2 Use-case Scenario login

Identifikasi	
Deskripsi	Admin masuk dalam mengakses fitur utama misalnya, mengunggah <i>dataset</i> , membuat prediksi, dan melatih model
Aktor	Admin
Prekondisi	Admin telah memiliki akun dengan <i>username</i> dan juga <i>password</i> yang valid
Alur normal	1. Admin membuka halaman <i>login</i> 2. Admin memasukan <i>username</i> dan <i>password</i> 3. Sistem memvalidasi <i>kredensial</i>

	4. Jika valid sistem memberikan akses ke <i>dashboard</i>
Alur alternatif	Jika <i>password</i> atau <i>username</i> salah, sistem memberikan <i>notifikasi error</i> dan meminta admin mencoba lagi
Poskondisi	Admin berhasil masuk ke dalam sistem

Tabel 3. 3 *Use-case Scenario upload dataset*

Identifikasi	
Deskripsi	Admin mengunggah file <i>dataset</i> pelanggan untuk digunakan dalam proses pelatihan model atau prediksi
Aktor	Admin
Prekondisi	1. Admin sudah <i>login</i> ke sistem 2. Format file <i>dataset</i> sesuai dengan spesifikasi sistem (misalnya CSV dengan <i>header</i> yang sesuai
Alur normal	1. Admin memilih menu unggah <i>dataset</i> 2. Admin memilih <i>dataset</i> dari perangkatnya 3. Sistem memvalidasi format file 4. Jika valid sistem menyimpan <i>dataset</i> ke <i>database</i> 5. Sistem menampilkan pratinjau <i>dataset</i> kepada admin
Alur alternatif	Jika format file tidak <i>valid</i> sistem memberikan <i>notifikasi error</i> dan meminta admin mengunggah kembali file
Poskondisi	<i>Dataset</i> berhasil disimpan dan di unggah dalam <i>database</i>

Tabel 3. 4 *Use-case Scenario* tambahkan kolom kepuasan

Identifikasi	
Deskripsi	Sistem menambahkan kolom "Kepuasan" berdasarkan aturan yang telah di tentukan (lama_berlangganan > 12 bulan dianggap "puas", jika < 10 dianggap "tidak puas")
Aktor	Admin
Prekondisi	<i>Dataset</i> sudah di unggah kedalam sistem
Alur normal	1. Admin memilih menu untuk menambahkan kolom kepuasan 2. Sistem menghitung kepuasan berdasarkan aturan (misalnya: IF lama_berlangganan > 12 bulan THEN "Puas") 3. Sistem menyimpan hasil kedalam <i>database</i> 4. Sistem menampilkan <i>dataset</i> yang sudah di perbaharui kedalam kolom kepuasan
Alur alternatif	Jika <i>dataset</i> belum di unggah sistem memberikan <i>notifikasi error</i>
Poskondisi	Kolom "kepuasan" berhasil di tambahkan kedalam <i>database</i>

Tabel 3. 5 *Use-case Scenario* latih model

Identifikasi	
Deskripsi	Admin melatih model <i>Desicion tree</i> berdasarkan <i>dataset</i> yang tersedia
Aktor	Admin
Prekondisi	<i>Dataset</i> dengan kolom "Kepuasan" sudah ada di sistem
Alur normal	1. Admin memilih menu latih model

	2. Sistem membagi <i>dataset</i> menjadi data pelatihan dan data pengujian 3. Sistem melatih model <i>Decision tree</i> berdasarkan data pelatihan 4. Sistem mengevaluasi model menggunakan data pengujian (<i>akurasi, precision, recall</i>) 5. Sistem menyimpan model yang di latih 6. Sistem menampilkan hasil evaluasi kepada admin
Alur alternatif	Jika <i>dataset</i> belum tersedia sistem memberikan <i>notifikasi error</i>
Poskondisi	Model <i>Decision tree</i> berhasil di latih dan tersimpan dalam sistem

Tabel 3. 6 *Use-case Scenario* prediksi kepuasan

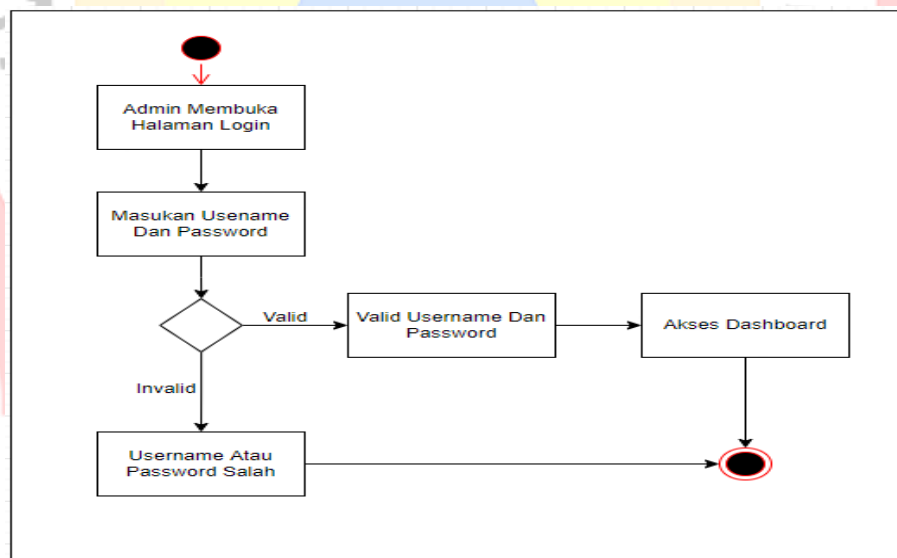
Identifikasi	
Deskripsi	Admin menggunakan model yang telah di latih untuk memprediksi tingkat kepuasan pelanggan baru berdasarkan <i>input</i> data
Aktor	Admin
Prekondisi	Model <i>Decision Tree</i> sudah di latih
Alur normal	1. Admin memilih menu prediksi 2. Admin memasukan data pelanggan baru (misalnya : jumlah transaksi, pendapatan bulanan) 3. Sistem menggunakan model untuk memprediksi kepuasan (Puas/Tidak puas) 4. Sistem menampilkan hasil prediksi kepada admin
Poskondisi	Hasil prediksi berhasil ditampilkan

Tabel 3. 7 *Use-case Scenario logout*

Identifikasi	
Deskripsi	Admin keluar dari sistem untuk mengakhiri sesi
Aktor	Admin
Prekondisi	Admin sudah <i>login</i> ke sistem
Alur normal	1. Admin memilih menu <i>logout</i> 2. Sistem menghapus sesi <i>login</i> admin 3. Admin di arahkan kembali ke halaman <i>login</i>
Poskondisi	Admin berhasil <i>logout</i>

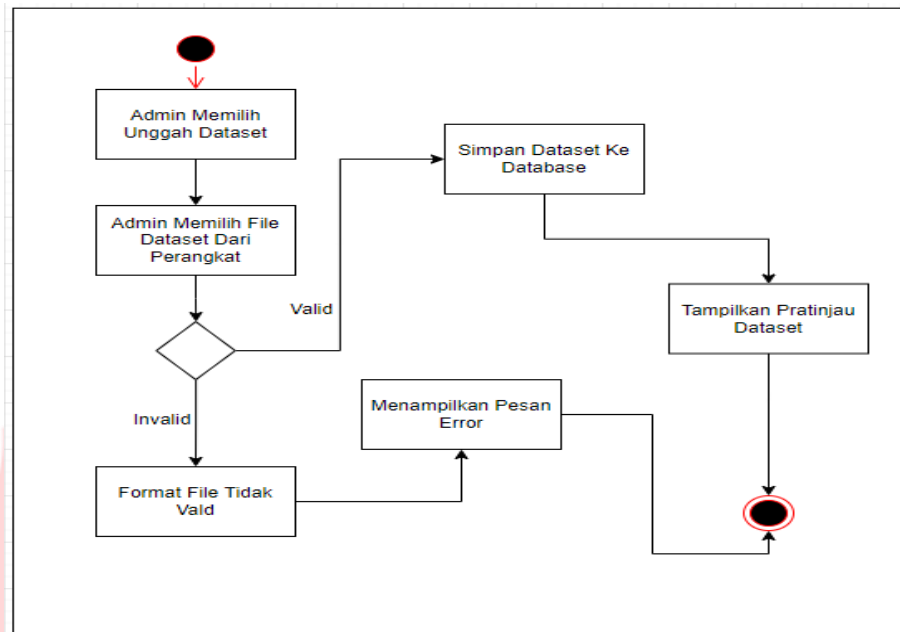
3.6.3 *Activity Diagram*

1. *Activity Diagram Login*



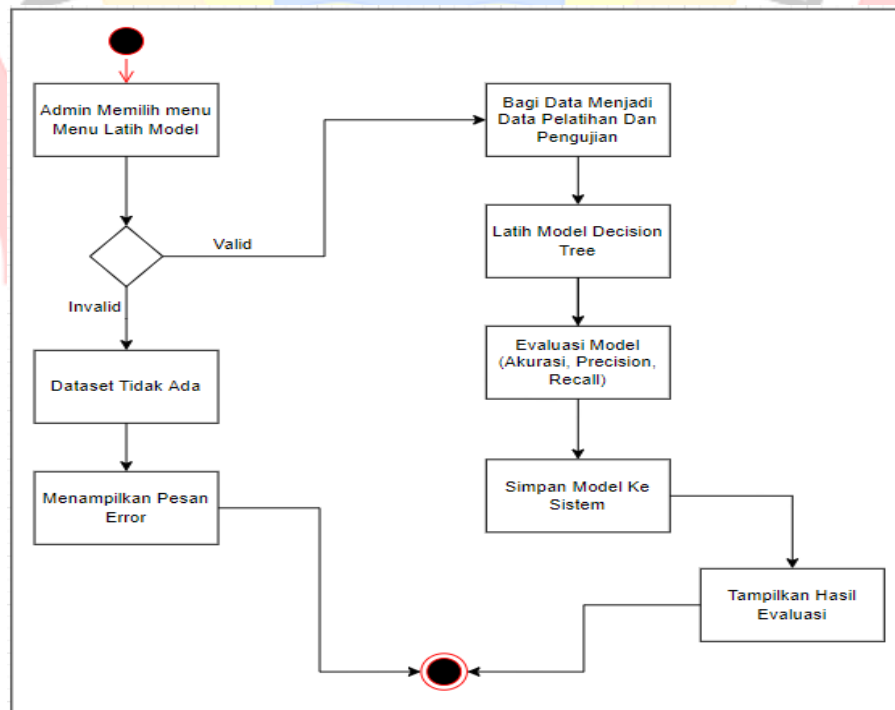
Gambar 3. 4 *Activity Diagram Login*

2. Activity Diagram Upload Dataset



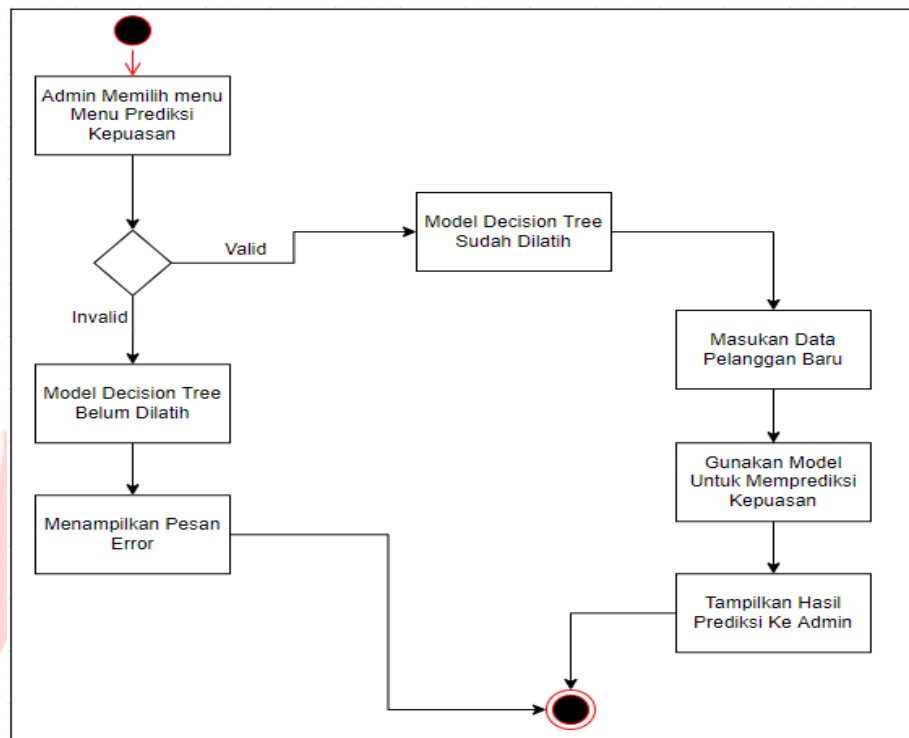
Gambar 3. 5 Activity Diagram Upload Dataset

3. Activity Diagram Latih Model



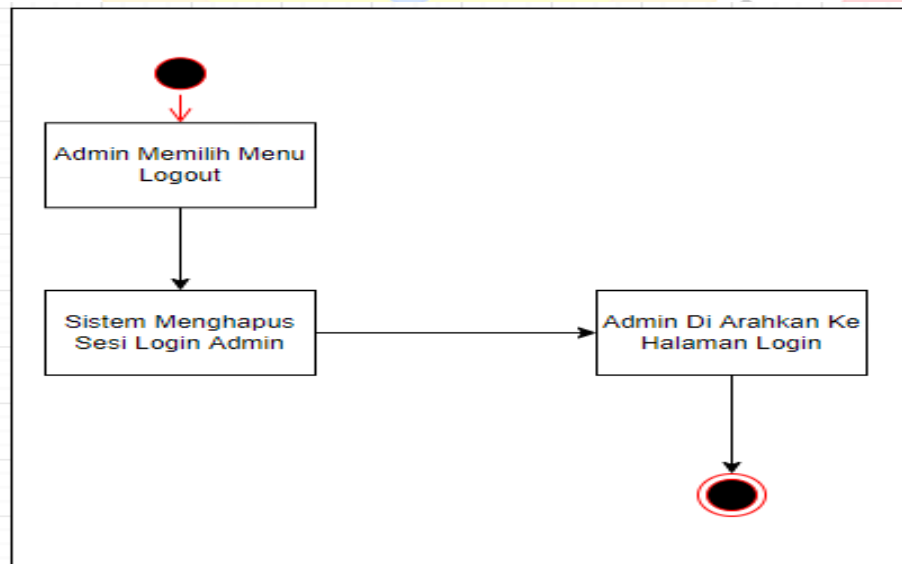
Gambar 3. 6 Activity Diagram Latih Model

4. Activity Diagram Prediksi Kepuasan



Gambar 3. 7 Activity Diagram Prediksi Kepuasan

5. Activity Diagram Logout

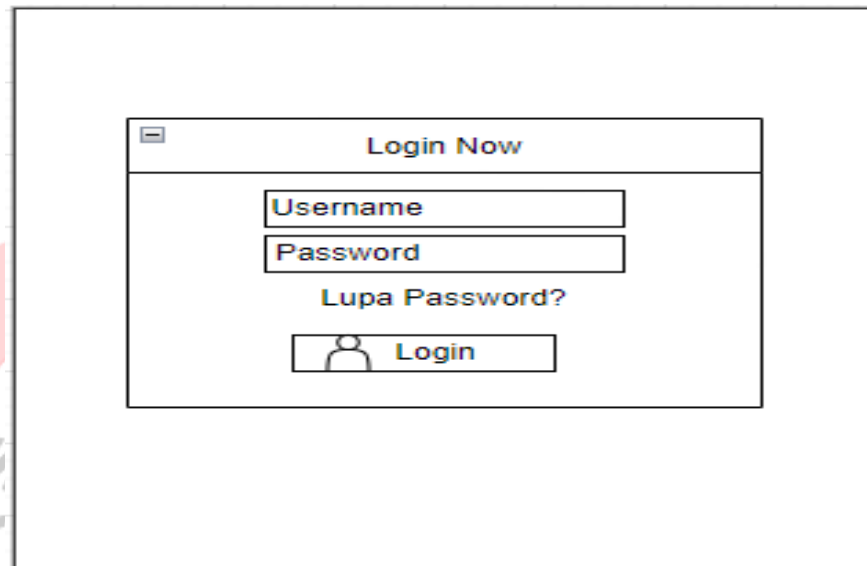


Gambar 3. 8 Activity Diagram Latih Model

3.7 Perancangan Layar

3.7.1 Perancangan Menu

1. *Login*

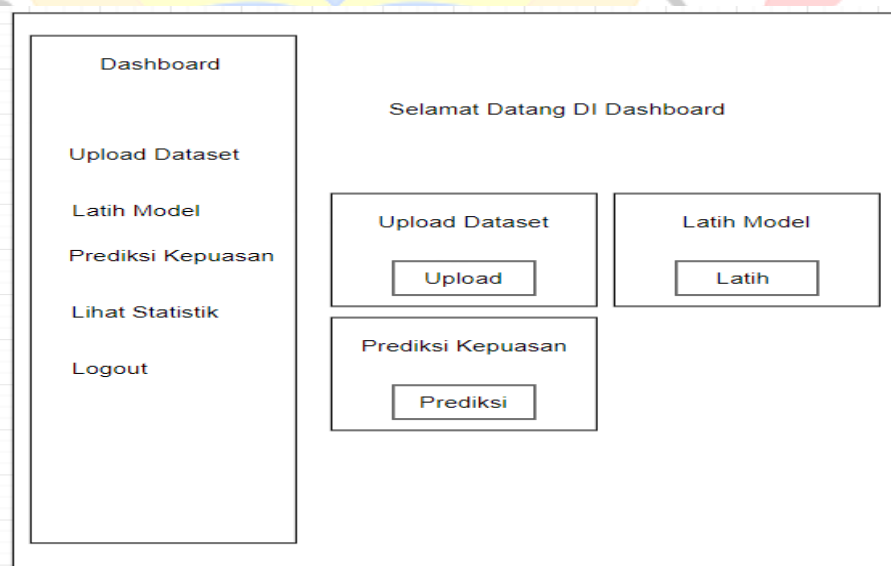


A wireframe of a login form titled "Login Now". The form contains two input fields for "Username" and "Password", a link "Lupa Password?", and a "Login" button with a user icon.

Login Now	
Username	
Password	
Lupa Password?	
	Login

Gambar 3. 9 Perancangan Menu *Login*

2. *Dashboard*

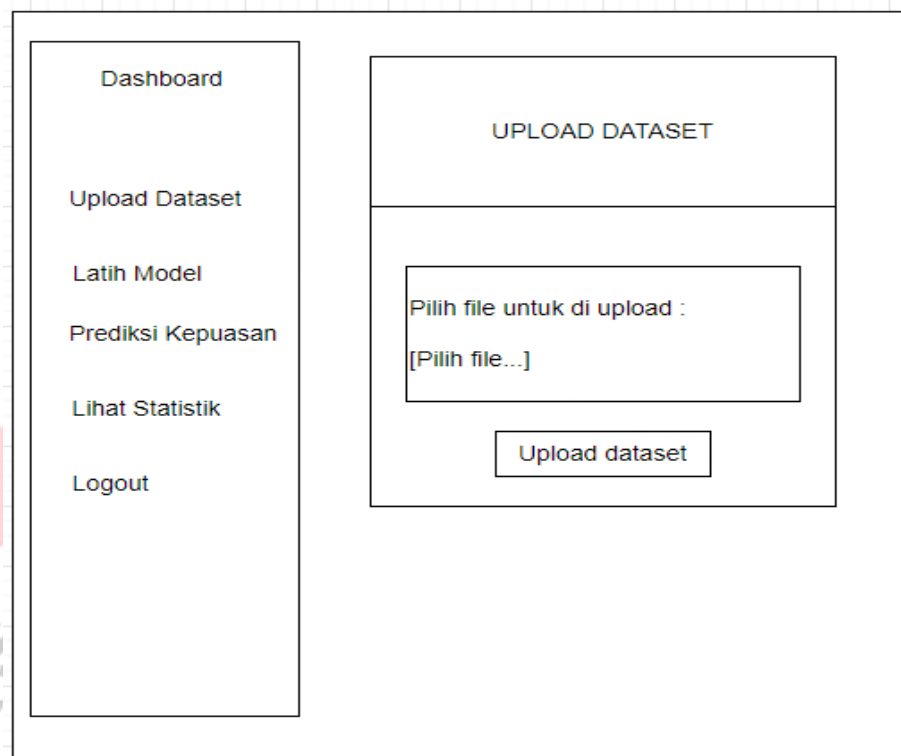


A wireframe of a dashboard layout. It features a sidebar menu on the left with options: Dashboard, Upload Dataset, Latih Model, Prediksi Kepuasan, Lihat Statistik, and Logout. The main content area has a greeting "Selamat Datang DI Dashboard" and three action buttons: "Upload Dataset" (with an "Upload" sub-button), "Latih Model" (with a "Latih" sub-button), and "Prediksi Kepuasan" (with a "Prediksi" sub-button).

Dashboard	Selamat Datang DI Dashboard	
Upload Dataset	Upload Dataset	Latih Model
Latih Model	Upload	Latih
Prediksi Kepuasan		
Lihat Statistik	Prediksi Kepuasan	
Logout	Prediksi	

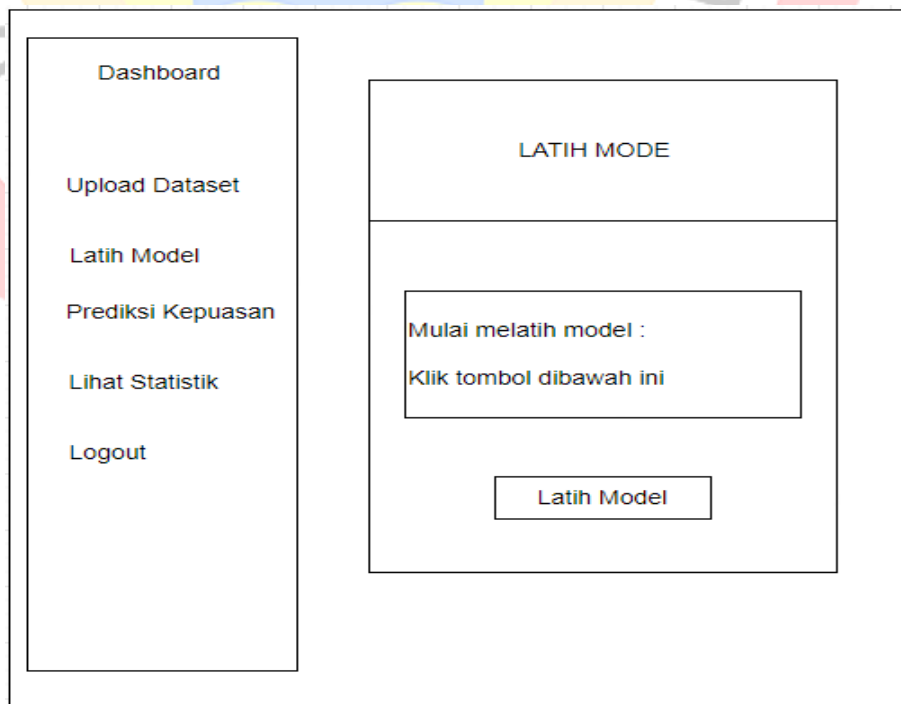
Gambar 3. 10 Perancangan Menu *Dashboard*

3. *Upload Dataset*



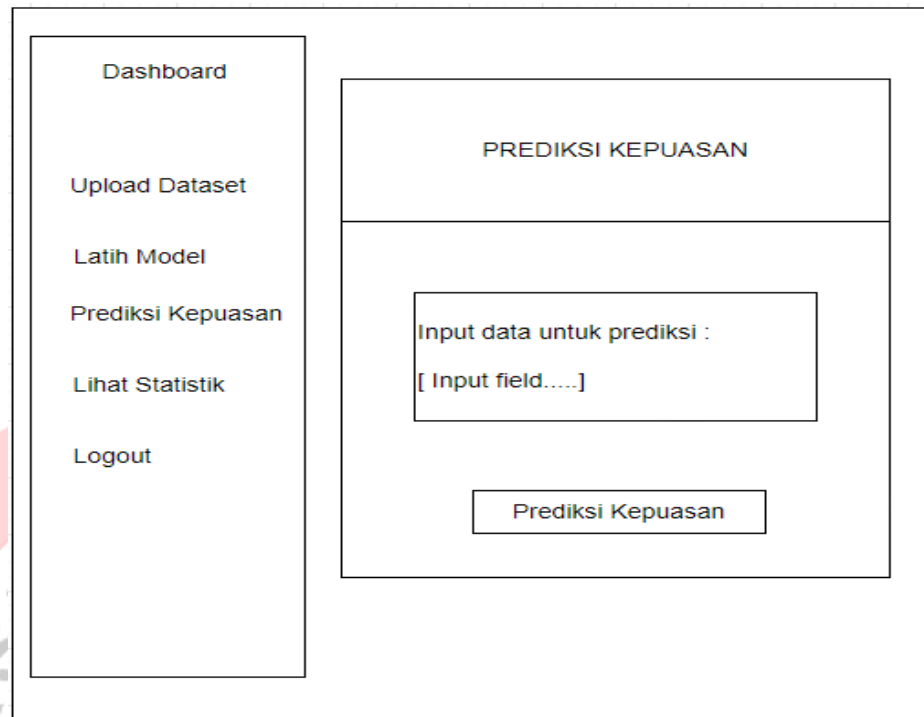
Gambar 3. 11 Perancangan Menu *Upload Dataset*

4. *Latih model*



Gambar 3. 12 Perancangna Menu Latih Model

5. Prediksi Kepuasan



Dashboard

Upload Dataset

Latih Model

Prediksi Kepuasan

Lihat Statistik

Logout

PREDIKSI KEPUASAN

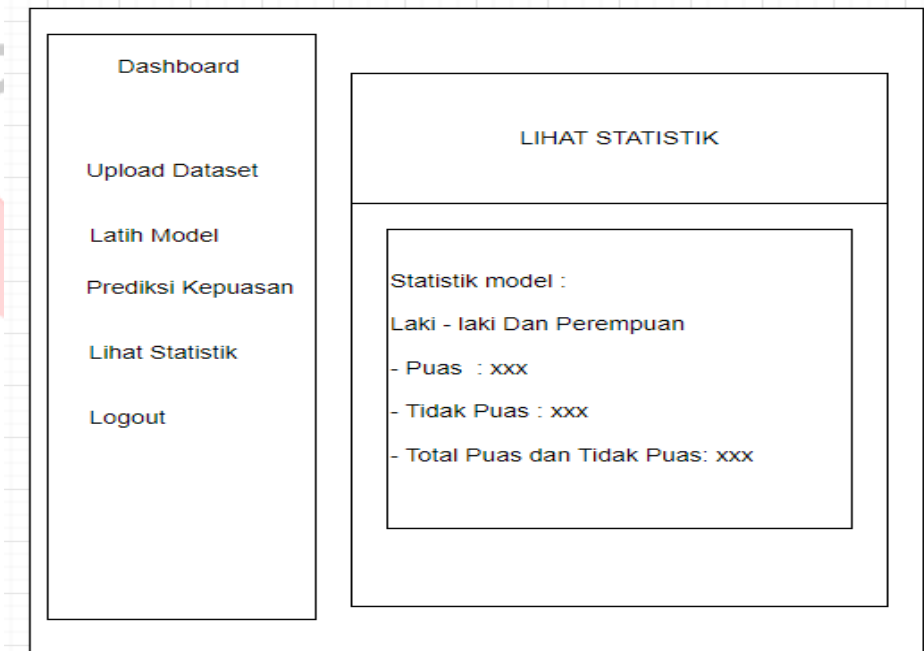
Input data untuk prediksi :

[Input field.....]

Prediksi Kepuasan

Gambar 3. 13 Perancangan Menu Prediksi Kepuasan

6. Lihat Statistik



Dashboard

Upload Dataset

Latih Model

Prediksi Kepuasan

Lihat Statistik

Logout

LIHAT STATISTIK

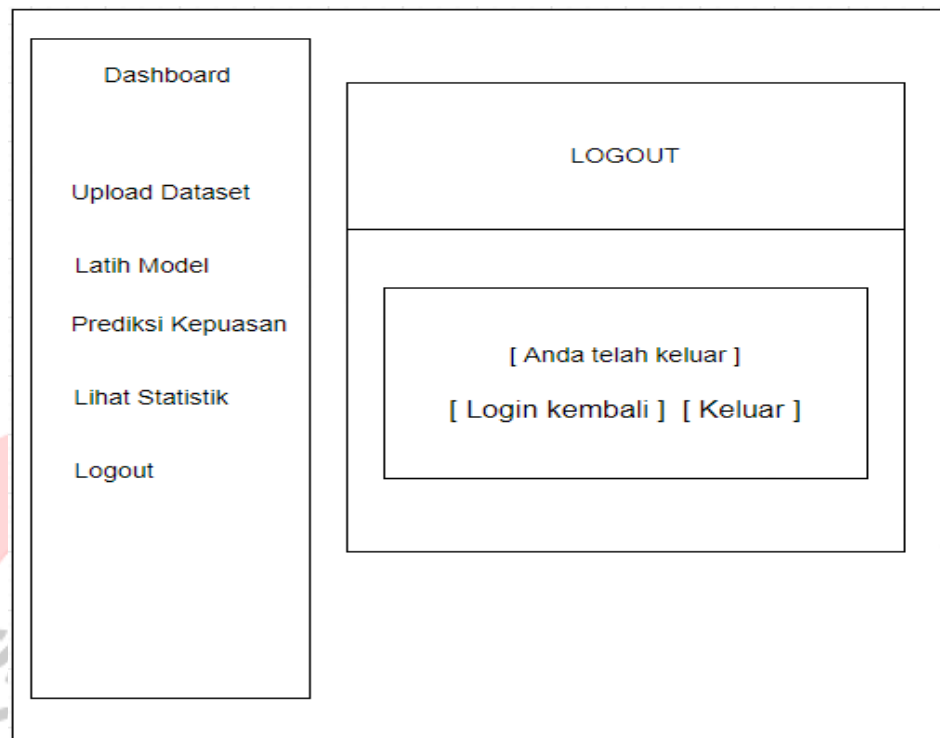
Statistik model :

Laki - laki Dan Perempuan

- Puas : xxx
- Tidak Puas : xxx
- Total Puas dan Tidak Puas: xxx

Gambar 3. 14 Perancangan Menu Lihat Statistik

7. Logout



Gambar 3. 15 Perancangan Menu *Logout*

3.7.2 Perancangan *Database*

1. Tabel-Tabel *Database*

a. Tabel pelanggan

Tabel 3. 8 Tabel *database* Pelanggan

Kolom	Type Data	Keterangan
nama_pelanggan	VARCHAR(20)	Nama pelanggan
User_ID	INT PRIMARY KEY	ID unik pelanggan
alamat	VARCHAR(20)	Alamat pelanggan
tikor_CST	VARCHAR(50)	Titik koordinat customer
ODP	VARCHAR(50)	Titik odp yang menjangkau wilayah pelanggan

Nama produk	VARCHAR(20)	Nama-nama produk perusahaan
<i>Bandwidth (mbps)</i>	INT	Kecepatan jaringan per- mbps
Harga (Rp)	INT	Harga produk
lama_berlangganan	INT	Lama berlangganan pelanggan
jenis_kelamin	ENUM('L', 'P')	Laki-laki (L) atau Perempuan (P)
kepuasan	ENUM('Puas', 'Tidak Puas')	Status kepuasan pelanggan (jika ada data <i>historis</i>)

b. Tabel *dataset*

Tabel 3. 9 Tabel *Dataset*

Kolom	Tipe Data	Keterangan
<i>id_dataset</i>	INT PRIMARY KEY	ID unik <i>dataset</i>
nama_file	VARCHAR(15)	Nama file <i>dataset</i> yang diunggah
tanggal_unggah	DATETIME	Tanggal dan waktu unggah <i>dataset</i>
jumlah_record	INT	Jumlah <i>record</i> dalam <i>dataset</i>

user_pengunggah	INT	ID user yang mengunggah <i>dataset</i> (relasi ke <i>user</i>)
-----------------	-----	---

c. Tabel prediksi

Tabel 3. 10 Tabel Prediksi

Kolom	Tipe Data	Keterangan
id_prediksi	INT PRIMARY KEY	ID unik prediksi
id_pelanggan	INT	ID pelanggan (relasi ke pelanggan)
tanggal_prediksi	DATETIME	Tanggal dan waktu prediksi
hasil_prediksi	ENUM('Puas', 'Tidak Puas')	Hasil prediksi (<i>output</i> model)
user_prediksi	INT	ID user yang melakukan prediksi (relasi ke <i>user</i>)

d. Tabel user

Tabel 3. 11 Tabel User

Kolom	Tipe Data	Keterangan
id_user	INT PRIMARY KEY	ID unik user
username	VARCHAR(20)	Nama pengguna
password	VARCHAR(12)	Password (di-hash)
role	ENUM('admin', 'user')	Peran pengguna dalam sistem (<i>admin</i> atau <i>user</i>)

3.8 Jadwal Penelitian

Gant Chart								
Proses	Tahap 1				Tahap 2			
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
Penentuan Judul								
Perancangan Teori								
Analisis Teori								
Analisis Permasalahan								
Desain Sistem								
Perancangan Sistem								
Pembahasan Dan hasil								
Kesimpulan								

Gambar 3. 16 Jadwal Penelitian

