

BAB II KERANGKA TEORITIS

A. Landasan Teori

1. Pengertian Data Mining

Mengutip (Tan, 2006 dalam Eko Prasetyo, 2013:2) menyatakan bahwa data mining adalah sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah data mining disebut juga knowledge discovery. Knowledge Discovery in Databases (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD (Turban, dkk 2001).

KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah data. Serangkaian proses tersebut memiliki tahap sebagai berikut (Tan, 2006) :

- a. Pembersihan data yaitu untuk membuang data yang tidak konsisten dan noise
- b. Integrasi data yaitu penggabungan data dari beberapa sumber
- c. Transformasi data yaitu data diubah menjadi bentuk yang sesuai untuk di mining
- d. Aplikasi teknik data mining yaitu proses ekstraksi pola dari data yang ada
- e. Evaluasi pola yang ditemukan yaitu proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan
- f. Presentasi pengetahuan yaitu dengan teknik visualisasi

Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna (Tan, 2006).

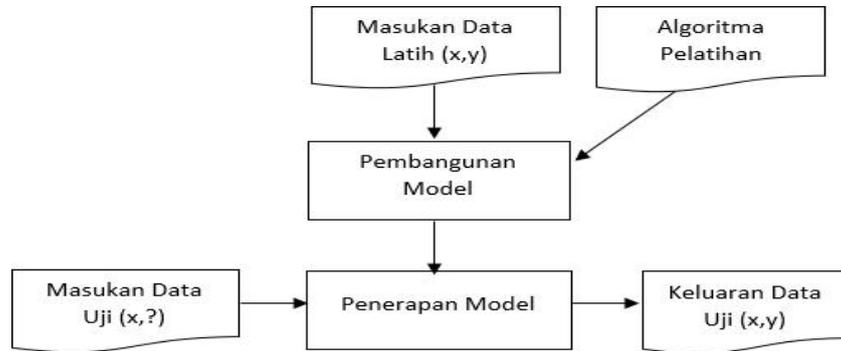
2. Pengertian Klasifikasi

Klasifikasi merupakan suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Aturan-aturan tersebut digunakan pada data-data baru untuk diklasifikasi.

Teknik ini menggunakan supervised induction, yang memanfaatkan kumpulan pengujian dari record yang terklasifikasi untuk menentukan kelas-kelas tambahan (Kusnawi, 2007). Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model sebagai prototipe untuk disimpan sebagai memori dan (2) penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya (Eko Prasetyo, 2012).

Klasifikasi dapat didefinisikan secara detail sebagai waktu pekerjaan yang melakukan pelatihan atau pembelajaran terhadap fungsi target / yang memetakan setiap vektor (set fitur) x ke depan satu dari sejumlah tabel kelas y yang tersedia pekerjaan pelatihan tersebut akan menghasilkan suatu model yang kemudian disimpan sebagai memori (Eko Prasetyo, 2013).

Model dalam dalam klasifikasi mempunyai arti yang sama dengan kotak hitam, dimana ada suatu model yang menerima masukan, kemudian mampu melakukan pemikiran terhadap masukan tersebut dan memberikan jawaban sebagai keluaran dari hasil pemikirannya. Kerangka kerja (*framework*) klasifikasi ditunjukkan pada Gambar 2.1. Pada gambar tersebut disediakan sejumlah data latih (x,y) untuk digunakan sebagai data pembangun model. Model tersebut kemudian dipakai untuk memprediksi kelas dari data uji $(x,?)$ sehingga diketahui kelas y yang sesungguhnya.



Gambar 2. 1 Proses Pekerjaan Klasifikasi

Klasifikasi digunakan untuk pembuatan model yang dapat melakukan pemetaan dari setiap himpunan variabel kesetiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan variabel yang baru didapat. Dalam pembangunan model selama proses pelatihan tersebut diperlukan adanya suatu algoritma untuk membangunnya yang disebut dengan Algoritma pelatihan. Ada banyak algoritma pelatihan yang sudah dikembangkan oleh para peneliti yaitu C4.5, Nearest Neighbor, Bayesian Classification, Neural Network, dll.

3. Pengembangan Sistem SDLC

Pendekatan sistem merupakan sebuah metodologi. Metodologi adalah satu cara yang direkomendasikan dalam melakukan sesuatu. Pendekatan sistem adalah metodologi dasar dalam memecahkan segala jenis masalah. Siklus hidup pengembangan sistem (System Development Life Cycle – SDLC) adalah aplikasi dari pendekatan sistem bagi pengembangan suatu sistem informasi (Raymond McLeod, 2007, p.200).

Terdapat beberapa tahapan pekerjaan pengembangan yang perlu dilakukan jika suatu proyek ingin memiliki kemungkinan berhasil yang besar. Tahapan-tahapan tersebut adalah:

- a. Perencanaan
- b. Analisis
- c. Desain
- d. Implementasi
- e. Penggunaan

Proyek dan sumber daya yang dibutuhkan untuk melakukan pekerjaan direncanakan kemudian disatukan. Sistem yang ada juga dianalisis untuk memahami masalah dan menentukan persyaratan fungsional dari sistem yang baru. Sistem baru ini kemudian dirancang dan diimplementasikan. Setelah implementasi, sistem kemudian digunakan, idealnya untuk jangka waktu yang lama.

Karena pekerjaan-pekerjaan di atas mengikuti satu pola yang teratur dan dilaksanakan dengan cara dari atas ke bawah, SDLC tradisional sering kali disebut sebagai pendekatan air terjun (waterfall approach). Aktivitas ini memiliki aliran satu arah menuju ke penyelesaian proyek.



Gambar 2. 2 Pola Melingkar dari Siklus Hidup Sistem

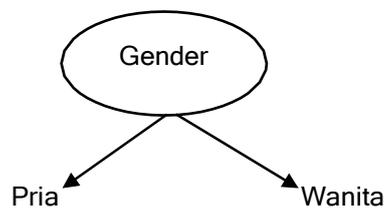
(Sumber: Raymond McLeod, 2007, p.200)

Gambar 2.2 mengilustrasikan sifat melingkar dari siklus hidup. Ketika sebuah sistem telah melampaui masa manfaatnya dan harus diganti, satu siklus hidup baru akan dimulai dengan diawali oleh tahap perencanaan.

Mudah bagi kita untuk melihat bagaimana SDLC tradisional dapat dikatakan sebagai suatu aplikasi dari pendekatan sistem. Masalah akan didefinisikan dalam tahap-tahapan perencanaan dan analisis. Solusi-solusi alternatif diidentifikasi dan dievaluasi dalam tahap desain. Lalu, solusi yang terbaik diimplementasikan dan digunakan. Selama tahap penggunaan, umpan balik dikumpulkan untuk melihat seberapa baik sistem mampu memecahkan masalah yang telah ditentukan.

4. Pengertian Metode Algoritma C4.5

Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal, bisa mengatasi *missing data*, bisa mengatasi data kontinu dan *pruning* (Larose, 2005).



Gambar 2. 3 Syarat pengujian fitur

Yang menjadi hal penting dalam induksi decision tree adalah bagaimana menyatakan syarat pengujian pada node. Ada 3 kelompok penting dalam syarat pengujian node:

a. Fitur biner

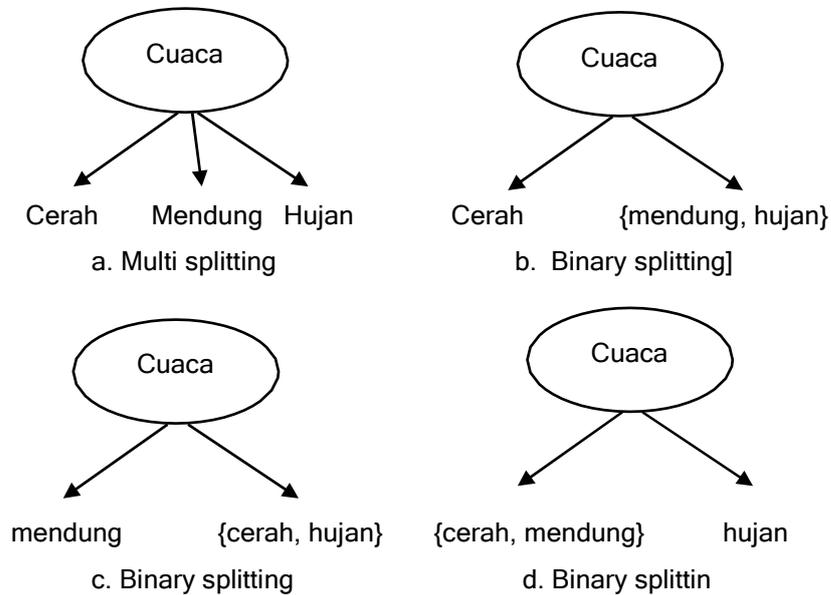
Fitur yang hanya mempunyai dua nilai berbeda disebut dengan fitur biner. Syarat pengujian ketika fitur ini menjadi node (akar maupun internal) hanya punya dua pilihan cabang. Contoh pemecahannya disajikan pada Gambar 2.3.

b. Fitur bertipe kategorikal

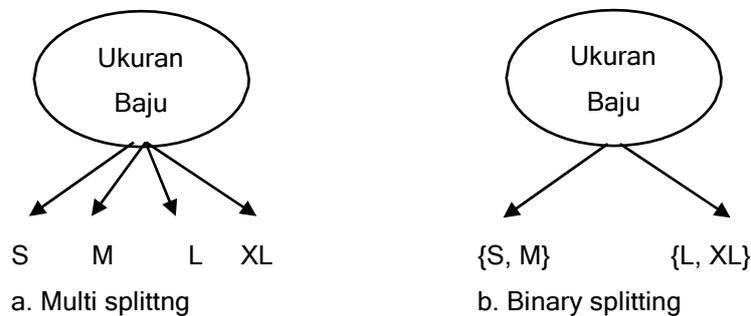
Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai beberapa nilai berbeda. Contohnya adalah fitur 'cuaca' mempunyai 3 nilai berbeda dan ini bisa mempunyai banyak kombinasi syarat pengujian pemecahan. Secara umum ada 2, yaitu pemecahan biner (*binary splitting*) dan (*multi splitting*). Kombinasinya disajikan seperti pada Gambar 2.4. Untuk pemecahan yang hanya membolehkan pemecahan biner, seperti algoritma CART, maka akan memberikan kemungkinan jumlah kombinasi pemecahan sebanyak 2^{k-1} , dimana k adalah jumlah nilai berbeda dalam fitur tersebut. Contoh pemecahannya disajikan pada Gambar 2.4 dan Gambar 2.5.

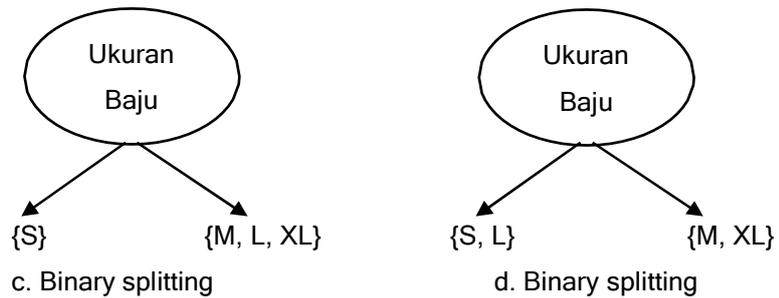
c. Fitur bertipe numerik

Untuk fitur bertipe numerik, syarat pengujian dalam node (akar maupun internal) dinyatakan dengan pengujian perbandingan ($A < v$) atau ($A \geq v$) dengan hasil biner, atau untuk multi dengan hasil berupa jangkauan nilai dalam bentuk $v_i \leq A < v_{i+1}$, untuk $i = 1, 2, \dots, k$. Untuk kasus pemecahan biner, maka algoritma akan memeriksa semua kemungkinan posisi pemecahan v dan memilih posisi v terbaik. Untuk cara multi, maka algoritma harus memeriksa semua kemungkinan jangkauan nilai *kontinyu*. Contoh pemecahan pada fitur numerik disajikan pada Gambar 2.6.



Gambar 2. 4 Syarat pengujian fitur bertipe nominal





Gambar 2. 5 Syarat pengujian fitur bertipe ordinal



Gambar 2. 6 Syarat pengujian fitur bertipe

Kriteria yang paling banyak digunakan untuk memilih fitur sebagai pemecah dalam algoritma C4.5 adalah rasio gain, yang diformulasikan oleh persamaan berikut:

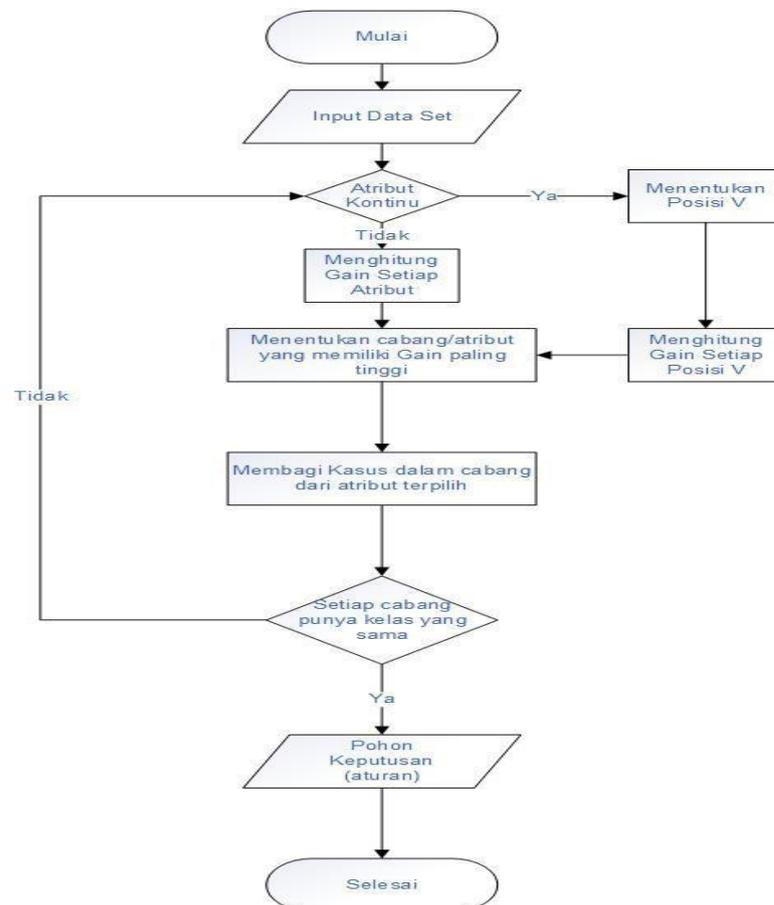
$$g(s, j) = \frac{g(s, j)}{g(s)}$$

Persamaan diatas menyatakan nilai rasio gain pada fitur ke- j . $SplitInfo(s, j)$ didapat dari $SplitInfo(s, j) = -\sum_{i=1}^k p(v_i|s) \log_2 p(v_i|s)$ dimana k menyatakan jumlah pemecahan .

Berikut ini adalah algoritma C4.5 menggunakan flowcart yang disajikan pada Gambar 2.6 dan berserta penjelasannya :

- Memasukan Data Set yang telah disediakan. Data Set adalah kumpulan objek dan atributnya.
- Pada proses cek keputusan dilihat apakah data set tersebut termasuk Atribut Kontinu atau bukan, Atribut Kontinu adalah atribut yang mempunyai jangkauan real.
- Jika Iya maka akan menentukan posisi V , setelah itu menghitung gain setiap V . Jika tidak maka menghitung gain setiap atribut.
- Setelah itu menentukan cabang/atribut yang memiliki gain paling tinggi.
- Setelah itu membagi kasus dalam cabang dari atribut yang terpilih.

- f. Jika setiap cabang mempunyai kelas yang sama, maka menghasilkan pohon keputusan. Jika tidak maka cabang atau atribut tersebut kembali melakukan cek atribut kontinu.



Gambar 2. 7 Flowchart Algoritma C4.5

Sumber : (Larose, 2005)

5. Pengertian Metode Prototype

Menurut (Ogedebe, 2012) menyampaikan bahwa prototyping merupakan metode pengembangan perangkat lunak, yang berupa model fisik kerja sistem dan berfungsi sebagai versi awal dari sistem. Dengan metode prototyping ini akan dihasilkan prototype sistem sebagai perantara pengembang dan pengguna agar dapat berinteraksi dalam proses kegiatan pengembangan sistem informasi. Agar proses pembuatan prototype ini berhasil dengan baik adalah dengan mendefinisikan kebutuhan awal. Prototype akan dihilangkan atau ditambahkan pada bagiannya sehingga sesuai dengan perencanaan dan analisis yang dilakukan oleh pengembang

sampai dengan uji coba dilakukan secara simultan seiring dengan proses pengembangan. Ada 4 metodologi prototyping yang paling utama yaitu :

1. Illustrative, menghasilkan contoh laporan dan tampilan layar.
2. Simulated, mensimulasikan beberapa alur kerja sistem tetapi tidak menggunakan data real.
3. Functional, mensimulasikan beberapa alur sistem yang sebenarnya dan menggunakan data real.
4. Evolutionary, menghasilkan model yang menjadi bagian dari operasional sistem.

Dibuatnya sebuah prototyping bagi pengembang sistem bertujuan untuk mengumpulkan informasi dari pengguna sehingga pengguna dapat berinteraksi dengan model prototype yang dikembangkan, sebab prototype menggambarkan versi awal dari sistem untuk kelanjutan sistem sesungguhnya yang lebih besar.

(Ogedebe, 2012), menegaskan telah ditemukan bahwa dalam analisis dan desain sistem, terutama untuk proses transaksi, dimana dialog yang ditampilkan lebih mudah dipahami. Semakin besar interaksi antara komputer dan pengguna, besar pula manfaat yang diperoleh ketika proses pengembangan sistem informasi akan lebih cepat dan membuat pengguna akan lebih interaktif dalam proses pengembangannya.

Prototyping dapat diterapkan pada pengembangan sistem kecil maupun besar dengan harapan agar proses pengembangan dapat berjalan dengan baik, tertata serta dapat selesai tepat waktu. Keterlibatan pengguna secara penuh ketika prototype terbentuk akan menguntungkan seluruh pihak yang terlibat, bagi pimpinan, pengguna sendiri serta pengembang sistem.

Manfaat lainnya dari penggunaan prototyping adalah :

1. Mewujudkan sistem sesungguhnya dalam sebuah replika sistem yang akan berjalan, menampung masukan dari pengguna untuk kesempurnaan sistem.
2. Pengguna akan lebih siap menerima setiap perubahan sistem yang berkembang sesuai dengan berjalannya prototype sampai dengan hasil akhir pengembangan yang akan berjalan nantinya.
3. Prototype dapat ditambah maupun dikurangi sesuai berjalannya proses pengembangan. Kemajuan tahap demi tahap dapat diikuti langsung oleh pengguna.

4. Penghematan sumber daya dan waktu dalam menghasilkan produk yang lebih baik dan tepat guna bagi pengguna.

6. Pengertian Pelanggan Potensial

Seseorang pelanggan atau konsumen adalah orang atau organisasi yang berinteraksi dengan produk, jasa, atau proses dan kemungkinan merupakan pelanggan aktif (Harvey Thompson, 2000: p42).

Seseorang dikatakan setia atau loyal apabila pelanggan tersebut menunjukkan perilaku pembelian secara teratur atau terdapat suatu kondisi dimana mewajibkan pelanggan membeli paling sedikit dua kali dalam selang waktu tertentu.

Pelanggan yang potensial umumnya akan melanjutkan pembelian produk atau jasa tersebut walaupun dihadapkan pada banyak alternative produk atau jasa yang lebih unggul di pandang dari berbagai sudut atributnya. Menurut Kotler didalam buku Zulkarnain, konsumen yang loyal tidak diukur dari berapa banyak dia membeli, tapi berapa sering dia melakukan pembelian ulang. termasuk merekomendasikan orang untuk membeli. (Zulkarnain, 2012).

B. Tinjauan Pustaka

Penelitian rujukan merupakan acuan yang dibutuhkan seorang peneliti untuk melakukan penelitian. Penelitian rujukan pada penelitian ini diambil berdasarkan kesamaan metode yang digunakan yaitu Algoritma C4.5. Banyak penelitian yang menggunakan metode ini dalam berbagai kasus. Antara lain :

1. PENERAPAN DATA MINING UNTUK MENGANALISA JUMLAH PELANGGAN AKTIF DENGAN MENGGUNAKAN ALGORITMA C4.5 (Annisak Izzati Jamhur Universitas Putera Indonesia

YPTK Padang) Berdasarkan uraian pada bab-bab yang sudah dibahas sebelumnya dapat diambil beberapa kesimpulan, yaitu : 1. Pemilihan variabel Total Pembelian, Jumlah Barang dan PO Pertahun, dapat menjadi kriteria penilaian terhadap pelanggan aktif dan tidak aktif dengan menggunakan Algoritma C Algoritma C4.5 dapat menghasilkan rule untuk menggambarkan proses yang terkait dengan pelanggan aktif dan tidak aktif. 3. Metode pohon keputusan yang diproses dengan WEKA lebih membantu perusahaan dalam hal menganalisa jumlah pelanggan aktif. Saran Mengingat keterbatasan yang dimiliki oleh penulis, baik pengetahuan, waktu, maupun pemikiran, maka penulis dapat memberikan beberapa gambaran sebagai saran. Saran

ini dapat digunakan untuk reverensi di masa yang akan datang maupun untuk pengembangan lebih lanjut mengenai analisa jumlah pelanggan aktif menggunakan Algoritma C4.5 adalah sebagai berikut : 1. Perlu adanya penambahan variabel tertentu yang dilakukan untuk menganalisa jumlah pelanggan aktif, agar rule yang dihasilkan bisa lebih tepat lagi. 2. Pada penelitian ini, penulis hanya mencoba salah satu teknik yaitu decision tree. Untuk mendapatkan hasil yang lebih baik lagi, dapat digunakan beberapa teknik data mining lainnya. 3. Dalam penelitian ini penulis hanya menggunakan software data mining WEKA, perlu dilakukan perbandingan lebih lanjut terhadap software aplikasi data mining lainnya, seperti Rapid Miner, Orange, dan lain-lain, guna melihat hasil dari pengolahan yang dilakukan terhadap software yang digunakan. 19

Kontribusi : jumlah barang

2. IMPLEMENTASI METODE TOPSIS DAN SAW DALAM MEMBERIKAN REWARD PELANGGAN (Agus Wardana Windarto Program Studi Sistem Informasi , STIKOM Tunas Bangsa Pematangsiantar)

Dalam sebuah industri penjualan, persaingan merupakan hal yang wajar. Banyak nya usaha-usaha dengan jenis yang sama membuat seseorang pengusaha harus memiliki strategi-strategi yang tepat dalam meningkatkan daya beli pelanggan dan menuai keuntungan penelitian ini bertujuan untuk mengimplementasikan algoritma dalam ilmu komputer untuk membuat sistem pendukung keputusan pembeian reward kepada pelanggan Depot Air minum. Dalam penelitian ini metode yang digunakan adalah TOPSIS dan SAW. Dimana sampel yang digunakan sebanyak 6 pelanggan dengan kriteria penilaian adalah status pembayaran, status keaktifan pelanggan, lama berlangganan, jumlah pembelian, dan waktu pembelian dan hasil perbandingan kedua metode tersebut , diperoleh hasil bahwa perhitungan yang dilakukan dengan metode TOPSIS lebih baik di bandingkan dengan metode SAW

Kontribusi : keaktifan pelanggan, lama berlangganan

3. SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PELANGGAN TERBAIK PADA TB. BANGUN JAYA MENGGUNAKAN METODE WEIGHTED PPRODUCT (Ristika Oktavina , Teknik Informatika, Fakultas Ilmu Komputer , Universitas Dian Nuswantoro)

Berdasarkan penelitian yang sudah ada sebelumnya sudah menentukan pelanggan terbaik, penulis bermaksud untuk membuat sistem dimana dapat menentukan pelanggan terbaik yang nantinya akan diberikan bonus pada pelanggan TB. Bangun Jaya masih dalam bentuk hard copy dan keputusan dari satu pihak saja sehingga proses yang dilakukan masih belum akurat. Sistem yang akan dibuat ini merupakan sistem pendukung keputusan dengan metode Weight Product. Latar belakang yang dipilih maka diperoleh suatu perumusan masalah yaitu bagaimana menentukan kriteria yang tepat dan menentukan pelanggan yang layak diberi bonus dalam penilaian terhadap pelanggan terbaik dengan sistem pendukung keputusan menggunakan metode weight product. Setelah dilakukan penyeleksian kriteria didapat kriteria yang akan digunakan yaitu Total Belanja, Pembayaran, Kuantitas Belanja, Loyalitas Kunjungan

Kontribusi : Total Belanja

4. ANALISA DAN PENERAPAN METODE ALGORITMA C4.5 UNTUK PREDIKSILOYALITAS PELANGGAN(Teguh Budi Santoso, Jurusan teknik Informatika, Fakultas Teknik)

Dari data table dibawah dapat dilihat nilai akurasi serta eror rate data training serta data testing, dimana data training memiliki nilai akurasi yang lebih tinggi dari pada data testing dengan menggunakan algoritma C4.5 nilai algoritma C4.5 dapat digunakan pada data set pelanggan. Berdasarkan analisa pengguna data mining dengan algoritma C4.5 dapat digunakan pada data set pelanggan kedalam kegiatan manajemen strategi sehingga dapat menahan selama mungkin pelanggannya dengan baik, algoritma c4.5 ini hendaknya dimasukkan kedalam personal computer yang ada sehingga dapat mengurangi kemungkinan pelanggan untuk berpindah operator atau untuk mempertahankan loyalitas pelanggan

Tabel 2. 1 Tabel Akurasi Jurnal

Data set	Akurasi (%)	Eror Rate (%)
Data Training	97,5%	2,5%
Data Testing	93,3%	6,6%

5. PENGUJIAN USABILITY WEBSITE MENGGUNAKAN SYSTEM USABILITY SCALE (Ika Aprilia H.N, Program Studi Teknik Informatika, Universitas Gajah Mada Yogyakarta)

Dari hasil pengujian sistem usability dapat menjadi langkah awal untuk awal evaluasi website.

6. SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PENERIMA AWARD DI AGEN TOKO TIKET ON-LINE MENGGUNAKAN METODE SIMPLE ADDITIVE WEIGHT (Agus Pamuji, Program studi Teknik Informatika, Indraprasta PGRI)

Pada saat menggunakan metode FMADM dengan menggunakan metode SAW terdapat kriteria yang di perlukan untuk menentukan siapa yang akan terseleksi sebagai penerima award. Adapun kriteria penerima award adalah rata-rata item pemesanan, lama berlangganan, jumlah tunggakan/tagihan, jumlah investasi transaksi, jumlah pembatalan, frekuensi complain..

7. PERBANDINGAN METODE NAÏVE BAYES CLASSIFIER DENGAN METODE DECISION TREE (C4.5) UNTUK MENGANALISA KELANCARAN PEMBIAYAAN (Triowali, Program studi Teknik Informatika, Bussines Institut Darmajaya)

Salah satu algoritma klasifikasi yang sering digunakan untuk mendapat banyak perhatian para peneliti dalam memprediksi pembiayaan bermasalah pada perbankan adalah naïve bayes clasiffier dan decision tree (C4.5). Tetapi kelemahan yang dihadapi pada kedua algoritma tersebut adalah lamanya waktu dan tingkat akurasi pediksi yang digunakan untuk melakukan prediksi. Masalah ini juga menjadi perhatian banyak peneliti untuk memperbaikinya agar kinerja waktu dan akurasi prediksi menjadi lebih singkat tetapi kinerja akurasi tetap baik. Dalam melakukan prediksi tingkat ketetapan pembiayaan, dengan menggunakan pemodelan metode naïve bayes classifier didapatkan tingkat akurasi tertinggi sebesar 50%. Sedangkan pda pemodelan decision tree (c4.5) memperoleh rata-rata tingkat akurasi tertinggi sebesar 56%. Oleh karena itu dapat disimpulkan berdasarkan tingkat akurasi, bahwa pemodelan metode decision tree (c4.5) lebih baik dalam melakukan pediksi pembiayaan pada KSSPS/BMT Al-fadhila.

8. IMPLEMENTASI DATA MINING UNTUK PEMILIHAN PELANGGAN POTENSIAL MENGGUNAKAN ALGORITMA K-MEANS (Randi Rian Putra, Fakultas Sains dan Teknologi , Universitas Pembangunan Panca Budi Medan)

Penelitian ini dilakukan untuk mmpelajari data mining menggunakan algoritma k-means yang di implementasikan dalam pemilihan pelanggan potensial di MC.Laundri.Hasil dari proses data mining ini dapat digunakan sebagai pertimbangan dalam mengambil keputusan leih lanjut tentang

penetapan pelanggan potensial. Dan hasil output yang di peroleh adalah “potensial” dan “tidak potensial.

9. PENERAPAN SIMPLE ADDITIVE WEIGHTING (SAW) UNTUK PRIORITAS PENJADWALAN OPERASI PASIEN DI RUANG OK UROLOGI RSCM JAKARTA (muhammad adiyatama, Program studi sistem informasi, sekolah tinggi ilmu komputer stikom binaniaga)

Dalam proses uji coba penelitian kali ini menggunakan paket kuesioner PSSUQ dimana pengguna (4 dokter penjadwal) mengisi kuesioner tersebut untuk mengetahui apakah pengguna bias menggunakan aplikasi dengan baik dan sesuai dengan tujuan penelitian dalam memprioritaskan penjadwalan pasien operasi di ruang ok urologi RSCM Jakarta.

10. IDENTIFIKASI PELANGGAN POTENSIAL PRODUK ASURANSI DENGAN TEKNIK KLASIFIKASI (Adhitya Akbar, Universitas Gadjah Mada)

Berdasarkan studi kasus data pembelian produk asuransi pada dunia nyata, didapat decision tree sebagai model klasifikasi yang lebih baik dari naïve bayes, dengan tingkat akurasi total (overall accuracy) sebesar 80.51% sedangkan naïve bayes hanya sebesar 69.32%.

Tabel 2. 2 Tabel Tinjauan Pustaka

NO.	TAHUN	JUDUL	PENYUSUN	KONTRIBUSI
1	2016	Penerapan data Mining untuk menganalisa jumlah pelanggan aktif dengan menggunakan algoritma c4.5 (Annisak Izzati Jamhur Universitas Putera Indonesia YPTK Padang)	(Annisak Izzati Jamhur Universitas Putera Indonesia YPTK Padang)	Kontribusi pada penelitian ini adalah mengambil variabel untuk di jadikan variabel Jumlah barang yang digunakan

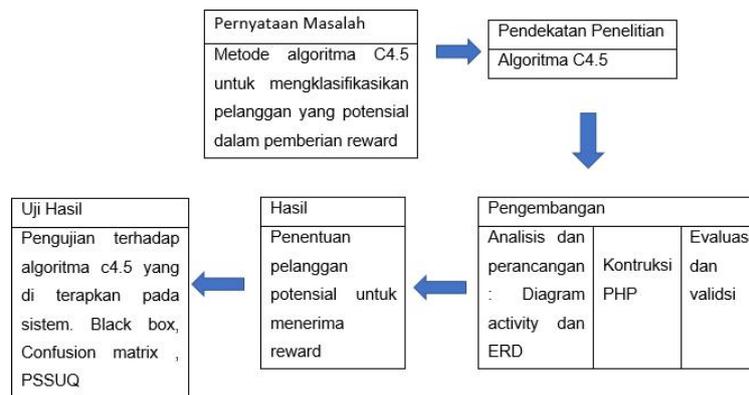
NO.	TAHUN	JUDUL	PENYUSUN	KONTRIBUSI
2	2017	IMPLEMENTASI METODE TOPSIS DAN SAW DALAM MEMBERIKAN REWARD PELANGGAN	(Agus Wardana Windarto Program Studi Sistem Informasi , STIKOM Tunas Bangsa Pematangsiantar)	Kontribusi pada penelitian ini adalah mengambil variabel untuk di jadikan variabel keaktifan pelanggan, untuk digunakan
3	2016	Sistem Pendukung Keputusan untuk menentukan pelanggan terbaik pada TB. Bangun Jaya Menggunakan Metode Weighted	(Ristika Oktavina , Teknik Informatika, Fakultas Ilmu Komputer , Universitas Dian Nuswantoro)	Kontribusi pada penelitian ini adalah mengambil variabel untuk dijadikan variabel JumlahTransaksi
4	2016	Analisa dan penerapan Metode Algoritma C4.5 untuk prediksi loyalitas pelanggan	Teguh Budi Santoso, Jurusan teknik Informatika, Fakultas Teknik	Kontribusi dalam penelitian ini adalah pengambilan uji hasil yang menggunakan confusion Matrix

NO.	TAHUN	JUDUL	PENYUSUN	KONTRIBUSI
5	2016	Pengujian unsability website menggunakan system usability scale	Ika Aprilia, Teknik Informatika , Universitas Gajah Mada, Yogyakarta	Kontribusi Penelitian ini adalah pengambilan pengujian sistem dengan SUS
6	2015	Sistem oendukung keputusan untuk menentukan penerima award di agen toko tiket online menggunakan metode Simple Additive Weight	Agus Pamuji, Program studi Teknik Informatika, Indraprasta PGRI	Kontribusi pada penelitian ini adalah mengambil variabel untuk di jadikan variable lama beralanggan untuk digunakan
7	2016	Perbandingan metode naïve bayes classifier dengan metode decision tree (c4.5) menganalisa untuk kelancaran Pembiayaan	Triowali, Program studi Teknik Informatika, Bussines Institut Darmajaya	Kontribusi pada penelitian ini adalah mengambil pada kelebihan perbandingan metode yaitu algorima c4.5

NO.	TAHUN	JUDUL	PENYUSUN	KONTRIBUSI
8.	2018	Implementasi data mining untuk memilih pelanggan potensial menggunakan Algoritma K-Means	Randi Rian Putra, Fakultas Sains dan Teknologi, Universitas Pembangunan Panca Budi Medan	Kontribusi pada penelitian ini adalah mengambil hasil output "potensial" dan "tidak potensial"
9.	2019	Penerapan Simple Additive Weight (SAW) untuk prioritas penjadwalan operasi pasien di ruang ukurologi Rscm Jakarta	muhammad adiyatama, Programstudi sistem informasi, sekolah tinggi ilmu computer stikom binaniaga	Kontribusi pada penelitian ini adalah mengambil uji pengguna dengan menggunakan PSSUQ
10.	2015	Identifikasi Pelanggan Potensial produk asuransi dengan teknik klasifikasi	Adhitya Akbar, Universitas Gadjah Mada	Kontibusi dalam penelitian ini yaitu hasil rule untuk menentukan pelanggan potensial

C. Kerangka Pemikiran

Berikut adalah kerangka pemikiran untuk memecahkan masalah penelitian ini yang dapat dilihat pada Gambar 2.8



Gambar 2. 8 Kerangka Pemikiran

Pendekatan menggunakan algoritma C4.5 menjadi dasar hasil penelitian, pengembangan terbagi menjadi tiga yaitu analisa perancangan, kontruksi sistem dengan menggunakan PHP dan setelah kontruksi selesai dilakukan, dilanjutkan dengan implementasi terdiri dari tiga tahap yaitu pertama analisa kebutuhan sistem degan menggunakan use case, diagram activity dan ERD (Entity Relational Diagram) juga menggunakan data penjualan sebagai bahan analisisnya. Kedua kontruksi PHP yaitu data transaksi yang di uji menggunakan PHP dengan algoritma C4.5. Ketiga tahap evaluasi dan validasi, tahap tersebut sistem menguji ketepatan agar memperoleh informasi yang tepat dari penelitian ini dengan menggunakan data penjualan. Setelah mendapatkan hasil, kemudian dilakukan uji hasil yaitu menguji ketepatan hasil perhitungan algortima c4.5 apakah sudah sesuai atau belum dengan menggunakan confusion matrix.

D. Hipotesis

Hipotesis dalam penelitian ini adalah penerapan metode Algoritma C4.5 diduga dapat menentukan pelanggan potensial dan pemberian reward.