

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1. Analisis Sistem

Sistem prediksi mengenai ketepatan pembayaran *customer* diproses menggunakan metode *decision tree*. Sistem prediksi ini diperlukan bagi seorang analis-penagihan dalam mengambil langkah-langkah terkait kebijakan pengiriman dan pembayaran barang kepada *customer*, sehingga resiko keterlambatan atau gagal bayar dari *customer* dapat diminimalisir.

Analisis-penagihan perusahaan memiliki beberapa indikator yang menjadi pertimbangan dalam menentukan kebijakan pengiriman dan pembayaran *customer*. Indikator pertama adalah jenis bisnis *customer* yang meliputi kontraktor, manufaktur, pariwisata, dan tambang. Indikator selanjutnya tentang pemakaian jasa broker, adanya *cashback*, tipe hari jatuh tempo, dan tipe termin pembayaran yang meliputi *cash on delivery*, tempo 7 hari, tempo 14 hari, dan tempo 30 hari. Selama ini analisa dilakukan hanya dengan metode tanya jawab dan kunjungan lokasi *customer*, hal ini memungkinkan adanya rekayasa jawaban dari pihak *customer*.

Dalam membuat penelitian ini penulis menggunakan data hasil penjualan PT Berdikari Jaya Bersama di rentang tahun 2021-2022 yang berjumlah 606 data, data tersebut dibagi menjadi dua bagian yaitu data latih sebanyak 420 data dan data uji sebanyak 186 data. Data tersebut selanjutnya diolah menggunakan metode *decision tree* hingga menghasilkan pohon keputusan dalam memprediksi ketepatan pembayaran *customer*. Diharapkan dari hasil penelitian ini, tim analis-penagihan PT Berdikari Jaya Bersama dapat memiliki tambahan pertimbangan yang lebih akurat dalam memprediksi ketepatan pembayaran *customer*.

3.1.1. Representasi Data

Data yang digunakan dalam pengujian sistem prediksi ketepatan pembayaran *customer* bersumber dari data hasil penjualan PT Berdikari Jaya Bersama di rentang tahun 2021-2022 yang berjumlah 606 data. Data-data tersebut berisikan informasi penjualan *customer* seperti informasi kontak, alamat *customer*, alamat pengiriman, *volume* dan catatan lainnya. Beberapa informasinya dapat digunakan sebagai atribut

dalam memprediksi ketepatan pembayaran *customer*, diantaranya jenis bisnis, broker, *cashback*, hari jatuh tempo dan termin. Mengenai informasi hasil ketepatan pembayaran di data-data tersebut dapat diketahui dari tanggal pembayaran lunas tagihan, apakah melewati atau tidak dari tanggal jatuh temponya. Apabila tidak melewati maka keterangannya “tepat waktu”, jika melewati maka keterangannya “terlambat”.

Tabel 3.1. Data Atribut

No	Atribut	Detail Atribut
1	Jenis Bisnis	<ol style="list-style-type: none"> 1. Kontraktor 2. Manufaktur 3. Pariwisata 4. Tambang
2	Broker	<ol style="list-style-type: none"> 1. Ya 2. Tidak
3	<i>Cashback</i>	<ol style="list-style-type: none"> 1. Ada 2. Tidak Ada
4	Hari Jatuh Tempo	<ol style="list-style-type: none"> 1. Hari Biasa 2. Hari Libur
5	Termin	<ol style="list-style-type: none"> 1. <i>Cash On Delivery (COD)</i> 2. Tempo 7 Hari 3. Tempo 14 Hari 4. Tempo 30 Hari
6	Hasil	<ol style="list-style-type: none"> 1. Tepat Waktu 2. Terlambat 3. Dominan Tepat Waktu 4. Dominan Terlambat

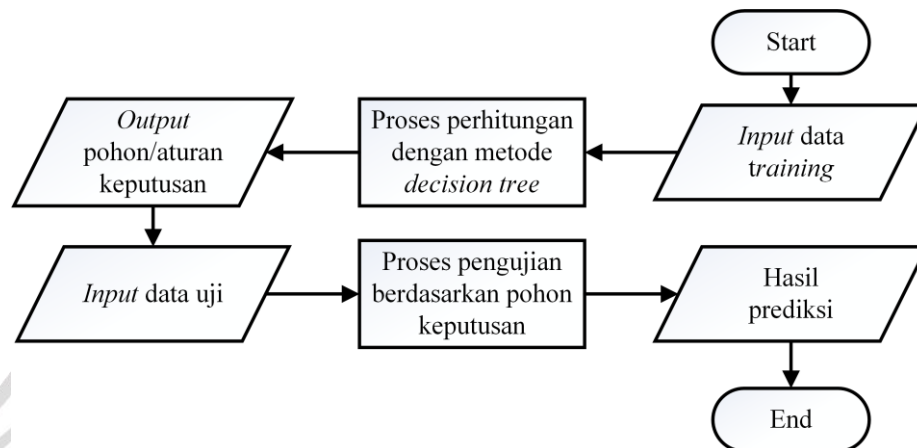
3.2. Perancangan Sistem

Setelah analisis sistem dilakukan maka dilanjutkan dengan perancangan sistem prediksi ketepatan pembayaran *customer*. Perancangan ini akan tersusun

dalam bentuk *flowchart*, diagram konteks, diagram berjenjang *data flow diagram* (DFD).

3.2.1. Flowchart

Berikut ini adalah *flowchart* dari sistem prediksi ketepatan pembayaran *customer* :



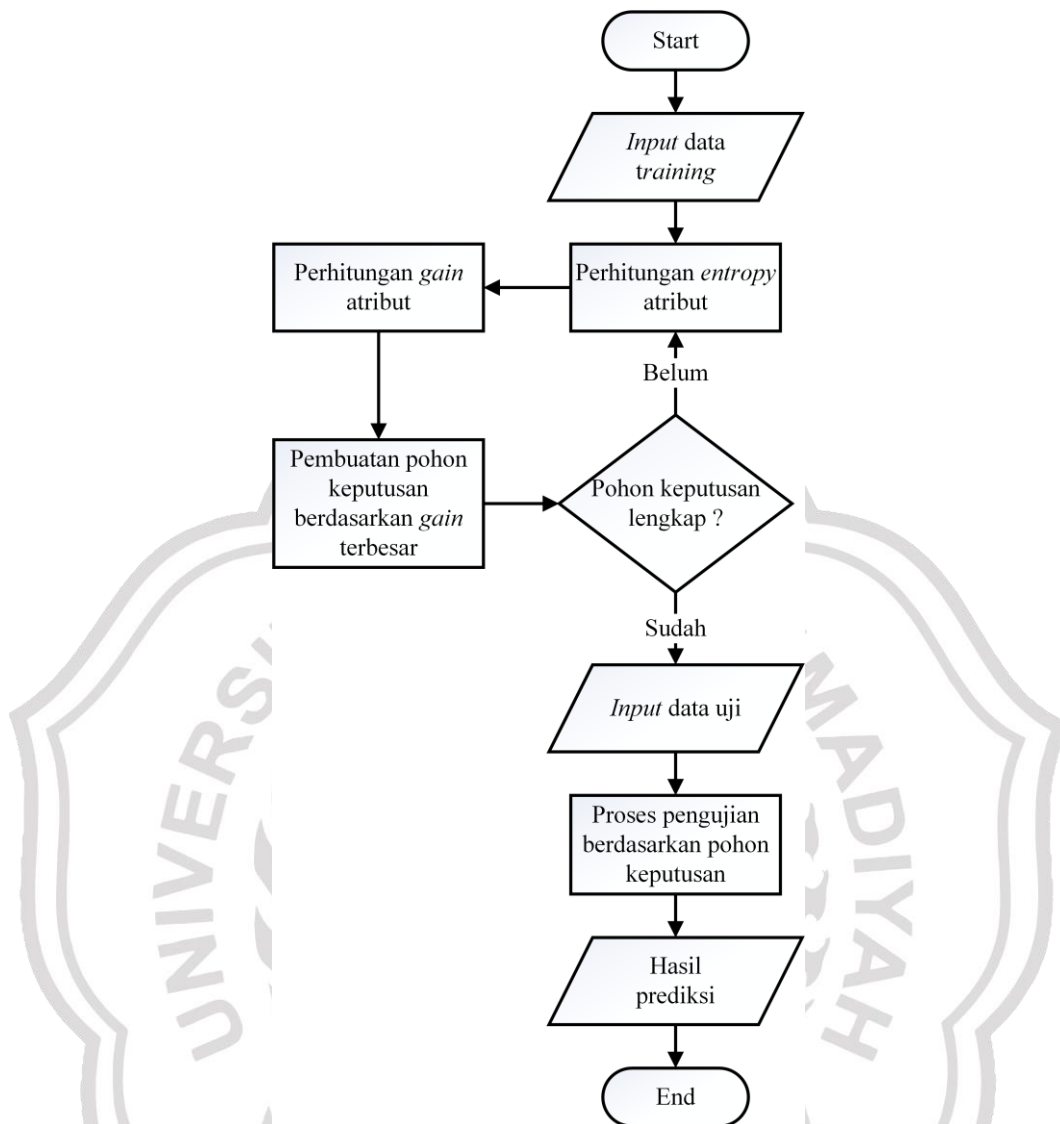
Gambar 3.1. *Flowchart* sistem prediksi

Berikut adalah penjelasan dari gambar 3.1 yang merupakan *flowchart* dari sistem prediksi ketepatan pembayaran *customer* :

1. *User* memasukkan data *training* yang berisikan data hasil penjualan PT Berdikari Jaya Bersama kedalam sistem. Data *training* yang dimasukkan hanya mengandung atribut yang telah ditentukan pada tabel 3.1.
2. Sistem memproses data *training* menggunakan metode *decision tree*.
3. Sistem melaporkan hasil *output* berupa pohon keputusan.
4. *User* memasukkan data uji, data uji dimasukkan satu per satu.
5. Sistem memproses data uji berdasarkan urutan dari pohon keputusan yang telah terbentuk.
6. Sistem melaporkan hasil prediksi dari data uji.

3.2.1.1. Flowchart Sistem Menggunakan Metode *Decision Tree*

Berikut ini adalah *flowchart* dari sistem prediksi ketepatan pembayaran *customer* yang menggunakan metode *decision tree* :



Gambar 3.2. Flowchart sistem prediksi dengan metode *decision tree*

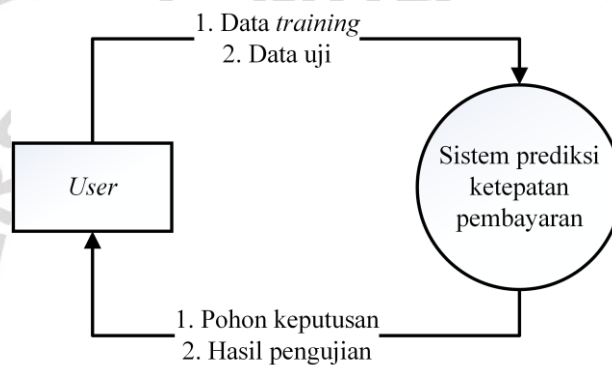
Berikut adalah penjelasan dari gambar 3.2 yang merupakan *flowchart* dari sistem prediksi ketepatan pembayaran *customer* dengan metode *decision tree* :

1. *User* memasukkan data *training* yang berisikan data hasil penjualan PT Berdikari Jaya Bersama kedalam sistem. Data *training* yang dimasukkan hanya mengandung atribut yang telah ditentukan pada tabel 3.1.
2. Sistem menghitung nilai *entropy* pada tiap atribut.
3. Sistem menghitung nilai *gain* pada tiap atribut.
4. *User* memilih *gain* terbesar sebagai *node* akar / *node* berikutnya.

5. Sistem menampilkan hasil pohon keputusan, apabila belum lengkap maka sistem akan mengulangi dari langkah ke-2 hingga pohon keputusan lengkap atau memiliki *leaf node* di tiap *node* paling akhir.
6. Jika sudah lengkap, *user* dapat memasukkan data uji satu per satu ke sistem
7. Sistem memproses data uji berdasarkan urutan dari pohon keputusan yang telah terbentuk.
8. Sistem melaporkan hasil prediksi dari data uji.

3.2.2. Diagram Konteks

Berikut ini adalah diagram konteks dari sistem prediksi ketepatan pembayaran *customer* :

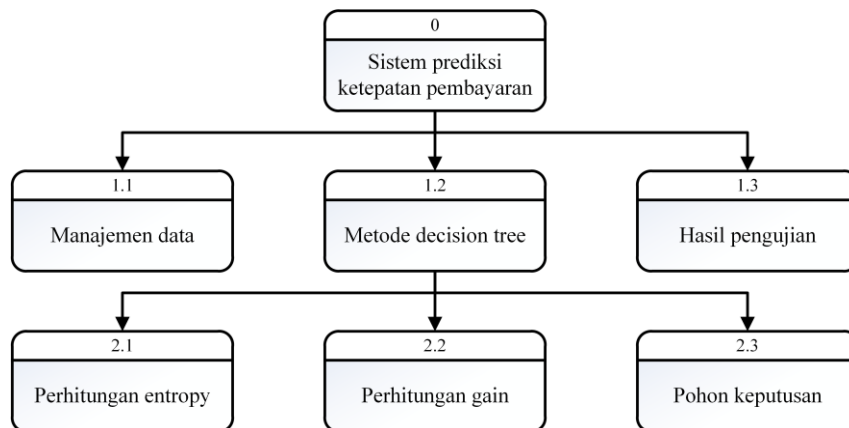


Gambar 3.3. Diagram konteks sistem prediksi

Gambar 3.3 merupakan diagram konteks dari sistem prediksi ketepatan pembayaran *customer*. Pada diagram di atas dijelaskan bahwa *user* yang dalam perannya sebagai tim analisa-penagihan adalah *entity* tunggal pada proses sistem ini. *User* dapat memasukkan data *training* dan data *uji* ke dalam sistem sesuai dengan atribut yang ditetapkan pada tabel 3.1 . Kemudian *user* akan mendapatkan keluaran berupa pohon keputusan dari sistem yang telah memproses data *training* menggunakan metode *decision tree* serta hasil pengujian dari data uji yang dimasukkan.

3.2.3. Diagram Berjenjang

Berikut ini adalah diagram berjenjang dari sistem prediksi ketepatan pembayaran *customer* :



Gambar 3.4. Diagram berjenjang sistem prediksi

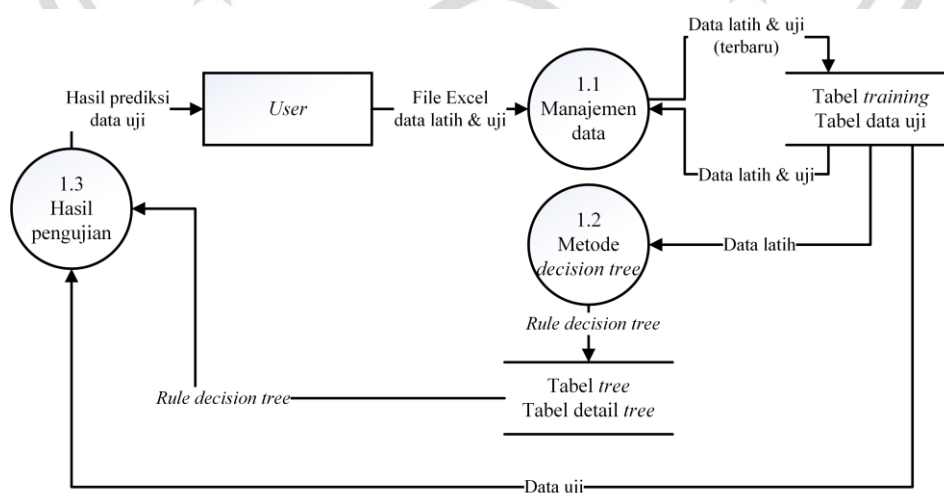
Gambar 3.4 merupakan diagram berjenjang dari sistem prediksi ketepatan pembayaran *customer*. Berikut adalah penjelasan dari gambar 3.4 :

1. *Top level* : konsep inti sistem ini adalah sistem prediksi ketepatan pembayaran *customer* dengan metode *decision tree*.
2. *Level 1* : berisi proses yang meliputi manajemen data berupa *input* data *training* dan data uji, metode *decision tree* dan hasil pengujian dari data uji.
3. *Level 2* : berisi proses dari metode *decision tree*, antara lain perhitungan *entropy*, perhitungan *gain*, dan keluaran berupa pohon keputusan.

3.2.4. Data Flow Diagram (DFD)

3.2.4.1. DFD Level 1

Berikut ini adalah DFD *level 1* dari sistem prediksi ketepatan pembayaran *customer* :



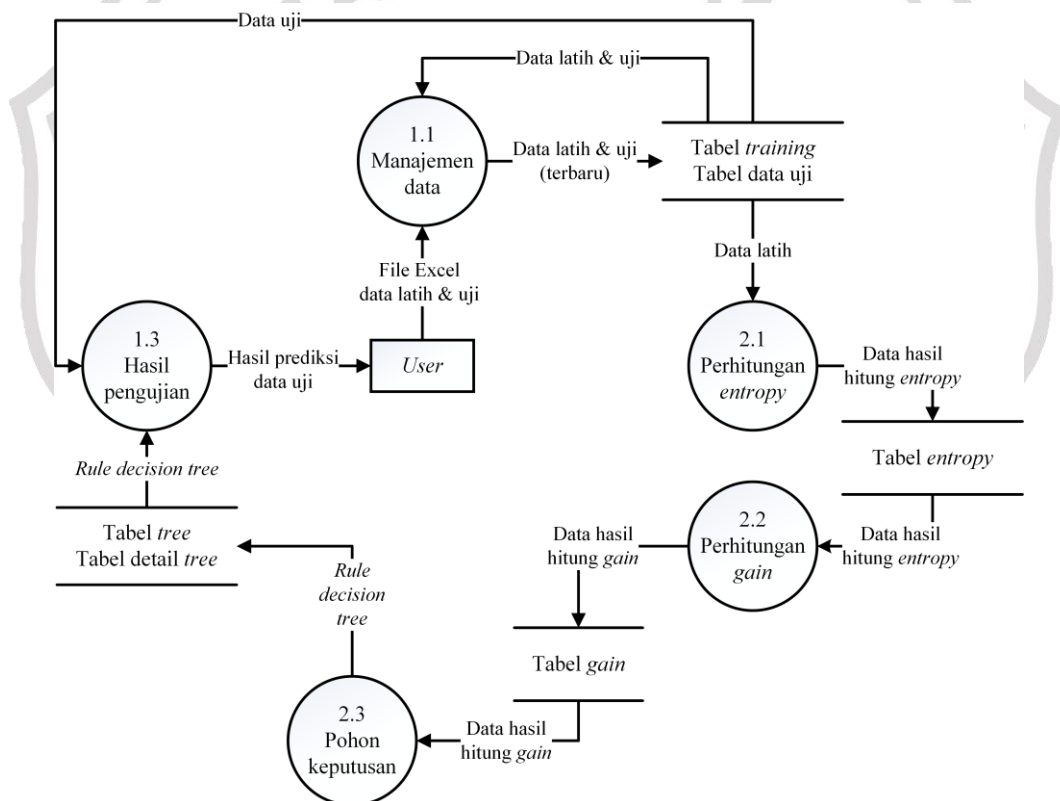
Gambar 3.5. DFD *level 1* sistem prediksi

Gambar 3.5 merupakan DFD *level 1* dari sistem prediksi ketepatan pembayaran *customer*. Di dalamnya terdapat beberapa proses antara lain :

- 1.1 *User* memasukkan data latih dan data uji ke dalam sistem. Masing-masing data tersebut disimpan ke tabel *training* dan tabel data uji
- 1.2 Sistem melakukan proses pengolahan menggunakan metode *decision tree* terhadap data yang sudah dimasukkan *user*.
- 1.3 Sistem menampilkan hasil pengujian terhadap data uji berdasarkan hasil pengolahan data di proses sebelumnya.

3.2.4.2. DFD *Level 2*

Berikut ini adalah DFD *level 2* dari sistem prediksi ketepatan pembayaran *customer* :



Gambar 3.6. DFD *level 2* sistem prediksi

Gambar 3.6 merupakan DFD *level 2* dari sistem prediksi ketepatan pembayaran *customer*. Di dalamnya terdapat beberapa proses antara lain :

- 2.1. Sistem melakukan perhitungan *entropy* di tiap atribut, hasil perhitungan disimpan ke dalam tabel *entropy* untuk diolah di proses selanjutnya.

- 2.2. Sistem melakukan perhitungan *gain* di tiap atribut setelah memperoleh nilai *entropy* dari proses sebelumnya. Hasil perhitungan disimpan pada tabel *gain*.
- 2.3. Sistem melakukan proses pembuatan pohon keputusan atau *rule decision tree*, data diperoleh dari tabel *gain*.

3.3. Perancangan Basis Data

Pada sistem prediksi ketepatan pembayaran *customer* mempunyai sebuah basis data atau *database* untuk menyimpan data-data yang ada. Di dalamnya terdapat struktur dari berbagai tabel.

3.3.1. Tabel *Training*

Tabel *training* digunakan untuk menyimpan data *training* yang akan diolah. Struktur dari tabel *training* dapat dilihat pada tabel 3.2 :

Tabel 3.2. Struktur tabel *training*

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
No_data	Int	11	
Bisnis	Varchar	100	
Broker	Varchar	100	
<i>Cashback</i>	Varchar	100	
Jatuh_tempo	Varchar	100	
Termin	Varchar	100	
Prediksi	Varchar	100	

3.3.2. Tabel Atribut

Tabel atribut digunakan untuk menyimpan data atribut yang digunakan pada sistem ini. Struktur dari tabel atribut dapat dilihat pada tabel 3.3 :

Tabel 3.3. Struktur tabel atribut

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Atribut	Varchar	50	

3.3.3. Tabel Detail Atribut

Tabel detail atribut digunakan untuk menyimpan data detail atribut, tabel ini berelasi dengan tabel atribut, tabel detail atribut mengambil nilai “id” dari tabel atribut. Struktur dari tabel detail atribut dapat dilihat pada tabel 3.4 :

Tabel 3.4. Struktur tabel detail atribut

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Atribut_id	Int	11	
Detail	Varchar	50	

3.3.4. Tabel Entropy

Tabel *entropy* digunakan untuk menyimpan data hasil perhitungan *entropy* semua atribut pada setiap *node*, sehingga tabel ini memiliki tabel duplikat yang berstruktur sama sejumlah banyaknya *node* yang dibuat. Tabel ini berelasi dengan tabel atribut melalui *field* “atribut_id”. Struktur dari tabel *entropy* dapat dilihat pada tabel 3.5 :

Tabel 3.5. Struktur tabel *entropy*

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Atribut_id	Int	11	
Detail	Varchar	50	
Jumlah	Int	11	
Tepat_waktu	Int	11	
Terlambat	Int	11	
<i>Entropy</i>	Float	11	
<i>Log_gain</i>	Float	11	

3.3.5. Tabel Gain

Tabel *gain* digunakan untuk menyimpan data hasil perhitungan *gain* pada setiap *node*, sehingga tabel ini memiliki tabel duplikat yang berstruktur sama

sejumlah banyaknya *node* yang dibuat. Tabel ini berelasi dengan tabel atribut melalui *field* “atribut_id”.

Struktur dari tabel *gain* dapat dilihat pada tabel 3.6 :

Tabel 3.6. Struktur tabel *gain*

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Atribut_id	Int	11	
<i>Gain</i>	Float	11	

3.3.6. Tabel *Tree*

Tabel *tree* digunakan untuk menyimpan urutan data *node* hasil pembentukan pohon keputusan, *field* “id” menjadi nomor urut dari *node* pohon keputusan. Tabel ini berelasi dengan tabel atribut melalui *field* “atribut_id”. Struktur dari tabel *tree* dapat dilihat pada tabel 3.7 :

Tabel 3.7. Struktur tabel *tree*

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Atribut_id	Int	11	

3.3.7. Tabel Detail *Tree*

Tabel detail *tree* digunakan untuk menyimpan detail data *node* hasil pembentukan pohon keputusan. Tabel ini berelasi dengan tabel *tree* melalui *field* “tree_id” serta *field* “next_tree_id” dan tabel detail atribut melalui *field* “detail_atribut_id”. Struktur dari tabel detail *tree* dapat dilihat pada tabel 3.8 :

Tabel 3.8. Struktur tabel detail *tree*

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
<i>Tree_id</i>	Int	11	
Detail_atribut_id	Int	11	
<i>Next_tree_id</i>	Int	11	

<i>Leaf_node</i>	Varchar	50	
------------------	---------	----	--

3.3.8. Tabel Data Uji

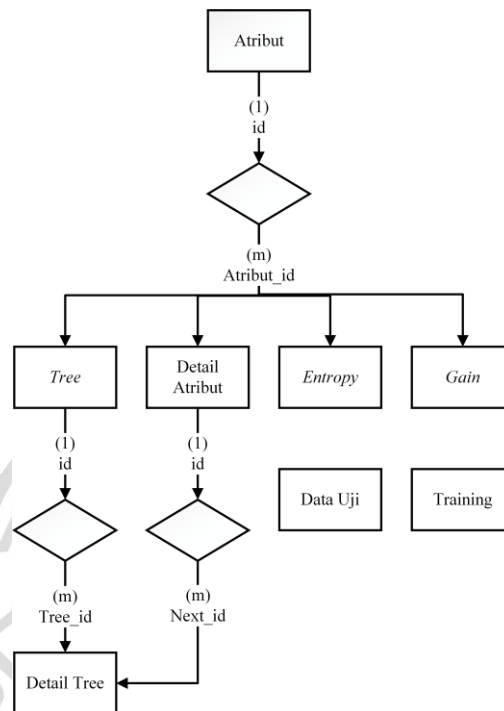
Tabel data uji digunakan untuk menyimpan data uji yang dimasukkan *user*. Struktur dari tabel data uji dapat dilihat pada tabel 3.9 :

Tabel 3.9. Struktur tabel data uji

Nama Field	Tipe Data	Length	Key
Id	Int	11	<i>Primary Key</i>
Perusahaan	Varchar	100	
Bisnis	Varchar	100	
Broker	Varchar	100	
<i>Cashback</i>	Varchar	100	
Jatuh_tempo	Varchar	100	
Termin	Varchar	100	
Prediksi	Varchar	100	

3.3.9. Relasi Tabel Database

Gambar relasi tabel pada sistem prediksi ketepatan pembayaran *customer* dapat dilihat pada gambar 3.7



Gambar 3.7. Relasi tabel *database*

3.4. Perancangan Pengujian

Perancangan dan pengujian ini akan menggunakan sejumlah data hasil penjualan PT Berdikari Jaya Bersama rentang tahun 2021-2022. Berikut ini adalah contoh pengujian pada penelitian ini.

Tabel 3.10. Tabel data latih

No	Bisnis	Broker	Cashback	Jatuh Tempo	Termin	Hasil
1	Pariwisata	Tidak	Tidak Ada	Hari Libur	Tempo 14 Hari	Tepat Waktu
2	Tambang	Tidak	Tidak Ada	Hari Libur	Tempo 14 Hari	Tepat Waktu
3	Kontraktor	Ya	Ada	Hari Biasa	Cash On Delivery	Tepat Waktu
4	Kontraktor	Ya	Ada	Hari Libur	Tempo 7 Hari	Tepat Waktu

5	Tambang	Tidak	Tidak Ada	Hari Libur	Tempo 30 Hari	Tepat Waktu
6	Manufaktur	Ya	Ada	Hari Biasa	Cash On Delivery	Tepat Waktu
7	Kontraktor	Tidak	Tidak Ada	Hari Biasa	Tempo 7 Hari	Tepat Waktu
8	Manufaktur	Ya	Ada	Hari Biasa	Tempo 14 Hari	Tepat Waktu
9	Pariwisata	Tidak	Tidak Ada	Hari Biasa	Tempo 30 Hari	Tepat Waktu
10	Tambang	Tidak	Tidak Ada	Hari Biasa	Cash On Delivery	Tepat Waktu
11	Tambang	Tidak	Tidak Ada	Hari Biasa	Tempo 14 Hari	Terlambat
12	Pariwisata	Tidak	Tidak Ada	Hari Biasa	Cash On Delivery	Terlambat
13	Manufaktur	Tidak	Tidak Ada	Hari Biasa	Tempo 7 Hari	Terlambat
14	Tambang	Tidak	Ada	Hari Biasa	Tempo 7 Hari	Terlambat
15	Manufaktur	Tidak	Ada	Hari Biasa	Cash On Delivery	Terlambat
16	Manufaktur	Ya	Ada	Hari Libur	Tempo 30 Hari	Terlambat
17	Kontraktor	Tidak	Tidak Ada	Hari Biasa	Tempo 30 Hari	Terlambat
18	Kontraktor	Tidak	Ada	Hari Biasa	Tempo 14 Hari	Terlambat
19	Pariwisata	Tidak	Ada	Hari Libur	Tempo 7 Hari	Terlambat

20	Pariwisata	Tidak	Tidak Ada	Hari Libur	Tempo 30 Hari	Terlambat
----	------------	-------	--------------	------------	------------------	-----------

Data yang terdapat pada tabel 3.10. merupakan data latih yang digunakan sebagai contoh pengujian. Proses pengujian dimulai dengan membuat pohon keputusan berdasarkan data latih. Sebelum melakukan proses perhitungan, berikut merupakan beberapa ketentuan pada penelitian ini dalam menentukan *leaf node* pada pohon keputusan :

1. Apabila *entropy* bernilai 0 pada *node* terpilih, dengan kondisi kelas “tepat waktu” lebih besar dari kelas “terlambat”, dan kelas “terlambat” bernilai sama dengan 0 (nol), maka *leaf node* nya “tepat waktu”.
2. Apabila *entropy* bernilai 0 pada *node* terpilih, dengan kondisi kelas “terlambat” lebih besar dari kelas “tepat waktu”, dan kelas “tepat waktu” bernilai sama dengan 0 (nol), maka *leaf node* nya “terlambat”.
3. Apabila *entropy* bernilai 0 pada *node* terpilih, dengan kondisi kelas “terlambat” dan kelas “tepat waktu” bernilai sama dengan 0 (nol), maka *leaf node* nya diambil dari jumlah data mayoritas pada *node* tersebut.
4. Apabila sampai pada *node* terakhir terdapat *entropy* yang tidak bernilai 0 (nol), dengan kondisi kelas “tepat waktu” lebih besar dari kelas “terlambat”, maka *leaf node* nya “dominan tepat waktu”.
5. Apabila sampai pada *node* terakhir terdapat *entropy* yang tidak bernilai 0 (nol), dengan kondisi kelas “terlambat” lebih besar dari kelas “tepat waktu”, maka *leaf node* nya “dominan terlambat”.
6. Apabila sampai pada *node* terakhir terdapat *entropy* yang tidak bernilai 0 (nol), dengan kondisi kelas “terlambat” sama dengan kelas “tepat waktu”, maka *leaf node* nya “dominan terlambat”. Karena prediksi terlambat bayar customer sama dengan prediksi tepat waktu, sehingga diambil prediksi paling negatif atau buruk.

Berikut adalah contoh proses perhitungan pada penelitian ini :

1. *Node 1*

Sebelum melakukan perhitungan *entropy* dan *gain* pada tiap *node*, data harus dikelompokkan terlebih dahulu sesuai dengan atribut dan detail atributnya. Pengelompokkan data artinya menghitung jumlah data berdasarkan kesamaan atribut dan detailnya. Contoh dari 20 data pada tabel 3.10, sebanyak 10 data mendapatkan hasil “tepat waktu” dan 10 data mendapatkan hasil “terlambat”. Kemudian data dengan atribut “bisnis” detail “kontraktor” diketahui berjumlah 5 data, dengan rincian 3 data mendapatkan hasil “tepat waktu” dan 2 data mendapatkan hasil “terlambat”. Selanjutnya menghitung nilai *entropy* pada tiap detail atribut. Contoh menghitung nilai *entropy* pada atribut “bisnis” detail “kontraktor” sebagai berikut :

$$E = - (\text{Tepat Waktu} / \text{Jumlah} \times \text{IMLOG2}(\text{Tepat Waktu} / \text{Jumlah}) + \text{Terlambat} / \text{Jumlah} \times \text{IMLOG2}(\text{Terlambat} / \text{Jumlah}))$$

$$E = - (90/211 \times \text{IMLOG2}(90/211) + 121/211 \times \text{IMLOG2}(121/211))$$

$$E = 0,98437298$$

Setelah mendapatkan nilai *entropy* pada tiap detail atribut, selanjutnya menghitung nilai *gain* pada tiap atribut. Contoh menghitung nilai *gain* pada atribut “bisnis” sebagai berikut :

$$G = \text{Entropy Total} - ((\text{Jumlah Kontraktor}/\text{Jumlah Total} \times \text{Entropy Kontraktor}) + (\text{Jumlah Manufaktur}/\text{Jumlah Total} \times \text{Entropy Manufaktur}) + (\text{Jumlah Pariwisata}/\text{Jumlah Total} \times \text{Entropy Pariwisata}) + (\text{Jumlah Tambang}/\text{Jumlah Total} \times \text{Entropy Tambang}))$$

$$G = 0,99089810 - ((211/606 \times 0,98437298) + (354/606 \times 0,99167197) + (25/606 \times 0,94268319) + (16/606 \times 0,89603823))$$

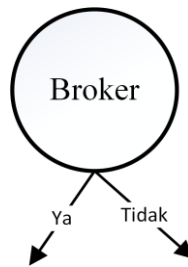
$$G = 0,00631350$$

Berikut adalah tabel hasil perhitungan *entropy* dan *gain node 1* yang terdapat pada tabel 3.11.

Tabel 3.11. Perhitungan *entropy* dan *gain node 1*

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
1	Total		20	10	10	1	
	Bisnis	Kontraktor	5	3	2	0,971	0,029
		Manufaktur	5	2	3	0,971	
		Pariwisata	5	2	3	0,971	
		Tambang	5	3	2	0,971	
	Broker	Ya	5	4	1	0,722	0,091
		Tidak	15	6	9	0,971	
	Cash-back	Ada	9	4	5	0,991	0,007
		Tidak Ada	11	6	5	0,994	
	Jatuh Tempo	Hari Biasa	13	6	7	0,996	0,007
		Hari Libur	7	4	3	0,986	
	Termin	Cash On Delivery	5	3	2	0,971	0,029
		Tempo 7 Hari	5	2	3	0,971	
		Tempo 14 Hari	5	3	2	0,971	
		Tempo 30 Hari	5	2	3	0,971	

Perhitungan *node 1* menghasilkan nilai terbesar *gain* pada atribut broker sebesar 0,091 , maka atribut broker menjadi *node* akar dengan semua detail atributnya dilanjutkan ke *node 2*. Berikut adalah tampilan sementara pohon keputusan pada gambar 3.8.



Gambar 3.8. Pohon keputusan *node 1*

2. *Node 2*

Melakukan langkah-langkah yang sama seperti di *node 1*, mulai dari pengelompokkan data, perhitungan *entropy*, dan perhitungan *gain*. Tetapi pada *node 2*, data yang dihitung didasarkan juga pada *node* induknya. Sehingga *node 2* dibedakan menjadi 2 tabel, masing-masing berdasarkan broker “ya” dan “tidak”. Sedangkan atribut broker tidak dihitung karena sudah menjadi *node* akar.

a. Atribut “broker” detail “ya”

Dari 20 data, diketahui sebanyak 5 data memiliki atribut “broker” detail “ya” dengan rincian 4 data hasil “tepat waktu” dan 1 data hasil “terlambat”. Kemudian dari 5 data, terdapat data dengan atribut “bisnis” detail “manufaktur” sebanyak 3 data dengan rincian 2 data hasil “tepat waktu” dan 1 data hasil “terlambat”. Pengelompokkan data dilakukan di semua atribut dan detailnya. Kemudian dilakukan perhitungan *entropy* dan *gain*.

Contoh proses perhitungan *entropy* atribut “bisnis” detail “manufaktur” :

$$E = - \left(\text{Tepat Waktu} / \text{Jumlah} \times \text{IMLOG2}(\text{Tepat Waktu} / \text{Jumlah}) + \text{Terlambat} / \text{Jumlah} \times \text{IMLOG2}(\text{Terlambat} / \text{Jumlah}) \right)$$

$$E = - \left(2/3 \times \text{IMLOG2}(2/3) + 1/3 \times \text{IMLOG2}(1/3) \right)$$

$$E = 0,9183$$

Contoh proses perhitungan *gain* atribut “bisnis” :

$$G = \text{Entropy Total} - \left((\text{Jumlah Kontraktor} / \text{Jumlah Total} \times \text{Entropy Kontraktor}) + (\text{Jumlah Manufaktur} / \text{Jumlah Total} \times \text{Entropy Manufaktur}) + (\text{Jumlah Pariwisata} / \text{Jumlah Total} \times \text{Entropy Pariwisata}) + (\text{Jumlah Tambang} / \text{Jumlah Total} \times \text{Entropy Tambang}) \right)$$

$$G = 0,7219 - \left((2/5 \times 0) + (3/5 \times 0,9183) + (0/5 \times 0) + (0/5 \times 0) \right)$$

$$G = 0,1710$$

Hasil perhitungan atribut lainnya terdapat pada tabel 3.12.

Tabel 3.12. Perhitungan *node* 2 (broker “ya”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
2	Total		5	4	1	0,7219	
	Bisnis	Kontraktor	2	2	0	0,0000	0,171
		Manufaktur	3	2	1	0,9183	
		Pariwisata	0	0	0	0,0000	
		Tambang	0	0	0	0,0000	
	Cash-back	Ada	5	4	1	0,7219	0
		Tidak Ada	0	0	0	0,0000	
	Jatuh Tempo	Hari Biasa	3	3	0	0,0000	0,322
		Hari Libur	2	1	1	1,0000	
	Termin	Cash On Delivery	2	2	0	0,0000	0,721
		Tempo 7 Hari	1	1	0	0,0000	
		Tempo 14 Hari	1	1	0	0,0000	
		Tempo 30 Hari	1	0	1	0,0000	

Perhitungan *node* 2 atribut “broker” detail “ya” pada tabel 3.12 menghasilkan nilai terbesar *gain* pada atribut termin sebesar 0,721, sehingga atribut termin dipilih sebagai *node* lanjutan dari atribut “broker” detail “ya”. Pada atribut termin terdapat juga beberapa detail dengan *entropy* yang bernilai 0, maka detail tersebut sudah dapat ditentukan *leaf node* sesuai dengan perbandingan jumlah datanya. Contoh detail “*cash on delivery*”, jumlah data hasil tepat waktu lebih besar dari jumlah data hasil terlambat, sehingga *leaf node* nya adalah tepat waktu.

b. Atribut “broker” detail “tidak”

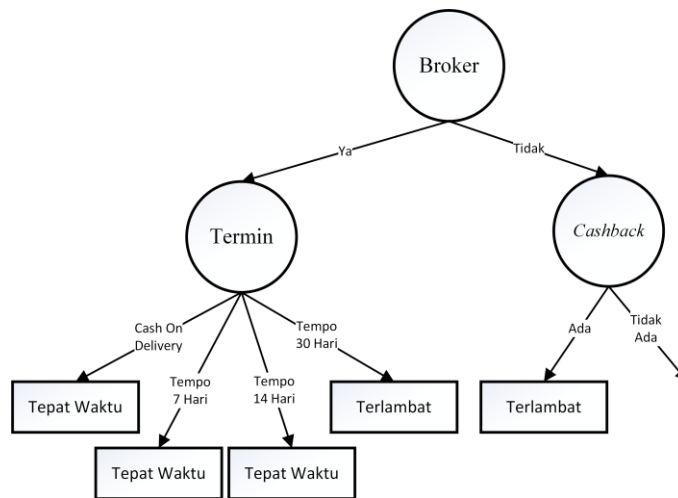
Dilanjutkan dengan menghitung *entropy* dan *gain* pada atribut “broker” detail “tidak”.

Tabel 3.13. Perhitungan *node* 2 (broker “tidak”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
2	Total		15	6	9	0,9710	
	Bisnis	Kontraktor	3	1	2	0,9183	0,140
		Manufaktur	2	0	2	0,0000	
		Pariwisata	5	2	3	0,9710	
		Tambang	5	3	2	0,9710	
	<i>Cash-back</i>	Ada	4	0	4	0,0000	0,242
		Tidak Ada	11	6	5	0,9940	
	Jatuh Tempo	Hari Biasa	10	3	7	0,8813	0,059
		Hari Libur	5	3	2	0,9710	
	Termin	Cash On Delivery	3	1	2	0,9183	0,037
		Tempo 7 Hari	4	1	3	0,8113	
		Tempo 14 Hari	4	2	2	1,0000	
		Tempo 30 Hari	4	2	2	1,0000	

Perhitungan *node* 2 atribut “broker” detail “tidak” pada tabel 3.13 menghasilkan nilai terbesar *gain* pada atribut *cashback* sebesar 0,242. sehingga atribut *cashback* dipilih sebagai *node* lanjutan dari atribut “broker” detail “tidak”

Berikut adalah tampilan sementara pohon keputusan pada gambar 3.9.



Gambar 3.9. Pohon keputusan *node 2*

3. *Node 3*

Node 3 melanjutkan perhitungan atribut *node 2* yang belum mendapatkan *leaf node*. Atribut yang sudah menjadi *node* induk, tidak akan dihitung lagi pada masing-masing perhitungan *entropy* dan *gain*. Cara pengelompokan data sama pada *node* sebelumnya, tetapi pada *node 3* dan selanjutnya pengelompokan dilakukan dengan menggabungkan *node* induk-induknya. Contoh dari 20 data terdapat 11 data yang memiliki gabungan atribut “broker” detail “tidak” dan atribut “cashback” detail “tidak ada” dengan rincian 6 data hasil “tepat waktu” dan 5 data hasil “terlambat”.

a. Atribut “broker” detail “tidak” => atribut “cashback” detail “tidak ada”

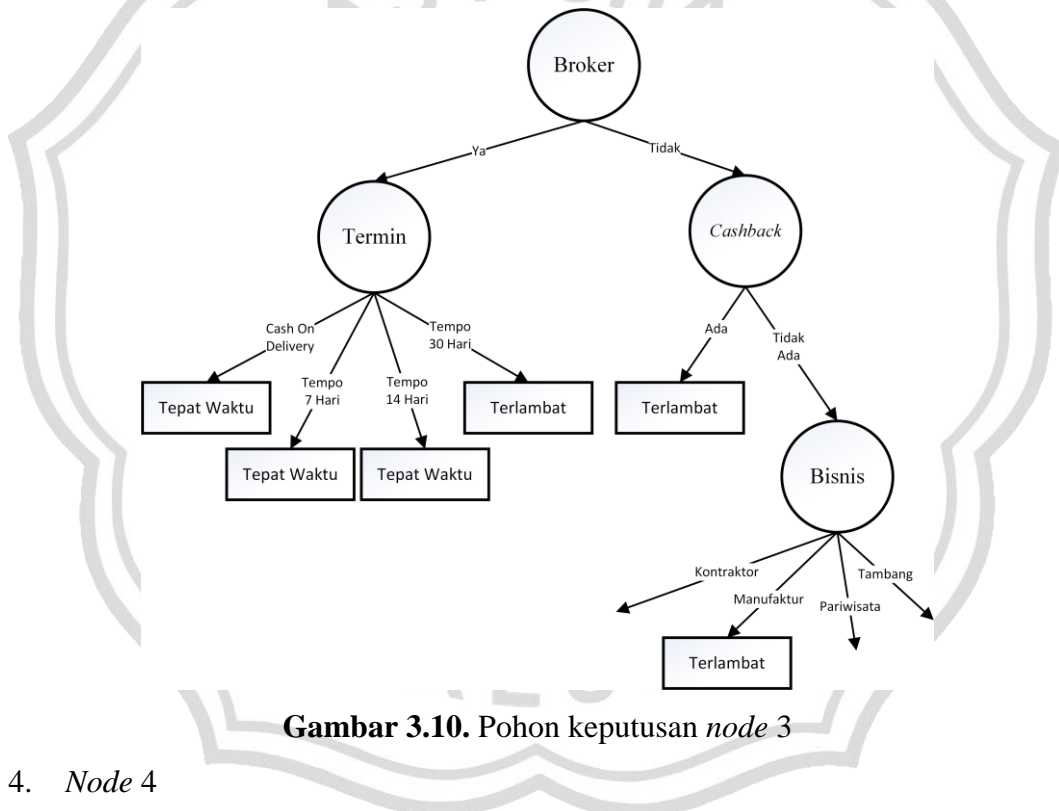
Tabel 3.14. Perhitungan *node 3* (broker “tidak”=>cashback “tidak ada”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
3	Total		11	6	5	0,9940	
	Bisnis	Kontraktor	2	1	1	1,0000	0,153
		Manufaktur	1	0	1	0,0000	
		Pariwisata	4	2	2	1,0000	
		Tambang	4	3	1	0,8113	
	Jatuh Tempo	Hari Biasa	7	3	4	0,9852	0,072
		Hari Libur	4	3	1	0,8113	

	Termin	Cash On Delivery	2	1	1	1,0000	0,016
		Tempo 7 Hari	2	1	1	1,0000	
		Tempo 14 Hari	3	2	1	0,9183	
		Tempo 30 Hari	4	2	2	1,0000	

Perhitungan *node* 3 atribut “broker” detail “tidak” => atribut “cashback” detail “tidak ada” pada tabel 3.14 menghasilkan nilai terbesar *gain* pada atribut bisnis sebesar 0,153.

Berikut adalah pohon keputusan sementara pada gambar 3.10.



Gambar 3.10. Pohon keputusan *node* 3

4. *Node* 4

Node 4 melanjutkan perhitungan atribut *node* 3 yang belum mendapatkan *leaf node*. Atribut yang sudah menjadi *node* induk, tidak akan dihitung lagi pada masing-masing perhitungan *entropy* dan *gain*. Cara pengelompokan data sama dengan pengelompokan data pada *node* 3, yaitu dengan menghitung jumlah data berdasarkan *node* induk-induknya.

- a. Atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “kontraktor”

Tabel 3.15. Perhitungan *node* 4 (broker “tidak”=>*cashback* “tidak ada”=>bisnis “kontraktor”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
4	Total		2	1	1	1,0000	
	Jatuh Tempo	Hari Biasa	2	1	1	1,0000	0
		Hari Libur	0	0	0	0,0000	
	Termin	Cash On Delivery	0	0	0	0,0000	1
		Tempo 7 Hari	1	1	0	0,0000	
		Tempo 14 Hari	0	0	0	0,0000	
		Tempo 30 Hari	1	0	1	0,0000	

Perhitungan *node* 4 atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “kontraktor” pada tabel 3.15 menghasilkan nilai terbesar *gain* pada atribut termin sebesar 1. Pada perhitungan di atas, diketahui bahwa detail atribut “*Cash On Delivery*” dan “Tempo 14 Hari” memiliki jumlah data 0 (nol). Sehingga dapat langsung menghasilkan *leafnode* sesuai dengan jumlah mayoritas data di *node* tersebut. Tetapi karena jumlah data “tepat waktu” dan “terlambat” bernilai sama, maka *leafnode* diputuskan sesuai ketentuan awal, yaitu “terlambat”.

- b. Atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “pariwisata”

Tabel 3.16. Perhitungan *node* 4 (broker “tidak”=>*cashback* “tidak ada”=>bisnis “pariwisata”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
4	Total		4	2	2	1,0000	
	Jatuh Tempo	Hari Biasa	2	1	1	1,0000	0
		Hari Libur	2	1	1	1,0000	
	Termin	Cash On Delivery	1	0	1	0,0000	0,5
		Tempo 7 Hari	0	0	0	0,0000	
		Tempo 14 Hari	1	1	0	0,0000	
		Tempo 30 Hari	2	1	1	1,0000	

Perhitungan *node* 4 atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “pariwisata” pada tabel 3.16 menghasilkan nilai terbesar *gain* pada atribut termin sebesar 0,5.

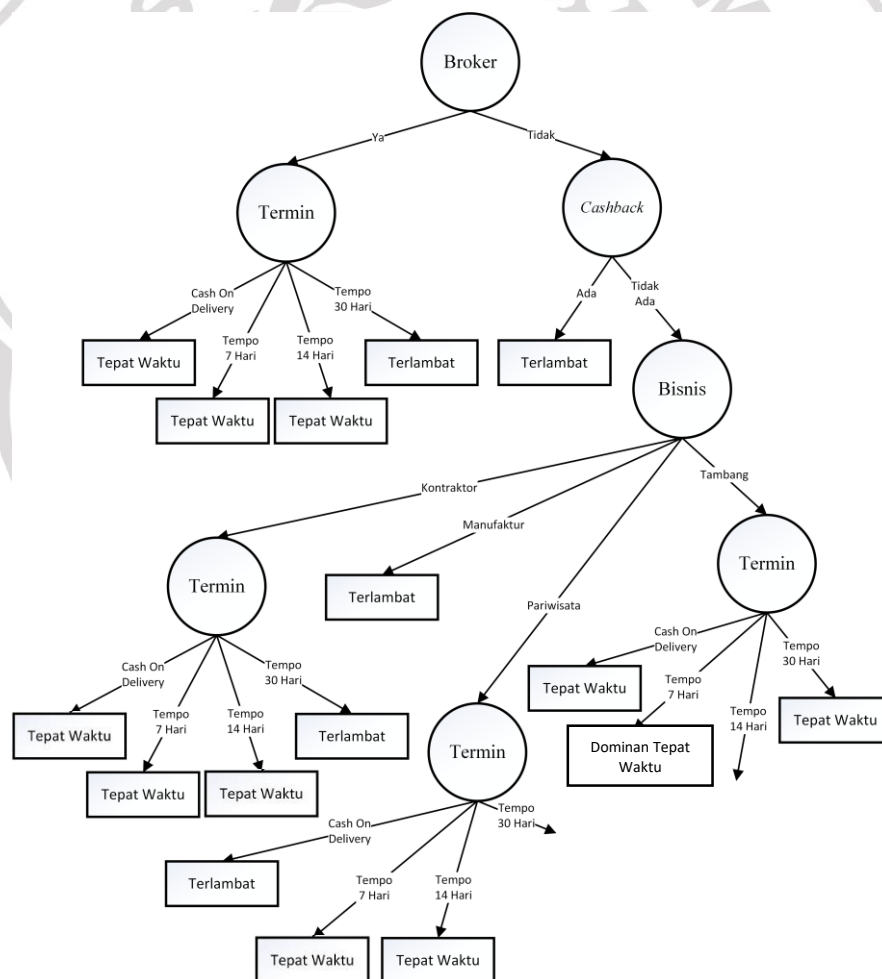
- c. Atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “tambang”

Tabel 3.17. Perhitungan *node* 4 (broker “tidak”=>*cashback* “tidak ada”=>bisnis “tambang”)

No-de	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain
4	Total		4	3	1	0,8113	
	Jatuh Tempo	Hari Biasa	2	1	1	1,0000	0,311
		Hari Libur	2	2	0	0,0000	

	Termin	Cash On Delivery	1	1	0	0,0000	0,311
		Tempo 7 Hari	0	0	0	0,0000	
		Tempo 14 Hari	2	1	1	1,0000	
		Tempo 30 Hari	1	1	0	0,0000	

Perhitungan *node* 4 atribut “broker” detail “tidak” => atribut “*cashback*” detail “tidak ada” => atribut “bisnis” detail “tambang” pada tabel 3.17 menghasilkan nilai terbesar *gain* sebesar 0,311 pada dua atribut. Maka harus memilih salah satu atribut saja. Proses pemilihannya dapat didasarkan pada pertimbangan peneliti, tidak melibatkan metode hitung. Disini penulis memilih atribut termin. Berikut adalah pohon keputusan sementara pada gambar 3.11.

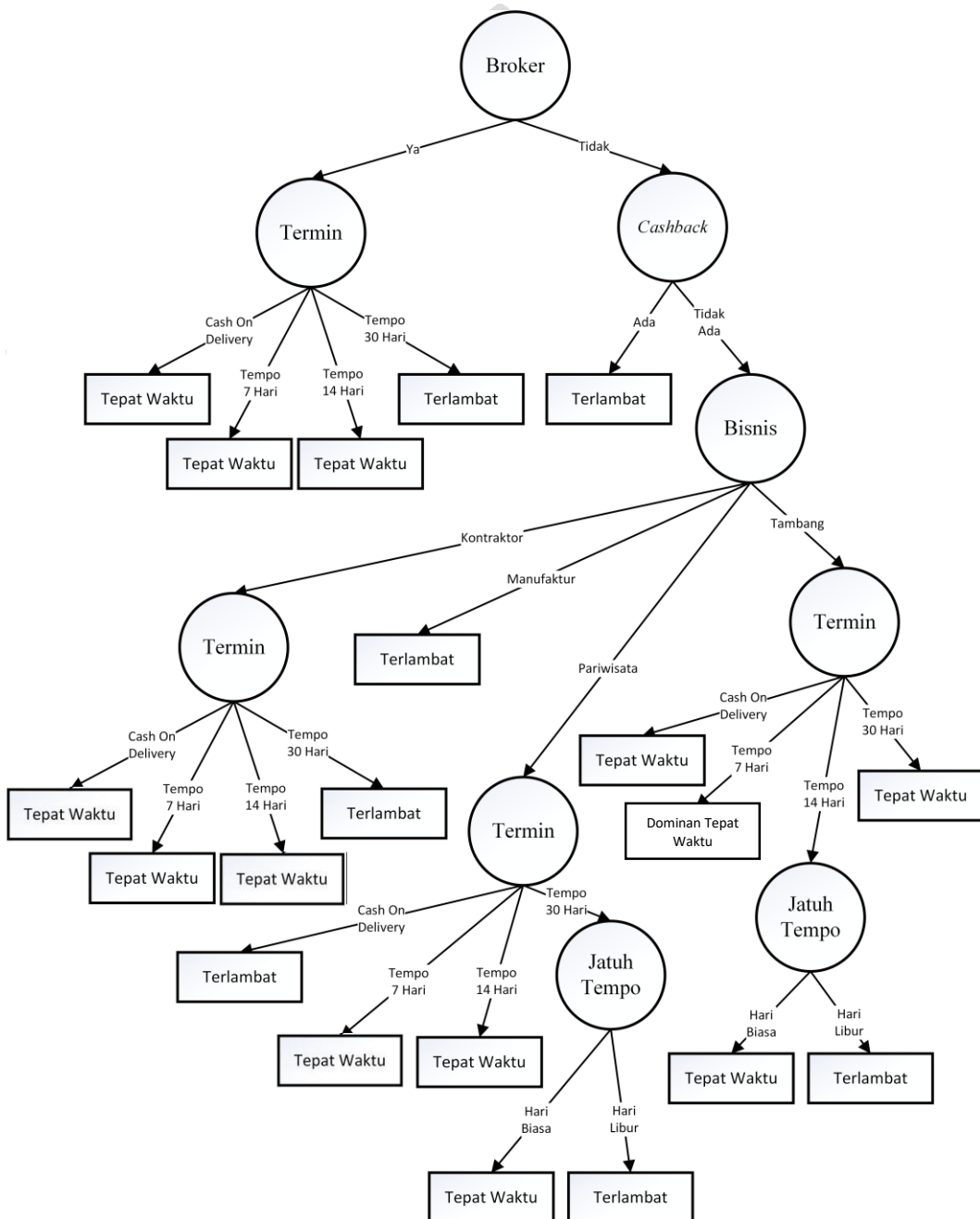


Gambar 3.11. Pohon keputusan *node* 4

	Jatuh Tempo	Hari Biasa	1	0	1	0,0000	1
		Hari Libur	1	1	0	0,0000	

Tabel 3.19 menghasilkan prediksi “terlambat” untuk atribut “jatuh tempo” detail “hari biasa” dan “tepat waktu” untuk atribut “jatuh tempo” detail “hari libur”.

Berikut adalah pohon keputusan akhir pada gambar 3.12.



Gambar 3.12. Pohon keputusan node 5 (akhir)

3.5. Perancangan Antarmuka Sistem

3.5.1. Halaman Atribut

Halaman ini memuat data-data atribut yang ada di sistem, terdapat tombol detail untuk mengarahkan ke halaman detail atribut. Tampilan rancangan halaman atribut terdapat pada gambar 3.13.

MINING-APP	Atribut							
WELCOME USER								
Prediksi Pembayaran Cust.	<table border="1"> <thead> <tr> <th>No</th> <th>Nama Atribut</th> <th>Aksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Bisnis</td> <td><input type="button" value="detail"/></td> </tr> </tbody> </table>		No	Nama Atribut	Aksi	1	Bisnis	<input type="button" value="detail"/>
No	Nama Atribut	Aksi						
1	Bisnis	<input type="button" value="detail"/>						
Atribut Data Latih Data Uji								

Gambar 3.13. Rancangan halaman atribut

3.5.2. Halaman Detail Atribut

Halaman ini memuat data detail yang dimiliki masing-masing atribut di sistem, terdapat juga *form* untuk menambahkan dan menghapus detail atribut. Tampilan rancangan halaman detail atribut terdapat pada gambar 3.14.

MINING-APP	Atribut "Bisnis"							
WELCOME USER	<input type="button" value="Detail atribut baru"/> <input type="button" value="Tambah"/>							
Prediksi Pembayaran Cust.	<table border="1"> <thead> <tr> <th>No</th> <th>Nama Atribut</th> <th>Aksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Kontraktor</td> <td><input type="button" value="hapus"/></td> </tr> </tbody> </table>		No	Nama Atribut	Aksi	1	Kontraktor	<input type="button" value="hapus"/>
No	Nama Atribut	Aksi						
1	Kontraktor	<input type="button" value="hapus"/>						
Atribut Data Latih Data Uji								

Gambar 3.14. Rancangan halaman detail atribut

3.5.3. Halaman Data Latih

Halaman ini memuat data-data latih yang dimasukkan oleh *user*, terdapat tombol tambah untuk mengarahkan ke halaman tambah data latih, tombol ubah untuk mengubah data latih serta tombol kalkulasi untuk mengarahkan ke halaman kalkulasi. Tampilan rancangan halaman data latih terdapat pada gambar 3.15.

MINING-APP	<h1>Data Latih</h1>							
WELCOME USER								
Prediksi Pembayaran Cust.	No Data	Bisnis	Broker	Cashback	Jatuh Tempo	Termin	Prediksi	Aksi
Atribut Data Latih Data Uji	1	Tambang	Ya	Ada	Hari Libur	COD	Tepat Waktu	<input type="button" value="Ubah"/>

Gambar 3.15. Rancangan halaman data latih

3.5.4. Halaman Tambah Data Latih

Halaman ini berisikan *form* untuk menambahkan data latih baru. Tampilan rancangan halaman tambah data latih terdapat pada gambar 3.16.

MINING-APP	<h1>Tambah Data Latih</h1>		
WELCOME USER			
Prediksi Pembayaran Cust.	Bisnis	Broker	Cashback
Atribut Data Latih Data Uji	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>
	Jatuh Tempo	Termin	Hasil
	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>
	<input type="button" value="Simpan"/>		

Gambar 3.16. Rancangan halaman tambah data latih

3.5.5. Halaman Ubah Data Latih

Halaman ini berisikan *form* untuk mengubah data latih. Tampilan rancangan halaman ubah data latih terdapat pada gambar 3.17.

MINING-APP	<h1>Ubah Data No. 106</h1>		
WELCOME USER			
Prediksi Pembayaran Cust.	Bisnis	Broker	Cashback
Atribut Data Latih Data Uji	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>
	Jatuh Tempo	Termin	Hasil
	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>	<input type="text" value="-- Pilih Data --"/>
	<input type="button" value="Simpan"/>		

Gambar 3.17. Rancangan halaman ubah data latih

3.5.6. Halaman Kalkulasi

Halaman ini berisikan data hasil perhitungan *entropy* dan *gain*, terdapat tombol pilih *node* untuk menjadikan sebuah atribut sebagai *node* di pohon keputusan. Tampilan rancangan halaman kalkulasi terdapat pada gambar 3.18.

MINING-APP		Kalkulasi						
WELCOME USER		Atribut	Gain	Pilih				
		Bisnis	0,.....	<input type="button" value="node"/>				
Prediksi Pembayaran Cust.		0,.....					
Atribut Data Latih Data Uji		Detail						
		Node	Atribut	Detail	Jumlah	Tepat Waktu	Terlambat	Entropy
		1	Bisnis	Tambang	211	90	121	0,98437

Gambar 3.18. Rancangan halaman kalkulasi

3.5.7. Halaman Rule Decision Tree

Halaman ini berisikan data *node* pohon keputusan yang saling berurutan dari *node* paling atas. Tampilan rancangan halaman *rule decision tree* terdapat pada gambar 3.19.

MINING-APP		Rule Decision Tree									
WELCOME USER											
Prediksi Pembayaran Cust.		Node	Prediksi								
Atribut Data Latih Data Uji		+ Bisnis									
		<table border="1"> <tr> <td>Kontraktor</td> <td></td> </tr> <tr> <td>+ Termin</td> <td></td> </tr> <tr> <td>Tempo 14 Hari</td> <td>Tepat Waktu</td> </tr> <tr> <td>Tambang</td> <td>Terlambat</td> </tr> </table>	Kontraktor		+ Termin		Tempo 14 Hari	Tepat Waktu	Tambang	Terlambat	
Kontraktor											
+ Termin											
Tempo 14 Hari	Tepat Waktu										
Tambang	Terlambat										

Gambar 3.19. Rancangan halaman *rule decision tree*

3.5.8. Halaman Data Uji

Halaman ini berisikan *form* untuk melakukan pengujian terhadap suatu data. Terdapat tombol uji untuk menampilkan hasil pengujian, dan tombol riwayat uji untuk melihat riwayat pengujian. Tampilan rancangan halaman data uji terdapat pada gambar 3.20.

MINING-APP		Data Uji		Riwayat Uji	
WELCOME USER		Perusahaan	<input type="text"/>	Hasil	
Prediksi Pembayaran Cust.		Bisnis	-- Pilih Data --	PERUSAHAAN : PT. ABCDE	
Atribut Data Latih Data Uji		Broker	-- Pilih Data --	BISNIS : TAMBANG	
		Cashback	-- Pilih Data --	BROKER : TIDAK	
		Termin	-- Pilih Data --	CASHBACK : TIDAK ADA	
		Jatuh Tempo	-- Pilih Data --	TERMIN : TEMPO 30 HARI	
			<input type="button" value="Uji"/>	JATUH TEMPO : HARI LIBUR	
				<input type="button" value="Prediksi TERLAMBAT"/>	

Gambar 3.20. Rancangan halaman data uji

3.5.9. Halaman Riwayat Uji

Halaman ini berisikan data riwayat pengujian yang pernah dilakukan oleh *user*. Tampilan rancangan halaman riwayat uji terdapat pada gambar 3.21.

MINING-APP		Riwayat Uji						Uji
WELCOME USER		Perusahaan	Bisnis	Broker	Cashback	Jatuh Tempo	Termin	Prediksi
Prediksi Pembayaran Cust.		PT. ABCDE	Tambang	Tidak	Tidak Ada	Hari Libur	30 Hari	Terlambat
Atribut Data Latih Data Uji								

Gambar 3.21. Rancangan halaman riwayat uji

3.6. Skenario Pengujian

Sebagai langkah berikutnya pada penelitian ini, penulis akan mengimplementasikan perhitungan metode *decision tree* di sistem aplikasi yang dibuat. Aplikasi akan menghitung keseluruhan data latih sebanyak 420 data dengan metode *decision tree*, kemudian aplikasi menghasilkan pohon keputusan. Selanjutnya, data uji sebanyak 186 data akan diproses dengan alur pohon keputusan yang dihasilkan aplikasi. Sehingga dapat diketahui apakah hasil keputusan prediksinya berbeda atau tidak antara hasil aslinya dengan hasil dari pohon keputusan. Dari perbandingan tersebut, dapat dihitung nilai akurasi dari pengujian metode *decision tree*. Berikut adalah rumus perhitungan nilai akurasinya :

$$\text{Akurasi} = \frac{\text{perbandingan bernilai "sama"}}{\text{jumlah seluruh data uji}} \times 100\% \quad \dots\dots(3.1)$$

Rumus 3.1 memerlukan jumlah perbandingan antara hasil dari data uji asli dengan hasil dari pohon keputusan yang bernilai sama. Kemudian dibagi dengan jumlah seluruh data uji yaitu 186 data. Hasilnya akan ditampilkan dalam bentuk persentase.

Berikut merupakan contoh data uji :

Tabel 3.20. Contoh data uji

No	Bisnis	Broker	Cashback	Jatuh Tempo	Termin	Hasil
1	Kontraktor	Ya	Ada	Hari Libur	Tempo 7 Hari	Tepat Waktu
2	Manufaktur	Tidak	Ada	Hari Biasa	Cash On Delivery	Tepat Waktu
3	Pariwisata	Tidak	Tidak Ada	Hari Libur	Tempo 30 Hari	Terlambat
4	Tambang	Tidak	Tidak Ada	Hari Biasa	Tempo 14 Hari	Terlambat
5	Manufaktur	Ya	Ada	Hari Biasa	Tempo 14 Hari	Terlambat

Masing-masing dari data uji pada tabel 3.20 akan dilakukan penelusuran hasil prediksinya menggunakan pohon keputusan yang sudah dihasilkan pada gambar 3.12. Contoh, pohon keputusan memiliki *node* akar atribut broker. Data uji nomor 1 memiliki atribut “broker” dengan detail “ya”, maka penelusuran dilanjutkan ke pohon keputusan *node* 2 atribut termin. Data uji nomor 1 memiliki atribut “termin” detail “tempo 7 hari”, pada penelusuran pohon keputusan dinyatakan mendapatkan hasil prediksi “tepat waktu”. Cara penelusuran tersebut dilakukan pada tiap data uji.

Berikut adalah hasil prediksi dari data uji lainnya :

Tabel 3.21. Hasil pengujian data uji

No	Bisnis	Bro-ker	Cash-back	Jatuh Tempo	Termin	Hasil	Pre-diksi	Sama ?
1	Kon-traktor	Ya	Ada	Hari Libur	Tempo 7 Hari	Tepat Waktu	Tepat Waktu	Sama
2	Manu-faktur	Tidak	Ada	Hari Biasa	Cash On Delivery	Tepat Waktu	Terlam-bat	Beda
3	Pariwi-sata	Tidak	Tidak Ada	Hari Libur	Tempo 30 Hari	Terlam-bat	Terlam-bat	Sama
4	Tam-bang	Tidak	Tidak Ada	Hari Biasa	Tempo 14 Hari	Terlam-bat	Tepat Waktu	Beda
5	Manu-faktur	Ya	Ada	Hari Biasa	Tempo 14 Hari	Terlam-bat	Tepat Waktu	Beda

Pada tabel 3.21 kolom “prediksi” adalah kolom hasil prediksi data uji yang telah ditelusuri menggunakan pohon keputusan di gambar 3.12, kemudian kolom “sama ?” adalah hasil perbandingan antara kolom “hasil” dan kolom “prediksi”. Selanjutnya dapat dihitung nilai akurasi dari pengujian di atas sebagai berikut :

$$\text{Akurasi} = \frac{\text{perbandingan bernilai "sama"}}{\text{jumlah seluruh data uji}} \times 100\%$$

$$\text{Akurasi} = \frac{2}{5} \times 100\%$$

$$\text{Akurasi} = 40\%$$

Didapatkan hasil nilai akurasi dari pengujian data uji sebesar 40%.