



**SISTEM INFORMASI REKOMENDASI PERUBAHAN STATUS
PELANGGAN MENGGUNAKAN ALGORITMA C4.5
(Studi Kasus: PT. Herona Express Jember)**

SKRIPSI

oleh :

Adinda Trisna Alfiraningrum

112410101027

**PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS JEMBER**

2015



**SISTEM INFORMASI REKOMENDASI PERUBAHAN STATUS
PELANGGAN MENGGUNAKAN ALGORITMA C4.5
(Studi Kasus: PT. Herona Express Jember)**

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat untuk menyelesaikan Pendidikan Sarjana (S1) Program Studi Sistem Informasi dan mencapai gelar Sarjana Komputer

oleh :

Adinda Trisna Alfiraningrum

112410101027

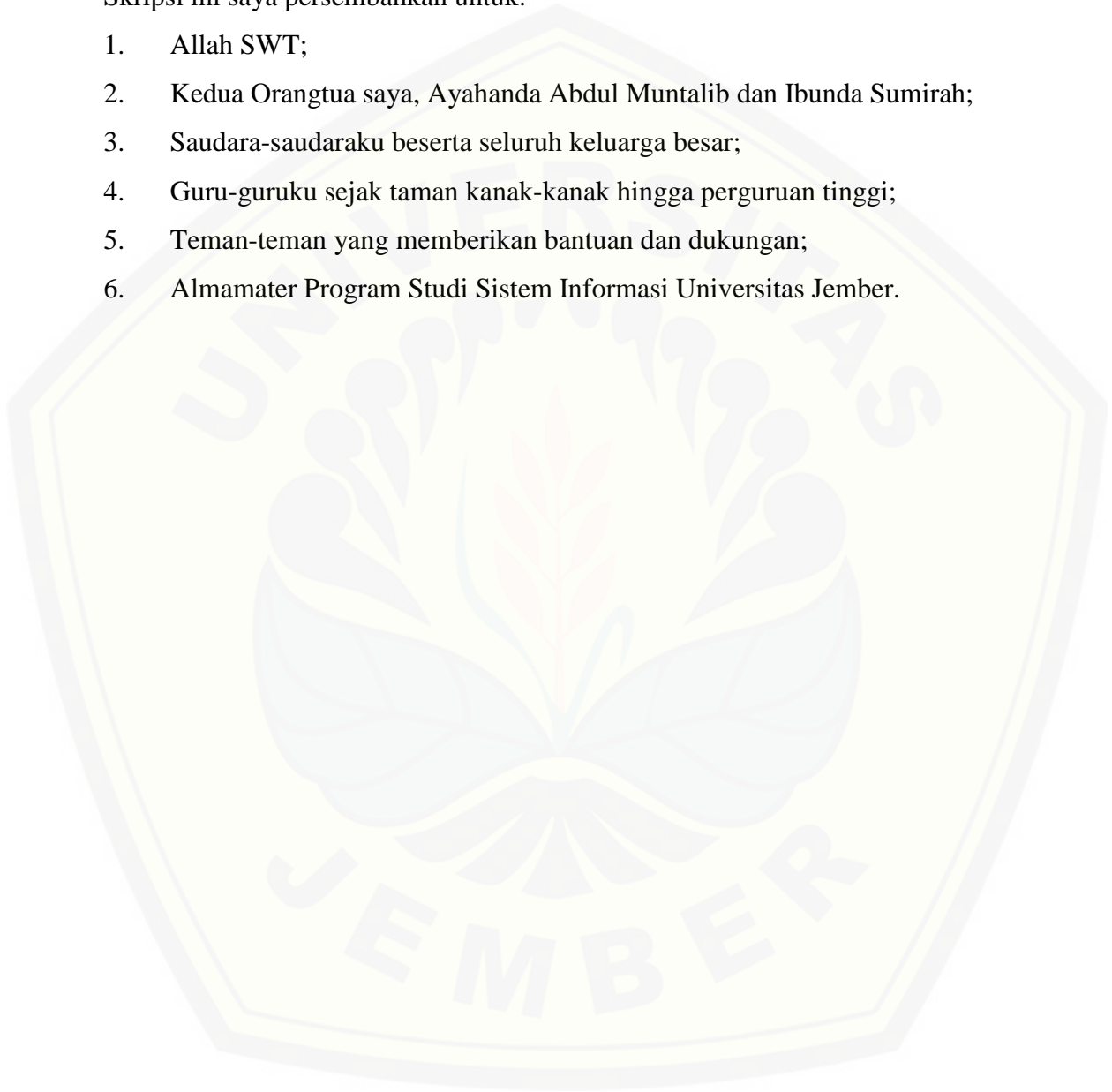
**PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS JEMBER**

2015

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

1. Allah SWT;
2. Kedua Orangtua saya, Ayahanda Abdul Muntalib dan Ibunda Sumirah;
3. Saudara-saudaraku beserta seluruh keluarga besar;
4. Guru-guruku sejak taman kanak-kanak hingga perguruan tinggi;
5. Teman-teman yang memberikan bantuan dan dukungan;
6. Almamater Program Studi Sistem Informasi Universitas Jember.



MOTO

إِنَّمَا أَمْرُهُ إِذَا أَرَادَ شَيْئًا أَنْ يَقُولَ لَهُ كُنْ فَيَكُونُ

“Sesungguhnya keadaan kekuasaannya apabila Ia menghendaki sesuatu hanyalah berkata kepadanya: ‘Jadilah!’, maka terjadilah ia.”

(QS. Yaasiin: 82)

“Man sara ‘ala ad-darbi washala”

“Barang siapa berjalan pada jalannya, dia akan sampai”

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Adinda Trisna Alfiraningrum

NIM : 112410101027

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Sistem Informasi Rekomendasi Perubahan Status Pelanggan Menggunakan Algoritma C4.5”, adalah benar-benar hasil karya sendiri, kecuali jika dalam pengutipan substansi disebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika di kemudian hari pernyataan ini tidak benar.

Jember, 28 Desember 2015

Yang menyatakan,

Adinda Trisna Alfiraningrum

PENGESAHAN PEMBIMBING

Skripsi berjudul “Sistem Informasi Rekomendasi Perubahan Status Pelanggan Menggunakan Algoritma C4.5”, telah diuji dan disahkan pada :

hari, tanggal : 28 Desember 2015

tempat : Program Studi Sistem Informasi Universitas Jember

Disetujui oleh :

Pembimbing I,

Pembimbing II,

Prof. Drs. Slamim, M.Comp.Sc.,Ph.D

Windi Eka Yulia Retnani S.Kom.,MT.

NIP 196704201992011001

NIP 198403052010122002

SKRIPSI

**SISTEM INFORMASI REKOMENDASI PERUBAHAN STATUS
PELANGGAN MENGGUNAKAN ALGORITMA C4.5
(Studi Kasus: PT. Herona Express Jember)**

oleh :

Adinda Trisna Alfiraningrum

112410101027

Pembimbing :

Dosen Pembimbing Utama : Prof. Drs. Slamir, M.Comp.Sc.,Ph.D

Dosen Pembimbing Pendamping : Winda Eka Yulia Retnani S.Kom.,MT.

PENGESAHAN

Skripsi berjudul “Sistem Informasi Rekomendasi Perubahan Status Pelanggan Menggunakan Algoritma C4.5”, telah diuji dan disahkan pada :

hari, tanggal : 28 Desember 2015

tempat : Program Studi Sistem Informasi Universitas Jember

Penguji I,

Penguji II,

Dr. Saiful Bukhori, ST., M.Kom

NIP 196811131994121001

Muhamad Arief Hidayat S.Kom., M.Kom.

NIP 198101232010121003

Mengesahkan

Ketua Program Studi,

Prof. Drs. Slamin, M.Comp.Sc.,Ph.D

NIP 196704201992011001

RINGKASAN

Persaingan antar kompetitor yang bergerak di bidang pengiriman barang melalui kereta api secara tidak langsung menuntut para pesaing bisnis dan penyedia jasa untuk selalu menciptakan inovasi baru agar menjadi lebih unggul. Pelanggan merupakan salah satu aset perusahaan yang harus dikelola dengan baik agar tidak pindah ke kompetitor lain. Salah satu cara mempertahankan pelanggan adalah dengan memberikan pelayanan khusus pada tingkatan pelanggan tertentu. Berdasarkan permasalahan tersebut, maka dibutuhkan sebuah sistem informasi yang dapat merekomendasikan perubahan status pelanggan dan saran berdasarkan *track record* transaksi pelanggan secara akurat sehingga dapat membantu perusahaan untuk memberikan pelayanan dengan lebih tepat dan efisien.

Prediksi rekomendasi perubahan status pelanggan yang diterapkan dalam sistem ini menggunakan analisis RFM (*Recency, Frequency* dan *Monetary*) dan *Data Mining Classification* Algoritma C4.5. Model prediksi tersebut akan menentukan kelas suatu pelanggan dan memberikan prediksi. Perhitungan prediksi tersebut menggunakan Algoritma C4.5 yang menghasilkan *rule* untuk selanjutnya digunakan sebagai dasar prediksi. Hasil dari penelitian ini adalah prediksi rekomendasi suatu pelanggan berdasarkan pola *track record* transaksi pelanggan sebelumnya.

PRAKATA

Puji syukur kehadiran Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Sistem Informasi Rekomendasi Perubahan Status Pelanggan”. Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan Strata Satu (S1) pada Program Studi Sistem Informasi Universitas Jember.

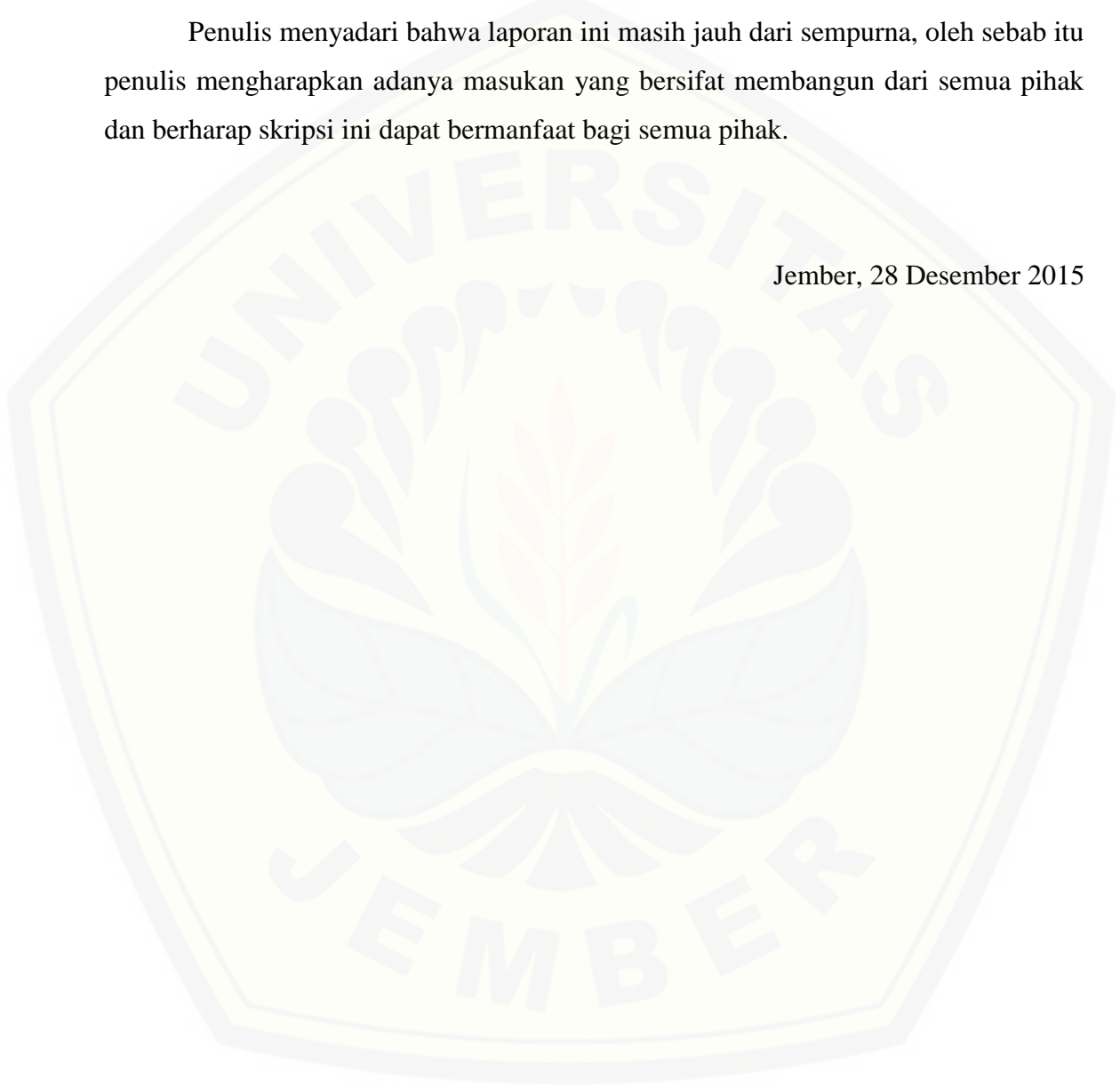
Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Prof. Drs. Slamini, M.Comp.Sc.,Ph.D., selaku Ketua Program Studi Sistem Informasi Universitas Jember;
2. Prof. Drs. Slamini, M.Comp.Sc.,Ph.D., selaku Dosen Pembimbing Utama dan Winda Eka Yulia Retnani S.Kom.,MT. selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan skripsi;
3. Anang Andrianto S.T.,MT., selaku Dosen Pembimbing Akademik yang telah membimbing selama penulis menjadi mahasiswa;
4. Seluruh Bapak dan Ibu dosen beserta staf karyawan di Program Studi Sistem Informasi Universitas Jember;
5. Ibunda tersayang Sumirah dan Ayahanda Abdul Muntalib yang telah memberikan banyak hal;
6. Adik Alfinda Qabliyah Fitri dan Almido Cahaya Perdaya;
7. Teman istimewa Mita, Kadek, Leli, Devi, Emas, Jefta, Maulana, Mas Angga, Mas Hadi, Mas Iyan, Mas Anggi, Mbak Vivi, Mas Fauzan, Mas Musa, Mas Rony, Adnan, Cece, Uni, Tari, Niak, Ipind, Anwar, Yopi, Nafta, Rozi, Sendi, Siska dan Aulia;
8. Semua mahasiswa Program Studi Sistem Informasi;
9. Keluarga kecil KKN 28 dan Ibu Kades Sidorejo sekeluarga;
10. Keluarga kecil Ulmiyah Jember;

11. Keluarga kecil Malang;
12. PT. Herona Express Jember;
13. Semua pihak yang tidak dapat disebutkan satu-persatu;

Penulis menyadari bahwa laporan ini masih jauh dari sempurna, oleh sebab itu penulis mengharapkan adanya masukan yang bersifat membangun dari semua pihak dan berharap skripsi ini dapat bermanfaat bagi semua pihak.

Jember, 28 Desember 2015



DAFTAR ISI

PERSEMBAHAN.....	ii
MOTO.....	iii
PRAKATA.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xvi
BAB 1. PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.3 Tujuan.....	4
1.4 Batasan Masalah.....	4
1.5 Sistematika Penulisan.....	4
BAB 2. TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu.....	6
2.2 Sistem Informasi Rekomendasi.....	7
2.3 Analisis RFM.....	7
2.4 <i>Data Mining Classification</i>	8
2.5 Algoritma C4.5.....	9
2.6 Metode pengujian Algoritma <i>Classification</i>	12
2.7 Model Waterfall.....	13
2.7.1 Analisa Kebutuhan.....	14
2.7.2 Desain.....	14
2.7.3 Implementasi.....	15
2.7.4 Pengujian.....	15
2.7.5 Pemeliharaan.....	15
BAB 3. METODOLOGI PENELITIAN.....	16
3.1 Pendekatan Penelitian.....	16

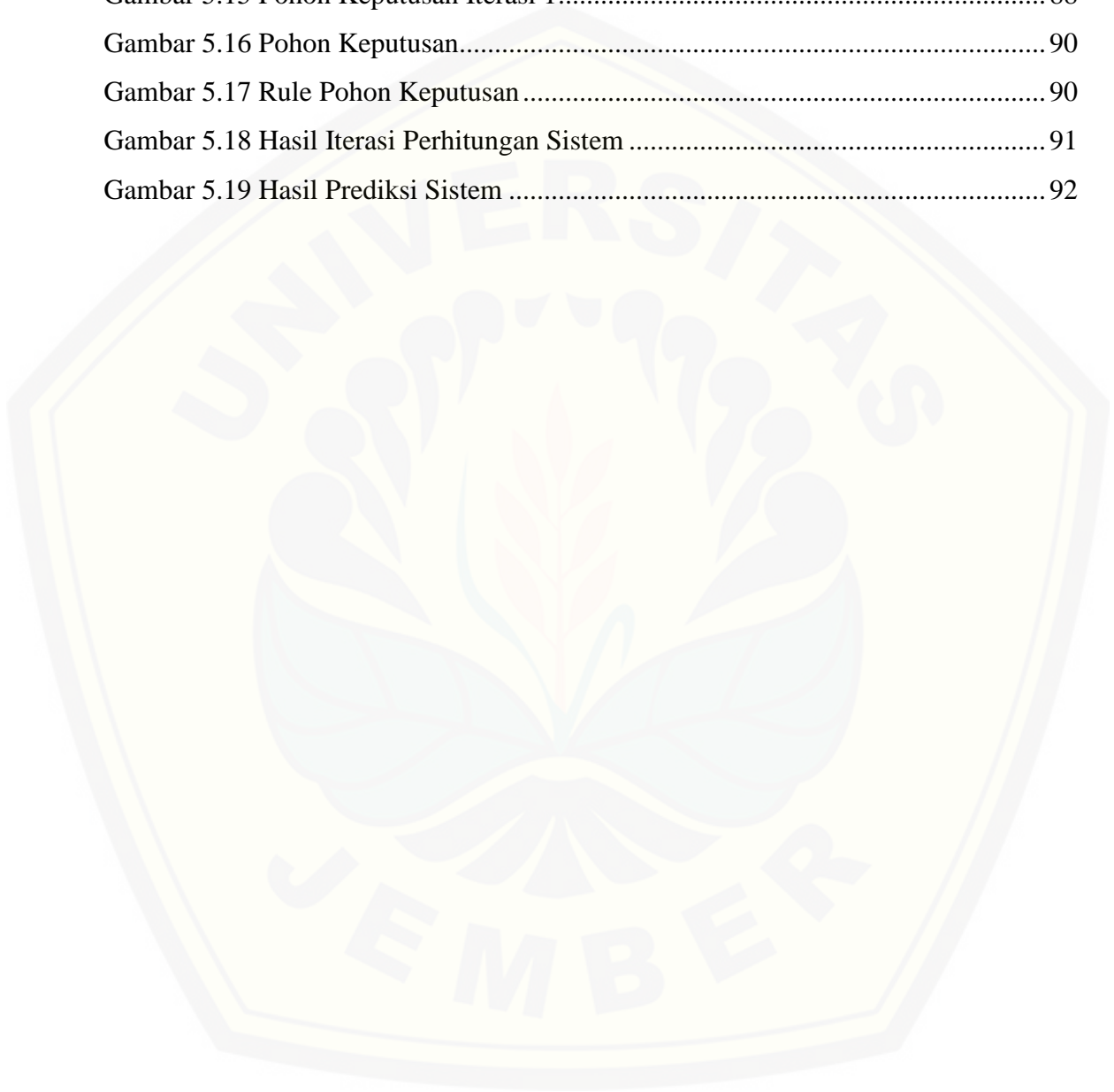
3.2.	Waktu dan Tempat.....	16
3.3.	Tahap Penelitian.....	16
3.3.1	Studi Literatur.....	17
3.3.2	Pengumpulan Data.....	18
3.3.3	Penerapan Algoritma C4.5 untuk Rekomendasi Perubahan Status Pelanggan.....	19
3.3.4	Pengembangan Sistem.....	23
BAB 4. ANALISIS DAN PENGEMBANGAN SISTEM.....		24
4.1	Pengumpulan Data.....	24
4.2	Algoritma C4.5 dan Analisis RFM untuk Prediksi Perubahan Status Pelanggan.....	24
4.2.1	Request Data.....	25
4.2.2	Proses Learning.....	25
4.2.3	Proses Pengujian.....	25
4.3	Pengembangan Sistem.....	26
4.3.1	Kebutuhan Fungsional.....	26
4.3.2	Kebutuhan Non-fungsional.....	26
4.4	Desain Sistem.....	27
4.4.1	Business Process.....	27
4.4.2	Usecase Diagram.....	28
4.4.3	Skenario.....	31
4.4.4	Activity Diagram.....	37
4.4.5	Sequence Diagram.....	40
4.4.6	Class Diagram.....	44
4.4.7	Entity Relationship Diagram.....	47
4.5	Penulisan Kode Program.....	48
4.6	Pengujian Kode Program.....	48
4.6.1	White Box Testing.....	48
4.4.2	<i>Black Box Testing</i>	63

BAB 5. HASIL DAN PEMBAHASAN.....	72
5.1 Hasil Implementasi Sistem Informasi Rekomendasi Perubahan Status Pelanggan.....	72
5.1.1 Halaman <i>Login</i>	72
5.1.2 Halaman <i>Dashboard</i>	73
5.1.3 Halaman Daftar Member.....	73
5.1.4 Halaman Daftar Persyaratan.....	74
5.1.5 Halaman Daftar Transaksi.....	74
5.1.6 Halaman Daftar Kota.....	75
5.1.7 Halaman Daftar Barang.....	76
5.1.8 Halaman Daftar Transaksi.....	76
5.1.9 Halaman <i>Data Set</i>	77
5.1.10 Halaman proses prediksi.....	78
5.2 Pembahasan.....	79
5.2.1 Hasil Implementasi Algoritma C4.5 dan analisis RFM pada sistem....	79
5.3 Pengujian Sistem.....	91
BAB 6. PENUTUP.....	94
6.1 Kesimpulan.....	94
6.2 Saran.....	95
DAFTAR PUSTAKA.....	96
LAMPIRAN A.....	98
LAMPIRAN B.....	136
LAMPIRAN C.....	146

DAFTAR GAMBAR

Gambar 2.1 Diagram Alir Waterfall	14
Gambar 3.1 Alur Penelitian.....	17
Gambar 3.2 <i>Flowchart</i> Proses <i>Learning</i>	19
Gambar 3.3 <i>Flowchart</i> pembangunan <i>tree</i>	20
Gambar 3.4 <i>Flowchart</i> pembuatan node	21
Gambar 3.5 <i>Flowchart</i> implementasi Algoritma C4.5 pada segmentasi pelanggan ...	22
Gambar 4.1 <i>Busines Process</i> Sistem.....	27
Gambar 4.2 <i>Usecase Sistem</i>	28
Gambar 4.3 Activity Diagram Fitur Prediksi	37
Gambar 4.4 Sequence Diagram Fitur Prediksi.....	41
Gambar 4.5 <i>Class Diagram</i>	46
Gambar 4.6 <i>Entity Relationship Diagram</i> (ERD).....	47
Gambar 4.7 <i>Listing Program</i> Fitur Prediksi	50
Gambar 4.8 Diagram Alir Fitur Prediksi Pada <i>csa_prediksi</i>	51
Gambar 4.9 Diagram Alir Fitur Prediksi Pada <i>m_prediksi</i>	51
Gambar 5.1 Halaman Login Sistem	72
Gambar 5.2 Halaman Dashboard	73
Gambar 5.3 Halaman Daftar Member.....	73
Gambar 5.4 Halaman Daftar Persyaratan.....	74
Gambar 5.5 Halaman Daftar Transaksi.....	75
Gambar 5.6 Halaman Daftar Kota	75
Gambar 5.7 Halaman Daftar Barang.....	76
Gambar 5.8 Halaman Daftar Transaksi oleh Superadmin	77
Gambar 5.9 Halaman Data Set.....	78
Gambar 5.10 Halaman Proses Prediksi	78
Gambar 5.11 Halaman Tambah Data Transaksi	80
Gambar 5.12 Halaman Form Prediksi Pelanggan Baru	80

Gambar 5.13 <i>Code Program Data Set</i>	81
Gambar 5.14 Code Program Perhitungan Prediksi Perubahan Status Pelanggan.....	84
Gambar 5.15 Pohon Keputusan Iterasi 1.....	88
Gambar 5.16 Pohon Keputusan.....	90
Gambar 5.17 Rule Pohon Keputusan.....	90
Gambar 5.18 Hasil Iterasi Perhitungan Sistem.....	91
Gambar 5.19 Hasil Prediksi Sistem.....	92



DAFTAR TABEL

Tabel 2.1 Model <i>Confusion Matrix</i>	12
Tabel 3.1 Atribut dan Nilai Atribut.....	18
Tabel 4.1 Definisi Aktor	29
Tabel 4.2 Definisi Usecase Sistem.....	29
Tabel 4.3 Skenario Prediksi Perubahan Status.....	32
Tabel 4.4 Test Case Fitur prediksi	56
Tabel 4.5 Pengujian Black Box.....	63
Tabel 5.1 Contoh Data <i>Set</i>	85
Tabel 5.2 Contoh Probabilitas.....	85
Tabel 5.3 Contoh Kelas Atribut <i>Recency</i>	86
Tabel 5.4 Contoh Batas Atribut <i>Numerik</i>	87
Tabel 5.5 Perhitungan Entropy dan Information Gain Nilai Batas Atribut Numerik .	87
Tabel 5.6 Hasil Perhitungan Semua Atribut	88
Tabel 5.7 Perhitungan Semua Atribut pada Iterasi 2	89
Tabel 5.8 Perhitungan Semua Atribut pada Iterasi 3	89
Tabel 5.9 <i>Confusion Matrix</i> Hasil Klasifikasi Perubahan Status Pelanggan	92

BAB 1. PENDAHULUAN

Bab ini merupakan langkah awal dari penulisan tugas akhir ini. Bab ini menjelaskan latar belakang, perumusan masalah, tujuan, batasan masalah, dan sistematika penulisan.

1.1 Latar Belakang

Teknologi informasi semakin maju seiring dengan perkembangan kehidupan manusia. Kemajuan teknologi yang cepat memberikan pengaruh cukup besar pada segala bidang, termasuk salah satunya dalam bidang bisnis industri maupun jasa. Persaingan antar kompetitor secara tidak langsung menuntut para pesaing bisnis dan penyedia jasa untuk selalu menciptakan inovasi baru agar menjadi lebih unggul. Target riil dari sebuah usaha adalah pelanggan atau pengguna barang dan jasa. Pelanggan merupakan salah satu aset penting perusahaan industri maupun penyedia jasa. Salah satu strategi pemasaran sebuah perusahaan industri maupun penyedia jasa yang paling umum digunakan adalah mempertahankan pelanggan. Pengelompok-an pelanggan akan mempermudah strategi pemasaran dari sebuah perusahaan. Strategi tersebut lebih menguntungkan daripada mencari pelanggan baru dinilai dari segi kepercayaan, hingga biaya yang dikeluarkan.

Pengiriman barang melalui kereta api merupakan salah satu inovasi baru pada masanya, yang juga banyak diminati oleh masyarakat kini. Banyaknya kompetitor yang bergerak dibidang yang sama membuat para penyedia jasa tersebut berlomba untuk menawarkan pelayanan yang paling baik. Salah satu contoh bisnis yang membutuhkan strategi tersebut adalah ekspedisi pengiriman barang menggunakan jasa kereta api PT. Herona Express Jember. PT. Herona Express merupakan perusahaan yang bergerak di bidang jasa titipan untuk pengiriman dokumen dan atau paket ukuran kecil hingga dalam jumlah besar (cargo), yang telah hadir sejak tahun 1998 (Express, 2013). PT. Herona Express adalah salah satu contoh Ekspedisi Muatan Kereta Api (EMKA) yang dapat mengirim ke lebih dari 50 kota di seluruh Jawa dan Bali. PT. Herona Express

Jember melayani pengiriman barang dengan tidak ada batasan minimal dengan tarif yang berbeda-beda tiap barang/*item*.

Suatu perusahaan akan berpikir bagaimana caranya mendapatkan pelanggan, dan menjadikan pelanggan sebagai aset tetap perusahaan. Salah satu cara yang dapat dilakukan adalah dengan memberikan penawaran baru setiap periode tertentu. Penawaran-penawaran tersebut membuat pelanggan tertarik dan ingin mencoba. Pelanggan yang telah mencoba kemudian merasa nyaman dan percaya pada suatu penyedia jasa, akan cenderung berlangganan di tempat tersebut. Berbeda dengan pelanggan yang masih mencari perbandingan dan mencoba-coba untuk menggunakan suatu penyedia jasa, akan cenderung pindah ke kompetitor lain guna mencari yang paling baik.

Pelayanan pada pelanggan yang memberikan keuntungan lebih banyak pada perusahaan tentu tidak ingin diperlakukan sama dengan pelanggan biasa. Pelayanan yang dimaksud adalah dari segi penawaran produk, penawaran harga, dan atau sejenisnya. Selama ini perusahaan hanya memberikan pelayanan yang berbeda pada golongan tertentu berdasarkan hubungan baik secara personal, dan bukan pada tingkat yang seharusnya.

Segmentasi pelanggan akan mempermudah perusahaan untuk meninjau pelanggannya. Hal ini sangat membantu mengingat banyaknya pengguna jasa pengiriman sangat banyak. Segmentasi pelanggan memberikan celah untuk lebih unggul dalam pelayanan. Status pelanggan yang baru dapat dilihat dari *record* transaksi atau perilaku pelanggan, yang tentunya akan menjadi bahan pertimbangan untuk memilah dan memilih mana pelanggan tetap dan bukan. Perubahan status tersebut digunakan sebagai tolok ukur dalam pemberian pelayanan. Pelayanan yang diberikan pada semua pelanggan tentunya adalah yang terbaik, namun untuk beberapa hal terdapat sedikit perbedaan. Pelanggan yang berpotensi lebih banyak memberikan keuntungan bagi perusahaan tentu diberikan pelayanan satu tingkat lebih baik dari pada pelanggan biasa.

Berdasarkan permasalahan tersebut, maka dibutuhkan sebuah sistem rekomendasi perubahan status pelanggan sehingga dapat mengetahui status pelanggan dengan lebih akurat dan dapat memberikan pelayanan yang tepat sasaran. Rekomendasi perubahan status pelanggan yang diterapkan pada sistem ini menggunakan Data Mining Algoritma C4.5 dengan analisis RFM. Algoritma C4.5 merupakan kelompok algoritma *decision tree* yang dapat digunakan untuk memprediksi suatu kejadian. Hasil klasifikasi menentukan kategori pelanggan yang baru berdasarkan *record* transaksi dengan dibantu analisis *recency*, *frequency*, dan *monetary* (RFM). Analisis RFM merupakan model analisis untuk menentukan segmentasi pelanggan berdasarkan waktu terakhir transaksi, seberapa sering transaksi dilakukan oleh pelanggan, dan jumlah dari transaksi tersebut.

1.2. Perumusan Masalah

Berdasarkan latar belakang diatas maka dirumuskan permasalahan sebagai berikut :

1. Bagaimana menerapkan Algoritma C4.5 dan analisis RFM pada sistem informasi segmentasi pelanggan pengguna jasa pengiriman barang?
2. Bagaimana merancang sistem untuk segmentasi pelanggan pengguna jasa pengiriman barang lebih akurat dengan memprediksi member pelanggan berdasarkan *record* transaksi dan mengoptimalkan pelayanan dengan menganalisis perilaku pelanggan?
3. Bagaimana membangun sistem untuk segmentasi pelanggan pengguna jasa pengiriman barang lebih akurat dengan memprediksi member pelanggan berdasarkan *record* transaksi dan mengoptimalkan pelayanan dengan menganalisis perilaku pelanggan?

1.3. Tujuan

Tujuan dari penelitian adalah :

1. Menerapkan Algoritma C4.5 dan analisis RFM dalam membuat rancang bangun sistem informasi perubahan status pelanggan pengguna jasa ekspedisi pengiriman barang.
2. Merancang sistem untuk mengoptimalkan pelayanan dengan memprediksi member pelanggan berdasarkan *record* transaksi dan menerapkan strategi pemasaran yang sesuai dengan prediksi perubahan status pelanggan.
3. Membangun sistem untuk mengoptimalkan pelayanan dengan memprediksi member pelanggan berdasarkan *record* transaksi dan menerapkan strategi pemasaran yang sesuai dengan prediksi perubahan status pelanggan.

1.4. Batasan Masalah

Batasan masalah dalam penelitian ini adalah :

1. Sistem ini digunakan untuk mengoptimalkan pelayanan dengan merekomendasikan status pelanggan yang baru berdasarkan *record* transaksi / perilaku pelanggan.
2. Sistem ini dikembangkan menggunakan salah satu teknik data *mining classification decision tree* yaitu Algoritma C4.5.
3. Objek yang dianalisis adalah para pengguna jasa pengiriman barang menggunakan kereta api PT. Herona Express Jember.

1.5. Sistematika Penulisan

Sistematika penulisan dalam penyusunan tugas akhir ini adalah sebagai berikut:

1. Pendahuluan

Bab ini terdiri atas latar belakang, perumusan masalah, tujuan, batasan masalah dan sistematika penulisan.

2. Tinjauan Pustaka

Bab ini berisi materi, informasi, dan kajian teori yang digunakan dalam penulisan.

3. Metodologi Penelitian

Bab ini menguraikan tentang metode apa yang dilakukan selama penelitian. Dimulai dari tahap pencarian permasalahan hingga pengujian sistem.

4. Analisis dan Pengembangan Sistem

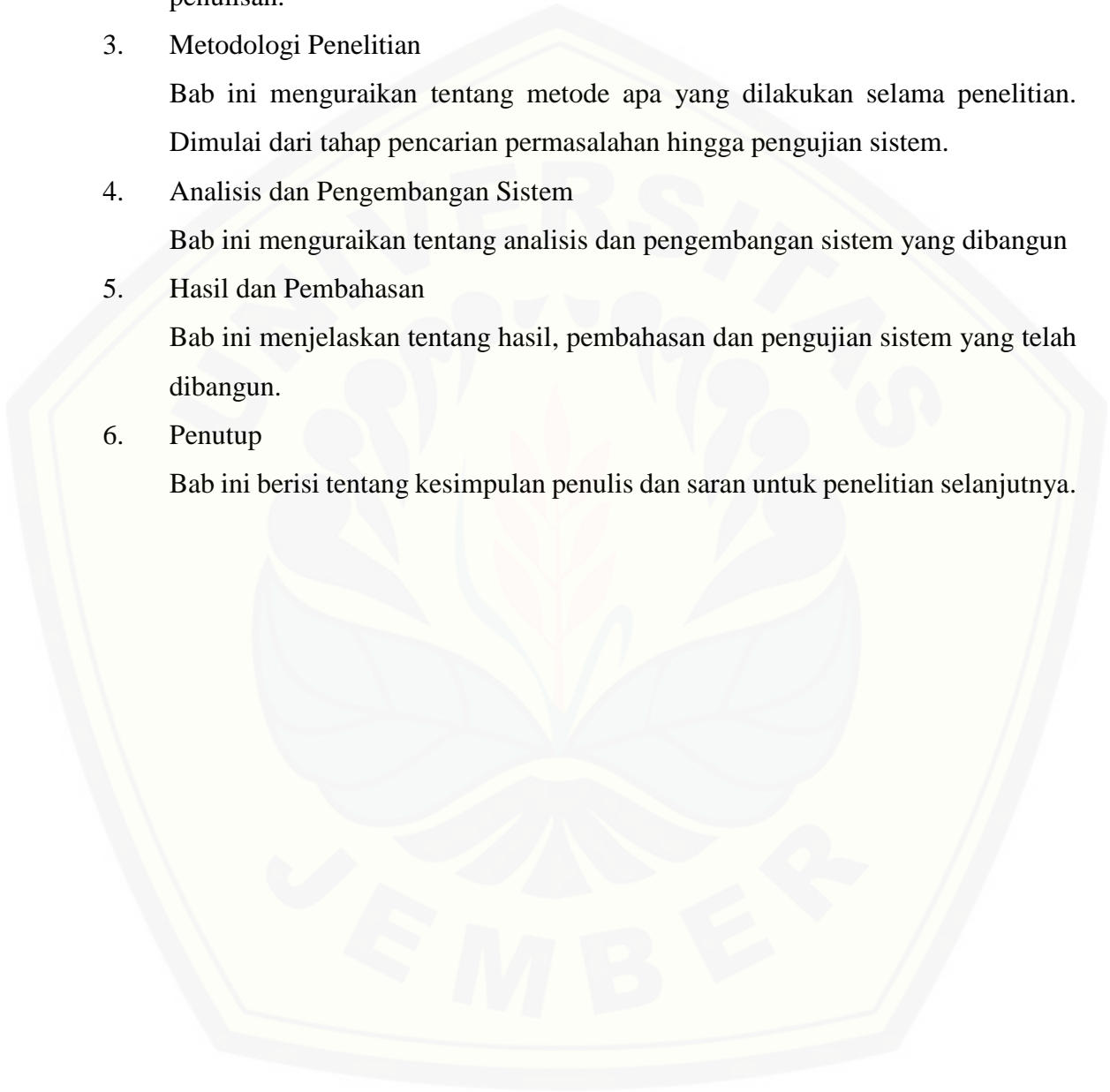
Bab ini menguraikan tentang analisis dan pengembangan sistem yang dibangun

5. Hasil dan Pembahasan

Bab ini menjelaskan tentang hasil, pembahasan dan pengujian sistem yang telah dibangun.

6. Penutup

Bab ini berisi tentang kesimpulan penulis dan saran untuk penelitian selanjutnya.



BAB 2. TINJAUAN PUSTAKA

Bab ini menguraikan lebih jauh mengenai teori-teori dan pustaka yang digunakan sebagai kerangka pemikiran dalam penelitian. Teori yang dibahas meliputi penelitian terdahulu, sistem informasi rekomendasi, analisis RFM dan Algoritma C4.5.

2.1 Penelitian Terdahulu

Algoritma C4.5 pernah digunakan untuk mengklasifikasikan predikat kelulusan mahasiswa fakultas tertentu dengan IPK sebagai salah satu parameter yang digunakan untuk memperhitungkan kelulusannya (Nugroho, 2014). Menurut hasil penelitian tersebut, algoritma C4.5 dapat mengklasifikasi mahasiswa menjadi beberapa kelas dan mengetahui kelompok mana yang diprediksi sebagai mahasiswa lulus dengan peringkat baik, atau yang di bawah rata-rata, dan mengetahui bagaimana cara untuk mengurangi kelompok mahasiswa di bawah rata-rata dengan strategi yang cocok.

Peneliti menyatakan bahwa Algoritma C4.5 dapat mengklasifikasi dengan atribut yang kondisional. Dalam penelitian yang berjudul “*Segmentation of Stock Trading Customers According to Potential Value*”, Sohn menggunakan algoritma K-Means, SOM, dan Fuzzy C-Means untuk mencari segmentasi konsumen pasar modal berdasarkan nilai potensialnya, yakni berdasarkan total nilai transaksi yang telah dilakukan dalam jangka waktu tiga bulan (Shon, 2004). Dari hasil perbandingan terhadap tiga metode tersebut, algoritma K-Means memiliki hasil segmentasi lebih akurat dengan atribut yang dibandingkan sama. Strategi pemasaran yang dapat diberikan untuk masing-masing segmentasi tersebut lebih tepat sasaran.

Berdasarkan penelitian yang berjudul “Klasifikasi Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5”, Algoritma C4.5 digunakan untuk mengklasifikasi pelanggan sebuah asuransi, dan mengelompokkan pelanggan dengan 2 kategori, yakni nasabah lancar dan nasabah tidak lancar (Sunjana, Klasifikasi Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5, 2010). Pola perilaku dari nasabah menjadi parameter dan untuk mendapatkan pola tersebut, menggunakan Algoritma C4.5.

Melihat pola yang telah terbentuk, dapat memperkirakan nasabah baru yang bergabung masuk kategori lancar atau tidak lancar sehingga memaksimalkan keuntungan perusahaan. Berdasarkan ulasan tersebut, maka metode data *mining classification decision tree* dinilai cocok digunakan untuk memprediksi kelompok pelanggan pengguna jasa pengiriman barang berdasarkan *track record* pelanggan.

2.2. Sistem Informasi Rekomendasi

Sistem informasi merupakan suatu sistem didalam organisasi yang mempertemukan kebutuhan pengolahan transaksi harian, mendukung operasi, bersifat manajerial, dan kegiatan strategis dari suatu organisasi dan menyediakan pihak luar tertentu dengan laporan-laporan yang diperlukan (Jogiyandi, 2005). Menurut (Kadir, 2003) sistem informasi adalah suatu sistem buatan manusia yang secara umum terdiri atas sekumpulan komponen berbasis komputer dan manual yang dibuat untuk menghimpun, menyimpan dan mengolah data serta menyediakan informasi keluaran kepada pemakai. Rekomendasi menurut Kamus Besar Bahasa Indonesia adalah penyuguhan; saran yang menganjurkan (membenarkan, menguatkan) (Setiawan, 2015). Rekomendasi merupakan saran yang dapat diberikan berdasarkan hasil analisis tertentu. Berdasarkan beberapa pernyataan tersebut dapat disimpulkan bahwa sistem informasi rekomendasi merupakan suatu media yang mengolah suatu komponen dan menghasilkan keluaran berupa informasi yang menyarankan sesuatu terhadap informasi yang telah diolah.

2.3. Analisis RFM

Berdasarkan penelitian yang berjudul “*Classifying the Segmentation of Customer Value via RFM Model and RS Theory*”, Ching mengelompokan pelanggan di C-Company dengan menggunakan analisis RFM (*Recency, Frequency and Monetary*) (Ching-Hsue Cheng, 2009). Menurut hasil dari penelitian tersebut, Cheng dan Chen menyimpulkan bahwa kelompok dasar dari banyaknya pelanggan tersebut dapat dibagi berdasarkan kapan transaksi terakhirnya, tingkat keseringan pelanggan

melakukan transaksi, dan besar nilai transaksi yang dilakukan. Atribut tersebut dapat membagi pelanggannya lebih spesifik. Model RFM diaplikasikan secara luas pada analisis pemasaran dan merupakan tool yang umum digunakan untuk membangun strategi pemasaran (Jo-Ting Wei, 2010). Segmentasi pelanggan berdasarkan analisis RFM menghasilkan kemampuan segmentasi antara 75% sampai 85% (Hughes, 2000).

Berikut adalah penjelasan mengenai RFM menurut Cheng:

1. Recency, yaitu kapan terjadinya transaksi paling terakhir yang dilakukan oleh pelanggan
2. Frequency, yaitu seberapa sering pelanggan melakukan transaksi. Contoh: satu kali dalam satu minggu, tiga kali dalam sebulan.
3. Monetary, berapa besar nilai jumlah transaksi yang dilakukan oleh pelanggan. (Ching-Hsue Cheng, 2009)

Segmentasi pelanggan tidak hanya penting untuk membedakan mana pelanggan tetap dan mana yang bukan, tetapi juga dapat menjadikan sebuah tolok ukur bagi perusahaan untuk memanfaatkan pelanggan sebagai salah satu strategi pemasaran. Mempertahankan pelanggan dengan mengetahui potensi masing-masing pelanggan bagi perusahaan lebih mudah daripada mempertahankan semua pelanggan dengan memberikan perlakuan yang sama. Proses data *mining* membantu perhitungan analisa jenis pelanggan pengguna jasa pengiriman barang.

2.4. Data Mining Classification

Data *mining* adalah suatu istilah sebuah proses yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data *mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksikan dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar (E. Turban, 2005).

Data *mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Tugas utama dari

data mining dikelompokkan menjadi *description*, *estimation*, *prediction*, *classification*, dan *assotiation* (Larose, 2005). Klasifikasi merupakan proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan objek yang digunakan. Salah satu teknik klasifikasi yang paling populer digunakan adalah *decision tree* (Han & Kamber, 2006). Data mining digunakan sebagai alat bantu untuk analisis pola data dengan jumlah data yang banyak. Penggunaan data mining dinilai sesuai untuk membagi pelanggan menjadi beberapa segmen dengan logika pengolahan data dalam jumlah besar dan menemukan keteraturan data atau pola yang tersembunyi.

2.5. Algoritma C4.5

Salah satu metode klasifikasi *decision tree* adalah Algoritma C4.5. Algoritma ini merupakan pengembangan dari Algoritma *Iterative Dichotomiser 3* (ID3). ID3 adalah salah satu teknik data mining klasifikasi *decision tree learning* yang paling dasar. Algoritma ini melakukan pencarian secara menyeluruh (*greedy*) pada semua kemungkinan pohon keputusan (Wahyudin, 2009). Hasil dari pengembangan ID3 yaitu, algoritma C4.5 dapat menghasilkan pohon keputusan yang lebih akurat. Algoritma ini dapat menangani data *training* dengan nilai atribut yang hilang, serta dapat menangani atribut kontiyu atau numerik.

Decision tree menggunakan representasi struktur pohon. *Node* paling atas pada *decision tree* disebut *root*. Setiap *node* merepresentasikan atribut, sedangkan cabangnya merepresentasikan nilai dari atribut, dan daunnya merepresentasikan kelas. *Decision tree* adalah *flowchart* seperti struktur *tree*, dimana tiap *internal node* menunjukkan sebuah *test* pada sebuah atribut, tiap cabang menunjukkan hasil dari *test*, dan *leaf node* menunjukkan *class-class* atau *class distribution* (Sunjana, 2010).

Terdapat 3 jenis *node* pada metode *decision tree*, yaitu:

1. Root Node

Root node merupakan bagian paling atas dari *decision tree*. *Root node* tidak memiliki *input* dan memiliki *output* lebih dari satu.

2. Internal Node

Internal node merupakan *node* percabangan. *Node* hanya terdapat satu *input* dan memiliki dua atau lebih *output*.

3. Leaf Node

Leaf node merupakan *node* akhir dari *decision tree*. *Node* ini memiliki satu *input* dan tidak memiliki *output*.

Pada tahap pembelajaran algoritma C4.5 memiliki prinsip kerja sebagai berikut:

1. Pembuatan pohon keputusan. Tujuan dari algoritma penginduksian pohon keputusan adalah mengkontruksi struktur data pohon yang dapat digunakan untuk memprediksi kelas dari sebuah kasus atau *record* baru yang belum memiliki kelas. C4.5 melakukan konstruksi pohon keputusan dengan metode *divide and conquer*. Pada awalnya hanya dibuat *node* akar dengan menerapkan algoritma *divide and conquer*. Algoritma ini memilih pemecahan kasus-kasus yang terbaik dengan menghitung dan membandingkan *gain ratio*, kemudian *node-node* yang terbentuk di level berikutnya, algoritma *divide and conquer* akan diterapkan lagi sampai terbentuk daun-daun.
2. Pembuatan aturan-aturan (rule set). Aturan-aturan yang terbentuk dari pohon keputusan akan membentuk suatu kondisi dalam bentuk *if-then*. Aturan-aturan ini didapat dengan cara menelusuri pohon keputusan dari akar sampai daun. Setiap *node* dan syarat percabangan akan membentuk suatu kondisi atau suatu *if*, sedangkan untuk nilai-nilai yang terdapat pada daun akan membentuk suatu hasil atau *then*.

Pemilihan atribut sebagai akar didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Perhitungan untuk mendapatkan pohon keputusan adalah sebagai berikut (Larose, 2005):

1. Perhitungan *entropy* kelas menggunakan rumus berikut.

$$Entropy_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} * Entropy(D_j) \quad \dots\dots\dots \text{persamaan(1)}$$

2. Perhitungan nilai *entropy* masing-masing atribut dapat dilihat pada rumus berikut:

$$Entropy(D) = \sum_{i=1}^m p_i \log_2(p_i) \quad \dots\dots\dots \text{persamaan(2)}$$

3. Perhitungan nilai *Gain*:

$$Gain(A) = Entropy(D) - Entropy_A(D) \quad \dots\dots\dots \text{persamaan(3)}$$

4. Perhitungan *split information* untuk menghitung *gain-ratio* dengan menggunakan rumus berikut:

$$SplitInfo_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} * \log_2 \left(\frac{|D_j|}{|D|} \right) \quad \dots\dots\dots \text{persamaan(4)}$$

5. Perhitungan *gain-ratio*:

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \quad \dots\dots\dots \text{persamaan(5)}$$

Dimana :

D : Himpunan Kasus (data)

A : Atribut (attribute)

|D_i| : Jumlah kasus pada partisi ke-i

|D| : Jumlah kasus dalam D

Proses partisi pohon keputusan akan berhenti saat:

1. Semua *record* dalam simpul N mendapat kelas yang sama.
2. Tidak ada atribut dalam *record* yang dapat dipartisi lagi.
3. Tidak ada *record* didalam cabang yang kosong

Penjelasan persamaan (1) adalah rumus yang digunakan untuk mencari *entropy* masing-masing kelas. *Entropy* mengukur tingkat ke-tidakpasti-an atau ke-acak-an dalam suatu data *set*. *Entropy* merupakan kebutuhan bit untuk menyatakan suatu kelas, yang jika semakin kecil nilai *entropy* maka semakin baik untuk digunakan dalam ekstraksi kelas. Selanjutnya hasil dari persamaan (1) dimasukkan dalam persamaan (2) yakni mencari *entropy* atribut. Persamaan (3) digunakan untuk mencari *gain*. *Gain* didapat dari pengurangan *entropy* umum dengan hasil *entropy* pada persamaan (2). Pencarian *split info* menggunakan persamaan (4). *Split info* menggunakan fungsi

logaritma dari atribut masing-masing kelas. Tahap akhir perhitungan merupakan variabel *gain ratio*, yang merupakan hasil pembagian dari *gain* dengan *split info* seperti pada persamaan (5).

2.6. Metode pengujian Algoritma Classification

Pengujian algoritma klasifikasi pada data *mining* dengan mengukur tingkat akurasi algoritma tersebut. Salah satu metode yang dapat digunakan adalah *confussion matrix*. *Confussion matrix* merupakan pengukuran tabel klasifikasi yang bersifat prediktif (Xhemali, 2009)

Tabel 2.1 Model *Confusion Matrix*

Actual	Prediction	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Tabel 2.1 menjelaskan bahwa data set tersebut menggunakan dua kategori kelas yakni *positive* dan *negative*. TP (True Positive) merupakan jumlah *record positive* yang bernilai benar, FP (False Positive) merupakan jumlah *record negative* yang bernilai benar, FN (False Negative) merupakan jumlah *record negative* yang bernilai salah dan TN (True Negative) merupakan jumlah *record negative* yang bernilai salah. Langkah pengujian dilakukan dengan menguji data *testing*, kemudian masukan ke dalam *confusion matrix*. *Confusion matrix* dapat menghitung *accuracy* dan *error rate*. *Accuracy* merupakan persentase dari prediksi yang benar, dan *error rate* merupakan kebalikan dari *accuracy* yaitu presentase jumlah *record* data yang tidak diklasifikasikan dengan benar (Xhemali, 2009). Perhitungan nilai *accuracy* dan *error rate* menggunakan persamaan berikut:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad \dots\dots\dots \text{persamaan(6)}$$

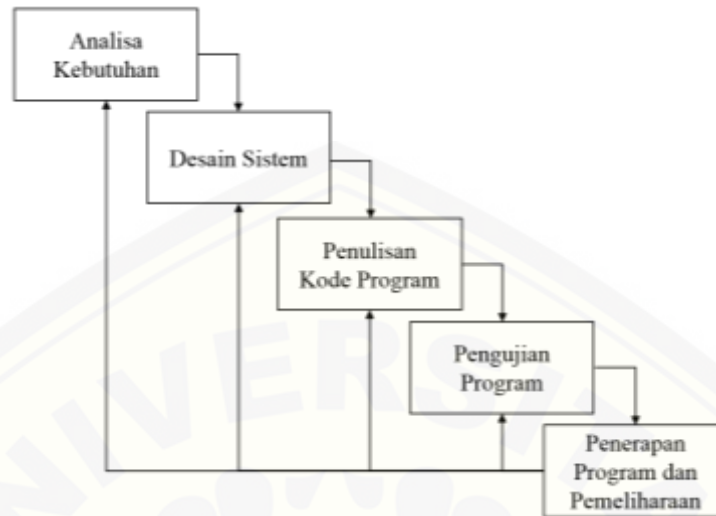
$$Error\ rate = \frac{FP+FN}{TP+FP+TN+FN} \times 100\% \quad \dots\dots\dots \text{persamaan(7)}$$

Keterangan:

- TP = True Positive
- FP = False Positive
- FN = False Negative
- TN = True Negative

2.7. Model Waterfall

Model pengembangan rancang bangun sistem informasi rekomendasi perubahan status ini menggunakan *Waterfall*. *Waterfall* model adalah sebuah contoh dari proses perencanaan, dimana semua proses kegiatan harus terlebih dahulu direncanakan dan dijadwalkan sebelum dikerjakan (Sommerville, 2011). Model *Waterfall* adalah model yang menyarankan sebuah pendekatan sistematis dan sekuensial dalam pembangunan sebuah perangkat lunak melalui tahapan-tahapan yang ada pada SDLC (System Development Life Cycle). Model ini dinilai tepat untuk digunakan dengan rancang bangun model yang sederhana dan tidak kompleks. Penulisan kebutuhan di awal yang telah jelas juga mendukung penggunaan model ini. Tahap umum model *Waterfall*, yakni analisa kebutuhan, desain, implementasi, pengujian dan pemeliharaan (Pressman, 2010). Diagram alir dari *Waterfall* ditunjukkan pada Gambar 2.5.



Gambar 2.1 Diagram Alir Waterfall

2.7.1 Analisa Kebutuhan

Analisis kebutuhan merupakan tahap awal yang dilakukan dalam model *waterfall*. Tahap ini merupakan tahap pencarian dan penulisan kebutuhan untuk membangun sistem. Pada tahap ini penulis terlebih dahulu melakukan analisis kebutuhan dengan melakukan observasi dan wawancara di tempat penelitian.

2.7.2 Desain

Pembangunan sistem ini digambarkan dengan menggunakan *Unified Modeling Language* (UML) yang dirancang dengan konsep *Object Oriented Programming* (OOP). Pemodelan UML yang digunakan adalah sebagai berikut:

1. *Business Process*
2. *Use Case Diagram*
3. *Use Case Scenario*
4. *Activity Diagram*
5. *Sequence Diagram*
6. *Class diagram*
7. *Entity Relationship Diagram* (ERD)

2.7.3 Implementasi

Pada tahap ini, kebutuhan sistem diimplementasikan dalam *coding* atau kode program. Penulisan kode program menggunakan bahasa pemrograman *Hyper Text Pre-Processor* (PHP), *Hyper Text Markup Language* (HTML), *Cascading Style Sheet* (CSS) dan *Javascript*. *Database Management System* (DBMS) yang digunakan adalah *MySQL* dengan *tool* XAMPP.

2.7.4 Pengujian

Tahap pengujian dilakukan untuk mengetahui kesalahan-kesalahan baik secara logika maupun fisik pada sistem yang telah dibangun. Pengujian dilakukan dengan *white box testing* dan *black box testing*. *white box testing* dilakukan oleh penulis program, dengan metode pengujian *cyclomatic complexity*. *Black box testing* dilakukan oleh user, yakni karyawan perusahaan tempat penelitian dengan menguji bagian fungsionalitas sistem.

2.7.5 Pemeliharaan

Tahap pemeliharaan suatu software diperlukan dalam sebuah pengembangan untuk menjaga sistem tetap berjalan dan berkembang sesuai dengan fungsinya. Pengembangan dilakukan ketika terdapat perubahan (*update*) atau perbaikan pada sistem.

BAB 3. METODOLOGI PENELITIAN

Bab ini menguraikan tentang metode yang dilakukan selama penelitian, pendekatan penelitian, waktu dan tempat penelitian dan tahap penelitian untuk menjawab rumusan masalah yang telah dijelaskan.

3.1. Pendekatan Penelitian

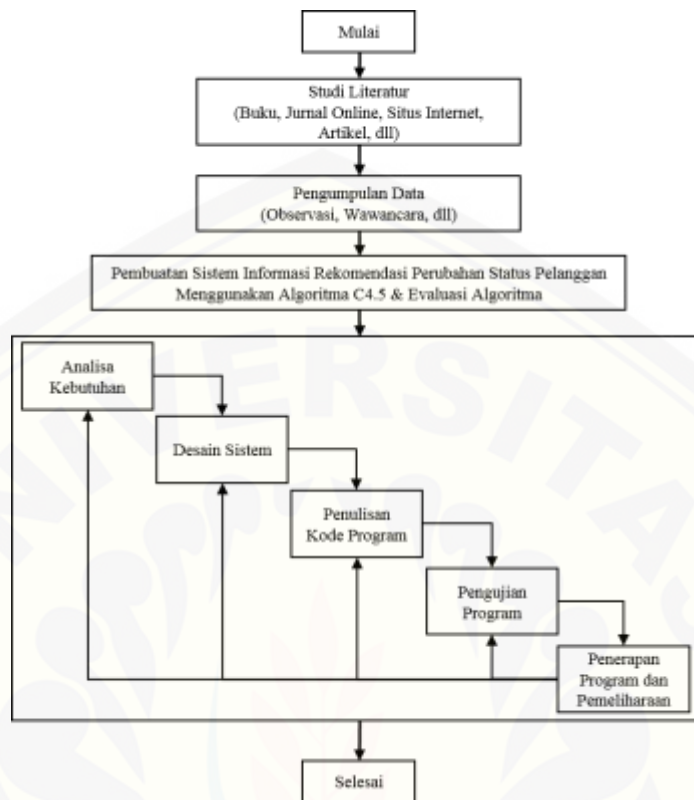
Pendekatan penelitian yang digunakan adalah metode penelitian kuantitatif dan kualitatif. Jenis pendekatan kuantitatif digunakan sebab penelitian ini menganalisa studi kasus banyaknya pelanggan pada jasa ekspedisi pengiriman barang menggunakan kereta, dan jenis pendekatan kualitatif digunakan untuk mengkaji kualitas dari teori yang sudah ada sebelumnya.

3.2. Waktu dan Tempat

Penelitian dilakukan di PT. Herona Express Jember. Waktu dilaksanakannya penelitian adalah selama enam bulan yaitu pada bulan Maret 2015 hingga bulan Agustus 2015.

3.3. Tahap Penelitian

Beberapa tahapan penelitian yang dilakukan adalah metode pengumpulan data, studi literatur dan pengembangan sistem. Metode yang diimplementasikan dalam membangun sistem dengan Algoritma C4.5 dalam menentukan status perubahan pelanggan.



Gambar 3.1 Alur Penelitian

Sumber: (Hasil Analisis, 2015)

Penjelasan Gambar 3.1 adalah beberapa tahap penelitian dimulai dengan studi literatur, pengumpulan data dengan beberapa cara seperti wawancara atau observasi untuk menganalisis objek. Selanjutnya adalah perancangan, pembangunan dan pengujian sistem dengan menggunakan Algoritma C4.5 melalui tahapan model *waterfall* dalam pengerjaan sistem. Berikut penjelasan lengkap dari Gambar 3.1:

3.3.1 Studi Literatur

Studi literatur merupakan tahap pertama melakukan penelitian. Studi literatur dilakukan untuk mendapatkan gambaran umum maupun khusus mengenai objek maupun teori pendukung dalam penelitian ini. Studi literatur yang digunakan yakni: buku pedoman, buku *online*, jurnal *online*, dan skripsi terkait dengan penelitian yang dibutuhkan.

3.3.2 Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam penelitian. Data-data yang dibutuhkan yakni:

a. Data primer

Data primer merupakan sumber data yang diperoleh secara langsung dari sumber informasi atau pihak pertama. Contoh data primer yang dibutuhkan pada penelitian ini adalah data member, data barang, data kota dan data karyawan untuk digunakan sebagai data pendukung pengolahan data transaksi sistem yang menerapkan algoritma C4.5.

Terdapat beberapa atribut dan kelas pada data set yang dibutuhkan dalam penelitian ini. Pemilihan atribut menggunakan analisis RFM dan ditambah dengan beberapa atribut biasa. Atribut yang digunakan meliputi *recency*, *frequency*, *monetary*, berat dan koli. Tabel 3.1 berikut merupakan rincian atribut dan nilai atribut yang digunakan sebagai data set.

Tabel 3.1 Atribut dan Nilai Atribut

No.	Atribut	Nilai
1.	<i>Recency</i>	Kategori (Baru atau Lama)
2.	<i>Frequency</i>	Kategori (Jarang atau Sering)
3.	<i>Monetary</i>	Numeris
4.	Berat	Numeris
5.	Koli	Numeris

Sumber : (Hasil Analisis, 2015)

Kelas yang digunakan terdiri dari dua kategori yakni berubah dan tetap. Penentuan kategori tersebut merupakan prediksi perubahan status pelanggan berdasarkan *track record* yang dilakukan oleh pelanggan pada periode tertentu.

b. Data Sekunder

Data sekunder merupakan data pendukung yang diperoleh dari sumber lain selain tempat penelitian namun masih berkaitan dengan objek penelitian.

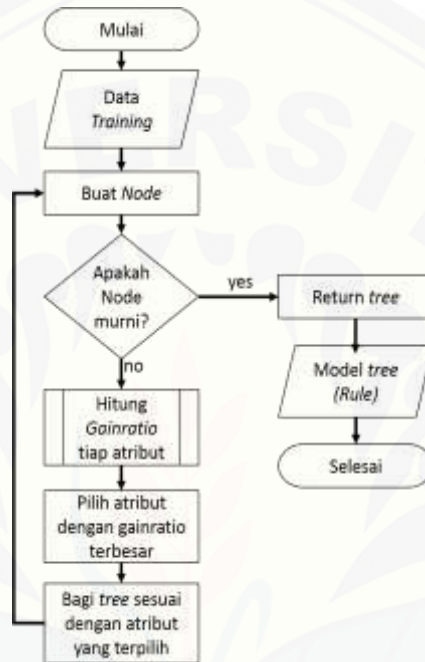
3.3.3 Penerapan Algoritma C4.5 untuk Rekomendasi Perubahan Status Pelanggan

Sistem informasi rekomendasi perubahan status memprediksi status baru dari pelanggan berdasarkan analisis *track record* dari pelanggan selama 6 bulan. Prediksi tersebut menggunakan *rule* yang dihasilkan dari proses data *mining* dengan menggunakan salah satu algoritma *decision tree* yakni Algoritma C4.5. Diagram alir proses *learning* dapat dilihat pada gambar 3.2.



Gambar 3.2 Flowchart Proses Learning

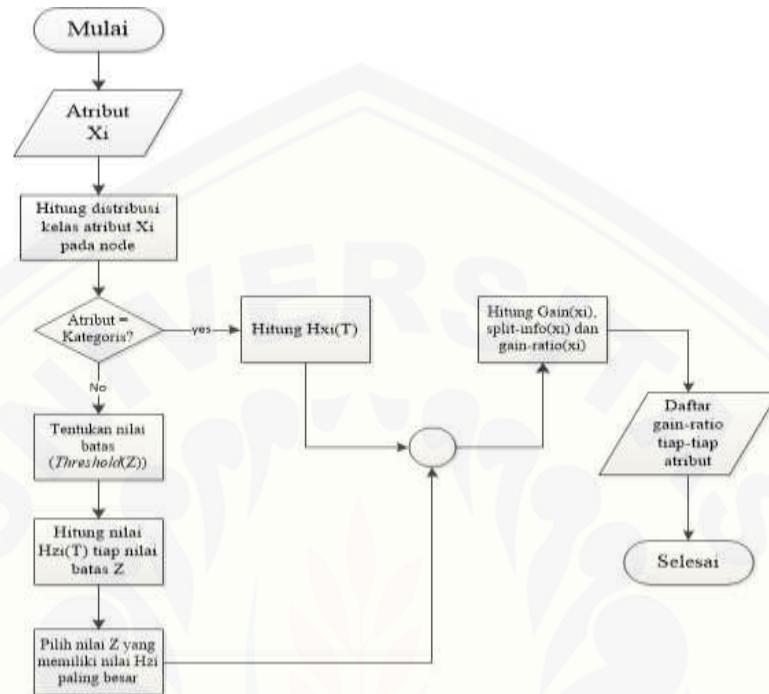
Penjelasan gambar 3.2 adalah setelah data *training* didapat, kemudian langkah berikutnya adalah membangun *tree* untuk membuat pohon keputusan. Proses dilanjutkan dengan mengonversi pohon keputusan menjadi sebuah *rule*. Proses pembangunan *tree* ditunjukkan pada gambar 3.3.



Gambar 3.3 Flowchart pembangunan *tree*

Penjelasan gambar 3.3 adalah setelah mendapat data *training*, kemudian membuat *node*. Jika *node* murni maka kembali ke *tree*, jika *node* tidak murni maka dilakukan perulangan dan menghitung kembali *gain ratio*. Pilih atribut dengan *gain ratio* paling besar, kemudian membuat kembali *tree* atau pohon keputusannya. Proses iterasi akan berakhir ketika *node* telah murni dan tidak dapat dibagi lagi. *Node* dikatakan murni apabila semua data didalamnya sudah memiliki kelas yang sama/hanya memiliki satu jenis kelas.

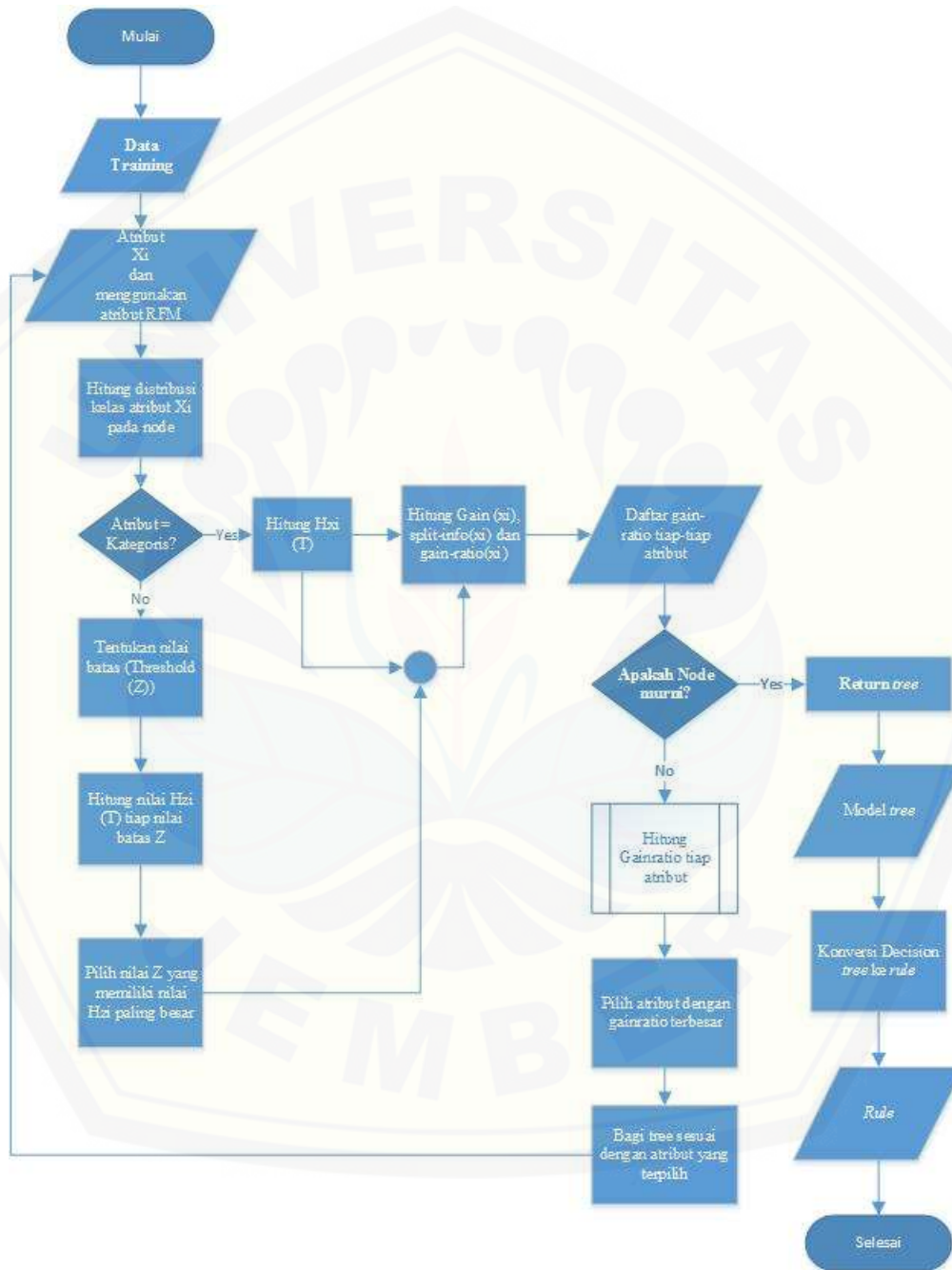
Pembuatan *node* dapat dilihat pada gambar 3.4.



Gambar 3.4 Flowchart pembuatan node

Penjelasan gambar 3.4 adalah cara membuat *node* diawali dengan mendapat data atribut x_i , lalu menghitung jumlah penyebaran kelas atribut x_i pada *node*. Kemudian cek apakah atribut merupakan kategori, jika tidak maka tentukan nilai batas z lalu hitung *entropy* tiap nilai batas Z . Pilih nilai batas yang memiliki nilai H_{zi} paling tinggi untuk menjadi nilai batas atribut yang akan dibandingkan dengan atribut lainnya setelah mendapatkan nilai *entropy* umum dengan menggunakan persamaan (1), menghitung *entropy* tiap atribut menggunakan persamaan (2). Selanjutnya menghitung *gain* (x_i) dengan persamaan (3), *split-info* (x_i) menggunakan persamaan (4) dan *gain-ratio* (x_i) menggunakan persamaan (5). Jika atribut merupakan numeris, cek terlebih dahulu *gain* batas mana yang paling besar. Setelah mendapatkan batas terbesar, ambil nilai atributnya, kemudian hitung nilai *split-info* (x_i) dan *gain-ratio* (x_i).

Diagram alir penerapan dari analisis RFM pada Algoritma C4.5 dijelaskan pada gambar 3.5.



Gambar 3.5 Flowchart implementasi Algoritma C4.5 pada segmentasi pelanggan

Penjelasan gambar 3.5 adalah proses penerapan Algoritma C4.5 pada segmentasi pelanggan dengan tahapan yang pertama adalah input data terlebih dahulu. Data *training* yang telah diinputkan dihitung dengan analisis RFM yakni menggunakan atribut pada RFM sebagai salah satu parameter. Proses prediksi *class* didapat dari *rule* yang telah dikonversi sebelumnya dari hasil membangun pohon keputusan. Prediksi *class* yang di dapat dari *record* transaksi data pelanggan kemudian digunakan untuk merekomendasikan kelas member apakah berubah atau tetap sesuai dengan perilaku pelanggan selama periode yang telah ditentukan.

3.3.4 Pengembangan Sistem

Pembuatan sistem pada penelitian ini menggunakan model *waterfall*. Model ini merupakan model sederhana dan banyak diimplementasikan oleh para pengembang software. Terdapat lima tahap pada model *waterfall* yang digunakan pada penelitian ini, yaitu analisa kebutuhan, desain, implementasi, pengujian dan pemeliharaan yang dilakukan secara urut. Setelah tahap pengumpulan data selesai, data akan dianalisis menggunakan RFM dan salah satu decision tree yaitu Algoritma C4.5. Tahap berikutnya adalah merepresentasikan analisis kebutuhan pada pemodelan UML (Unified Modelling Language). Pemodelan yang digunakan meliputi: *Business Process*, *Usecase Diagram*, Skenario, *Sequence Diagram*, *Activity Diagram*, *Class Diagram* dan ERD (Entity Relationship Diagram). Setelah tahap penulisan desain selesai, dilanjutkan dengan implementasi dalam *code program*. Hasil perancangan diuji menggunakan *white box* dan *black box*.