

ANALISIS STATUS PELANGGAN MENGGUNAKAN METODE *DECISION TREE* DENGAN ALGORITMA C4.5 STUDI KASUS PADA PT. WISANAMITRA ARGAKARYA

Kevin Wirawan¹⁾ dan Akhmad Budi²⁾

¹⁾ Alumni Program Studi Sistem Informasi

²⁾ Staf Pengajar Program Studi Teknik Informatika

Institut Bisnis dan Informatika Kwik Kian Gie

Jl. Yos Sudarso Kav.87 Sunter Jakarta Utara 14350

<http://www.kwikkiangie.ac.id>

akhmad.budi@kwikkiangie.ac.id

ABSTRACT

Along with the development of information technology related to business development that provides a lot of choice of alternative solutions to the organization's use of information technology and the company has become one of the main requirements as a means to increase the value added in the business process. Use of Data Mining can be interpreted as a method of how to improve the quality of business decisions based on data-related systems. Data Mining processes data into information, and then they are processed more further into knowledge that will help the process of analysis of the company effectively. This research used C4.5 algorithm to form a decision tree. The decision tree is a classification and prediction methods. Decision tree method changes the very large fact into a decision tree that represents the rules. The purpose of this research is analyzing the customers of PT. Wisanamitra Argakarya.

Kata Kunci: *Decision Tree, C4.5 algorithm, Decision Support System, Analysis.*

1. PENDAHULUAN

Perkembangan peradaban hidup manusia tidak bisa dilepaskan dengan perkembangan teknologi yang diperlukan dalam membantu kehidupan manusia. Manusia selalu berpikir untuk menciptakan teknologi yang dapat meringankan pekerjaan mereka. Teknologi tidak selalu harus dalam bentuk alat yang secara fisik terlihat. Teknologi juga bisa berupa sesuatu yang secara fisik tidak terlihat antara lain kerangka kerja (*framework*), metode, konsep pemikiran dan lain-lain. Beberapa contoh teknologi yang berguna bagi manusia antara lain teknologi di bidang pertanian, kedokteran, informasi dan komunikasi, sumber daya energi, transportasi dan lain-lain. Diantara teknologi-teknologi yang disebutkan, teknologi informasi dan komunikasi merupakan teknologi yang perkembangannya sangat pesat. Perkembangan tersebut ditandai dengan berkembangnya perangkat keras (*hardware*) dan perangkat lunak (*software*).

Penggunaan teknologi informasi dan komunikasi pada organisasi dan perusahaan sudah menjadi salah satu kebutuhan utama sebagai alat untuk meningkatkan nilai tambah dalam proses

bisnis. Penggunaan teknologi informasi dan komunikasi melibatkan data sebagai bahan baku informasi yang berguna bagi pengambilan keputusan. Sehubungan dengan semakin ketatnya persaingan bisnis, analisis data merupakan salah satu faktor yang menentukan. Perusahaan harus dapat menganalisis data perusahaan dengan tepat dan cermat agar dapat bertahan dalam persaingan bisnis. Analisis data yang dilakukan perusahaan akan digunakan dalam pengambilan keputusan dimana keputusan yang dipilih memiliki tingkat resiko yang minimum. Ketersediaan informasi yang cepat dan tepat akan sangat membantu para eksekutif pengambilan keputusan dalam menentukan kebijakan serta dapat menjadi salah satu faktor yang memposisikan organisasi atau perusahaan beberapa langkah lebih unggul dibanding kompetitornya.

Perusahaan perlu melakukan analisis data karena dengan adanya analisis data ini dapat digunakan untuk menilai kinerja perusahaan, dan digunakan untuk membandingkan kondisi perusahaan dari tahun sebelumnya dengan tahun sekarang. Dengan melakukan analisis data, pihak

perusahaan dapat mengetahui apakah kinerja perusahaan tersebut meningkat atau tidak sehingga perusahaan mempertimbangkan keputusan yang akan diambil untuk tahun yang akan datang sesuai dengan kinerja perusahaannya.

Analisis data dilakukan dengan mengikuti teori, aturan, metode atau algoritma tertentu yang sudah diakui manfaatnya. Penggunaan teori, aturan, metode atau pun algoritma tertentu tersebut terhadap data yang ada mengacu pada suatu bidang di komputasi yaitu *data mining*.

Data Mining berisi pencarian pola yang diinginkan dalam database besar untuk membantu dalam pengambilan keputusan diwaktu yang akan datang. Penggunaan *Data Mining* atau dapat diartikan sebagai suatu metode bagaimana meningkatkan kualitas pengambilan keputusan bisnis berdasarkan sistem yang berkaitan dengan data. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisis data yang kemudian dapat dipelajari dengan teliti dan bermanfaat. Dan salah satu teknik dari bagian *data mining* adalah klasifikasi, yaitu menentukan sebuah record data baru ke salah satu dari beberapa kategori atau kelas yang telah didefinisikan sebelumnya. Dan salah satu bagian teknik dari klasifikasi adalah *decision tree* (pohon keputusan).

Decision Tree merupakan representasi sederhana dari teknik klasifikasi yang merupakan proses pembelajaran suatu fungsi tujuan yang memetakan tiap himpunan atribut ke satu dari kelas yang didefinisikan sebelumnya. *Decision tree* merupakan salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Dengan kemampuannya untuk *break down* proses pengambilan keputusan yang kompleks menjadi lebih simple. *Decision Tree* juga dapat menemukan hubungan yang tidak terlihat secara nyata (tersembunyi) antara sejumlah calon variabel input dengan sebuah variabel target. Selain itu *decision tree* dapat memadukan antara eksplorasi data dan pemodelan, sehingga sangat baik sebagai langkah awal dalam proses pemodelan.

Algoritma C4.5 merupakan salah satu cara yang dapat dilakukan untuk membangun aturan yang direpresentasikan dalam bentuk *decision tree*. Fakta yang digunakan berasal dari tabel keputusan yang kemudian diproses dengan Algoritma C4.5, sehingga menghasilkan *decision tree*. Algoritma ini juga dapat menyederhanakan *knowledge* (pengetahuan) yang dimiliki sistem sehingga proses inferensi dapat menjadi lebih cepat.

Wisnamitra Argakarya merupakan sebuah perusahaan terbatas yang bergerak dalam instalasi *air conditioner* untuk gedung dan rumah mewah.

Untuk membantu kegiatan operasionalnya PT. Wisnamitra Argakarya sudah memiliki sistem informasi berbasis komputer. Perusahaan ini terus memantau status pelanggan yang terus bertambah dengan beragam permintaan. Sistem berjalan di PT. Wisnamitra Argakarya sudah terkomputerisasi dan sudah memiliki *database* untuk menyimpan datanya. Banyaknya jumlah transaksi harian menyebabkan adanya kesulitan dalam menangani data yang terus bertambah. Situasi ini menyebabkan hasil laporan yang kurang efektif dan akurat sehingga berdampak terhadap keterlambatan dan kurang efektifnya pengambilan keputusan.

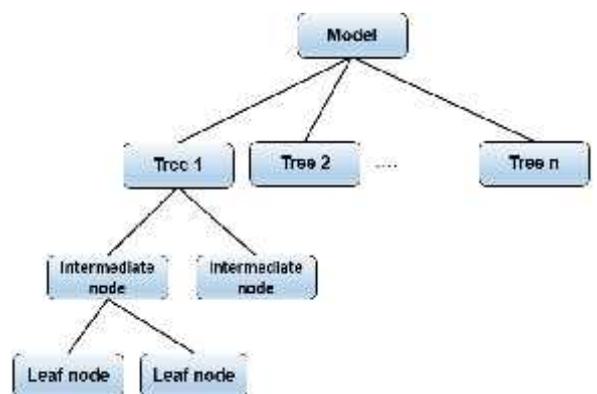
2. LANDASANTEORI/ KERANGKA PEMIKIRAN

Data mining merupakan pemilihan suatu *data* atau dengan istilah menggali pengetahuan dari jumlah *data* yang banyak [5]. Selain itu, penjelasan lain tentang *data mining* adalah proses menemukan pola yang menarik, dan pengetahuan dari *data* yang berjumlah besar [5]. *Data mining* merupakan teknologi baru yang sangat berguna untuk membantu perusahaan menemukan informasi yang sangat penting dari gudang data perusahaan. Beberapa aplikasi data mining berfokus pada prediksi, aplikasi tersebut meramalkan apa yang akan terjadi dalam situasi baru berdasarkan data yang menggambarkan apa yang terjadi di masa lalu [17].

Beberapa fungsi *data mining*, yaitu [9] :

1) Classification

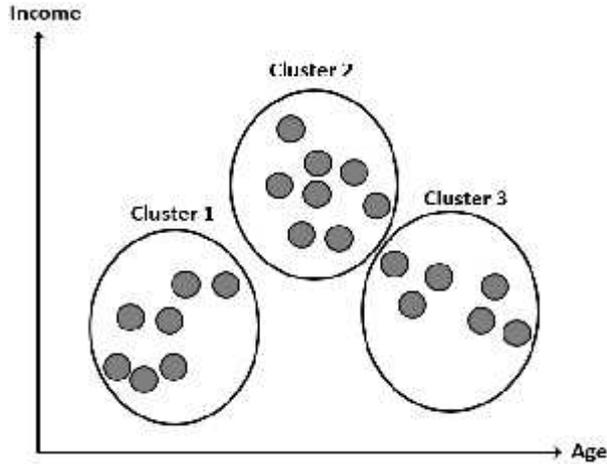
Berfungsi untuk mengklasifikasikan suatu *class* kedalam kategori yang akan dipilih.



Gambar 1. Classification – Decision Tree [9]

2) *Clustering*

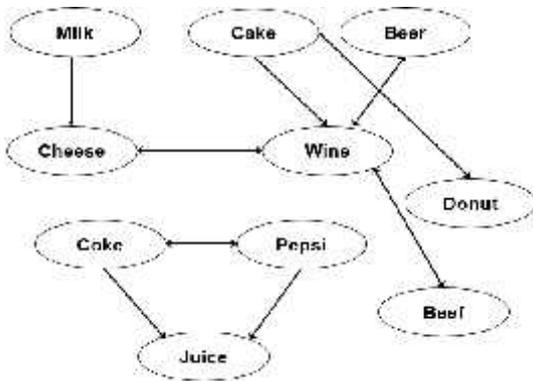
Berfungsi untuk mencari pengelompokan atribut kedalam segmentasi – segmentasi berdasarkan similaritas.



Gambar 2. Clustering [9]

3) *Association*

Berfungsi untuk mencari keterkaitan antara atribut atau *item set* berdasarkan jumlah *item* yang muncul dan *rule association* yang ada.



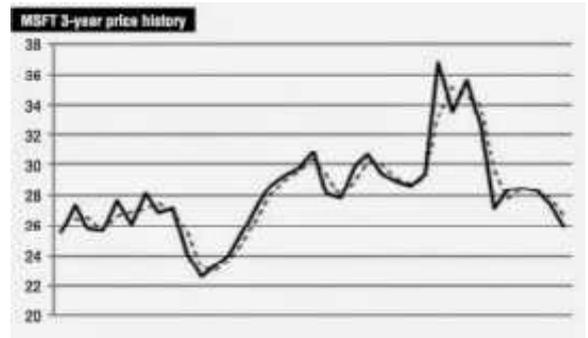
Gambar 3. Product Association [9]

4) *Regression*

Berfungsi untuk mencari prediksi dari suatu pola yang ada, hampir mirip dengan klasifikasi.

5) *Forecasting*

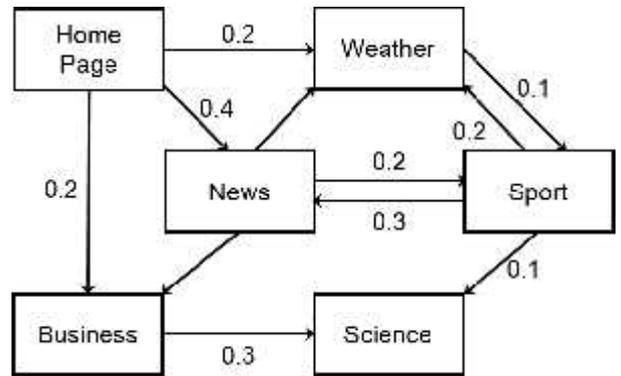
Berfungsi untuk peramalan waktu yang akan datang berdasarkan trend yang telah terjadi di waktu sebelumnya.



Gambar 4. Forecasting [9]

6) *Sequence Analysis*

Berfungsi untuk mencari pola urutan dari rangkaian suatu kejadian.



Gambar 5. Web Navigation Sequence [9]

7) *Deviation Analysis*

Berfungsi untuk mencari kejadian langka atau kejadian yang belum pernah terjadi sebelumnya dan kejadian yang sangat berbeda dari keadaan *normal*.

2.1 Decision Tree

Algoritma *decision tree* didasarkan pada pendekatan *divide-and-conquer* untuk klasifikasi suatu masalah. Algoritma tersebut bekerja dari atas ke bawah, mencari pada setiap tahap atribut untuk membaginya ke dalam bagian terbaik *class* tersebut, dan memproses secara rekursif submasalah yang dihasilkan dari pembagian tersebut. Strategi ini menghasilkan sebuah *decision tree* yang dapat diubah menjadi satu set *classification rules* [17].

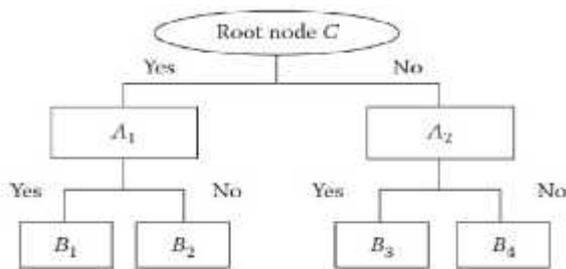
Decision tree merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (*tree*) di mana setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas. Node

yang paling atas dari *decision tree* disebut sebagai *root* [4].

Pada *decision tree* terdapat 3 jenis *node*, yaitu:

- a. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada input dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
- b. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu input dan mempunyai *output* minimal dua.
- c. *Leaf node* atau *terminal node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu input dan tidak mempunyai *output*.

Seperti ditunjukkan dalam Gambar 6, *decision tree* tergantung pada aturan *if-then*, tetapi tidak membutuhkan parameter dan metrik. Struktur sederhana dan dapat ditafsirkan memungkinkan *decision tree* untuk memecahkan masalah atribut *multi-type*. *Decision tree* juga dapat mengelola nilai-nilai yang hilang atau data *noise* [1].



Gambar 6. Contoh Struktur *Decision Tree* [1]

2.2 Algoritma C 4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk *decision tree*. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode *decision tree* mengubah fakta yang sangat besar menjadi *decision tree* yang merepresentasikan aturan. [6]

Salah satu algoritma yang digunakan untuk membangun *decision tree* yang berbasis algoritma induksi *decision tree* yaitu C4.5. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk *decision tree*.

Secara umum algoritma C4.5 untuk membangun *decision tree* adalah sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang.

4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti di bawah ini :

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

Dengan :

- S : Himpunan kasus
- |Si| : Jumlah kasus pada partisi ke i
- A : Atribut
- |S| : Jumlah kasus pada S
- n : Jumlah partisi atribut A

Sedangkan perhitungan nilai *entropy* dapat dilihat dari rumus 2 berikut :

$$E(S) = \sum_{i=1}^n -p * \log_2 p$$

Dengan :

- S : Himpunan kasus
- n : Jumlah partisi s
- A : Fitur
- pi : Proporsi dari Si terhadap S

3. METODOLOGI PENELITIAN

3.1 Objek Penelitian

Adapun objek dalam penelitian ini adalah PT Wisanamitra Argakarya. Objek penelitian ini beralamat di Sentra Bisnis Artha Gading A6-B No. 27 Jalan Boulevard Artha Gading, Kelapa Gading, Jakarta 14240. PT Wisanamitra Argakarya didirikan pada tahun 1988. PT Wisanamitra Argakarya merupakan *authorized dealer* dari Daikin di Indonesia (PT. Daikinaircon). PT Wisanamitra Argakarya telah memiliki pengalaman lebih dari 20 tahun di bidang air conditioning, serta dengan ditunjang oleh tenaga profesional menjadikan PT Wisanamitra Argakarya spesialis di bidang disain, sales, instalasi dan *after sales service* AC Daikin VRV System. Seperti halnya perusahaan lain, untuk tetap bisa bersaing di kerasnya persaingan di bidang bisnis *air conditioning*, maka PT Wisanamitra Argakarya menetapkan kepuasan pelanggan sebagai prioritas

utama dan bertekad untuk menjadi yang terbaik diantara perusahaan lain yang menjadi pesaing PT Wisanamitra Argakarya dalam bidang disain, sales, instalasi dan *after sales service* AC Daikin VRV System. Dalam menjalankan kegiatan operasionalnya PT Wisanamitra Argakarya melibatkan beberapa personil yang saling bekerja sama sesuai dengan *job description* masing-masing. Beberapa personil tersebut diatur dengan struktur organisasi.

3.2 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah dengan melakukan survey langsung pada PT. Wisanamitra Argakarya dan melakukan observasi pada dokumen dan pelaporan yang dibutuhkan.

Untuk mengetahui laporan-laporan apa saja yang dibutuhkan itu sendiri penulis melakukan wawancara langsung kepada General Manager PT. Wisanamitra Argakarya yaitu Bapak Raphael Edward dengan tujuan untuk menanyakan informasi apa saja yang ingin ditampilkan dalam pelaporan terkait pelanggan aktif dan tidak aktif yang terdaftar pada PT. Wisanamitra Argakarya.

3.3 Teknik Pengumpulan Data

Proses pengumpulan data dilakukan dengan cara:

1. Observasi Langsung
Cara ini dilakukan dengan melihat secara langsung bagaimana proses berjalannya system dilakukan dalam upaya mengumpulkan berbagai dokumen dan data-data untuk menghasilkan berbagai laporan yang dibutuhkan.
2. Wawancara Terstruktur
Cara ini dilakukan dengan melakukan wawancara langsung dengan beberapa pihak terkait untuk mengetahui kebutuhan informasi dan sistem yang ingin digali dan hendak dikembangkan. Sumber data yang diambil merupakan sumber data primer dimana data akan diperoleh secara langsung dari General Manager PT. Wisanamitra Argakarya yaitu Bapak Raphael Edward.

4. HASIL DAN PEMBAHASAN

4.1 Hasil

Dalam penelitian ini, pengolahan data dilakukan dengan memperhatikan beberapa hal yang menjadi bahan pertimbangan, yaitu :

- a) Mencari *node* dan *leaf*, serta *rule* yang dihasilkan dari *decision tree*.
- b) Atribut pada kasus tidak semuanya digunakan hanya 7 atribut dan 1 atribut target saja.
- c) Data yang merupakan input harus diuji.
- d) Data harus dipastikan keakuratannya dan bebas dari kesalahan.
- e) Terdapat 205 pelanggan yang harus ditentukan statusnya.

Dengan jumlah pelanggan yang cukup banyak dengan atribut target dari data ini adalah :

Tabel 1. Atribut Target

Atribut	Deskripsi
Yes	Pelanggan memiliki status member yang aktif
No	Pelanggan memiliki status member yang tidak aktif

Data yang digunakan dalam penelitian ini adalah data histori pelanggan yang diambil dari PT. Wisanamitra Argakarya. Data yang diperoleh mulai dari periode bulan Maret 2011 sampai dengan bulan Maret 2014. Data yang telah diperoleh memiliki jumlah transaksi lebih dari 10.000 transaksi yang berasal dari 205 pelanggan. Dari transaksi yang ada terdapat keterangan yang penting untuk proses analisis. Keterangan yang ada dapat dilihat pada tabel di bawah ini.

Tabel 2. Atribut Pelanggan PT. Wisanamitra Argakarya

No	Atribut	Deskripsi
1	No. Batch	Nomor surat jalan yang tertera pada transaksi
2	Tanggal Transaksi	Tanggal terjadinya transaksi
3	Keterangan	Alamat proyek yang melakukan transaksi
4	Proyek	Kode proyek yang dimiliki masing-masing pelanggan
5	Status	Status pembayaran transaksi
6	Total Pembiayaan	Jumlah biaya yang ada pada suatu transaksi
7	Tanggal Selesai	Tanggal pelunasan transaksi

Setelah pengolahan data dilakukan, keterangan yang ada dapat dijadikan sebagai variabel / atribut. Variabel atau atribut tersebut akan menentukan

hasil yang diinginkan untuk mengukur kriteria-kriteria yang akan ditentukan. Atribut yang ada dapat dilihat pada tabel di bawah ini.

Tabel 3. Penentuan Atribut

No	Atribut	Deskripsi
1	Tanggal transaksi	Atribut ini digunakan untuk menentukan variabel "Frekuensi Transaksi". Tanggal transaksi bertujuan untuk mengetahui kapan awal dan akhir transaksi proyek pada masa periode data yang ada. Dengan histori transaksi sesuai tanggal, maka dapat mengetahui seberapa sering frekuensi transaksi yang dilakukan oleh proyek dalam suatu periode dari transaksi yang dilakukan pada periode tertentu.
2	Kode proyek	Atribut ini untuk memilah dan memudahkan dalam proses pemilahan transaksi yang dilakukan masing-masing proyek.
3	Keterangan	Dalam hal ini, keterangan yang diberikan adalah alamat proyek. Alamat yang diberikan oleh proyek dapat membantu menentukan variabel "Status Pelanggan" karena dari alamat yang diberikan dapat diketahui apakah lokasi proyek merupakan perumahan pribadi atau lokasi proyek merupakan sebuah perusahaan.
4	Status	Maksud dari atribut ini adalah kejelasan mengenai kelunasan pembayaran yang dilakukan oleh proyek. Variabel "Status Transaksi" dapat ditentukan dari atribut ini. Jika dalam surat jalan status berisikan selesai, berarti pembayaran lancer dan begitu pun sebaliknya.
5	Total pembiayaan	Atribut ini membantu menentukan variabel "Jumlah Amount". Variabel ini ditentukan dari total seluruh amount yang telah diberikan proyek kepada perusahaan.
6	Tanggal selesai	Variabel "Jenis Pesanan" dapat diketahui dengan tanggal selesai yang dicetak pada surat jalan. Dengan ini, perusahaan dapat menentukan berapa lama proyek telah bekerja sama dengan perusahaan.

Tabel 4. Pemilahan Data

Tanggal	Proyek	Keterangan	No. Transaksi	Amount	Status	Tanggal Selesai
30-Jan-14	PJ-217 IDR	Diamond Golf Kav. 70-71	14.000178	3,623,599.98	Proses	Belum
3-Feb-14	PJ-204 IDR	Pinisi Permai II Blok H6 No.8	14.000197	5,479,300.00	Selesai	3-Apr-14
5-Feb-14	PJ-270 IDR	Permata Hijau A7, Jakarta	14.000217	3,504,000.00	Proses	Belum
5-Feb-14	PJ-271 IDR	Jl. Martimbang 5 / 9,	14.000219	2,299,800.00	Selesai	5-Aug-14
8-Feb-14	PJ-272 IDR	Sekolah Kencana III No.3,	14.000247	3,522,600.00	Selesai	8-Oct-14
8-Feb-14	PJ-275 IDR	Jl. Diamond Golf No. 56	14.000250	2,397,600.00	Proses	Belum
12-Feb-14	PJ-167 IDR	Jl. Imam Bonjol 18	14.000269	2,500,000.00	Proses	Belum
15-Feb-14	PJ-204 IDR	Jl. Simprug Golf III No. 161	14.000270	4,500,000.00	Proses	Belum

4.2 Pembahasan

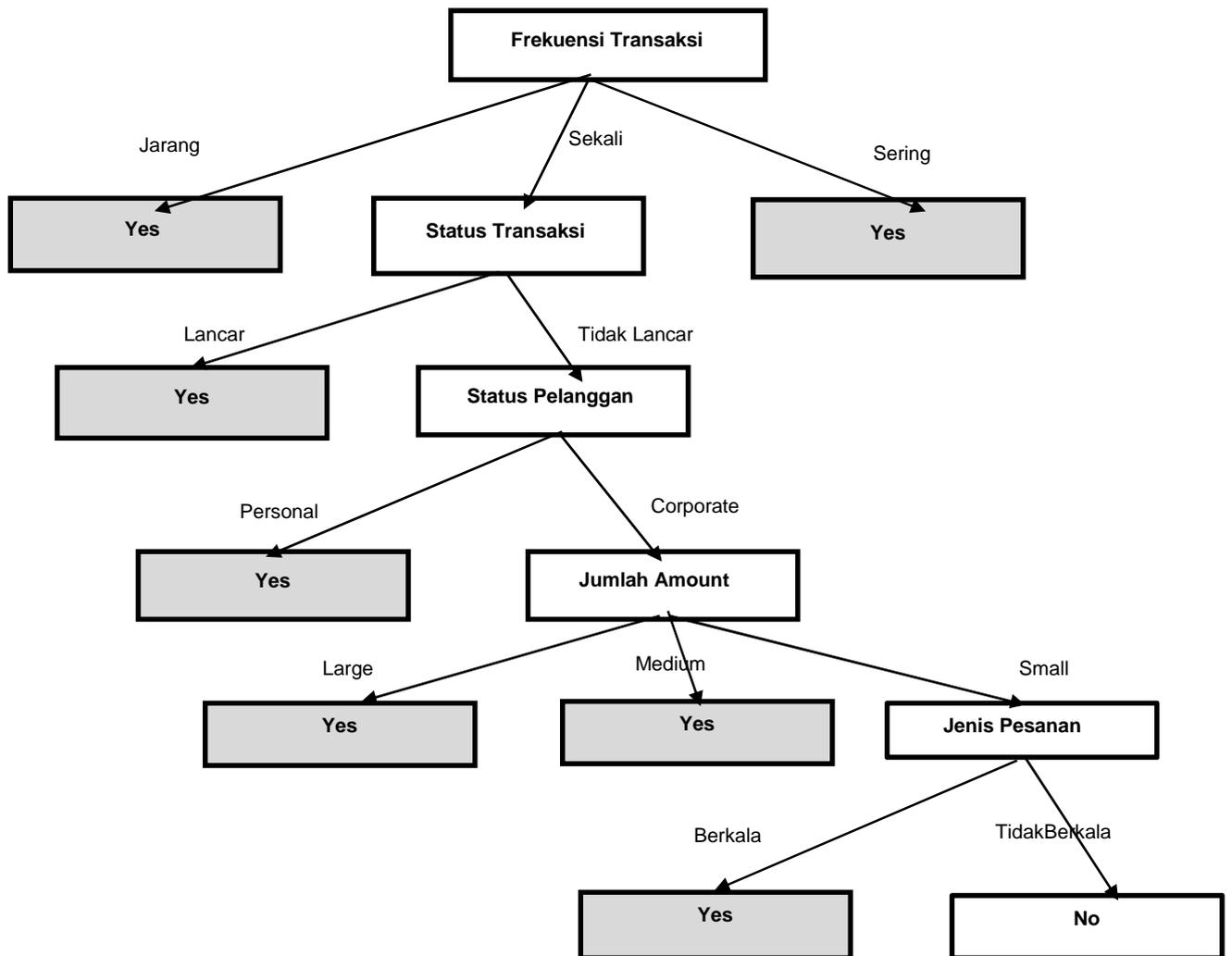
Pada penelitian ini akan dibuat *decision tree* untuk menentukan pelanggan mana yang memenuhi syarat sebagai member atau tidak dengan melihat atribut kode proyek, status pelanggan, jumlah *amount*, jenis pesanan, status transaksi, frekuensi transaksi, dan *Yes/No*. Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada dan dicari *entropy*-nya.

Untuk analisa ini, penulis menggunakan bahasa algoritma C4.5 disebabkan karena bahasa algoritma ini merupakan satu-satunya bahasa algoritma bisa menerapkan hasil *entropy* dan *gain* ke dalam pohon keputusan / *decision tree*. Dengan kata lain, Algoritma C4.5 dan *decision tree* merupakan dua model yang tak terpisahkan, karena untuk membangun sebuah *decision tree*, dibutuhkan algoritma C4.5.

Berikut ini adalah contoh dari analisis data mengenai langkah-langkah dalam pembentukan *decision tree* dengan menggunakan algoritma C4.5

1. Menghitung jumlah kasus, jumlah kasus untuk keputusan *Yes*, jumlah kasus untuk keputusan *No*, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut kode proyek, status pelanggan, jumlah *amount*, jenis pesanan, status transaksi, frekuensi transaksi, dan *Yes/No*. Setelah itu lakukan perhitungan *Gain* untuk masing-masing atribut.
2. Dari hasil perhitungan dapat diketahui bahwa atribut dengan *gain* tertinggi adalah *Frekuensi Transaksi*, yaitu sebesar 0.18011 Dengan demikian *Frekuensi Transaksi* dapat dijadikan *node akar*. Ada 3 nilai atribut dari *Frekuensi Transaksi* yaitu *Sekali*, *Jarang*, dan *Sering* dari ke-3 atribut itu *Jarang* dan *Sering* sudah mengklasifikasikan menjadi *Yes* sedangkan *Sekali* masih diperlukan perhitungan lagi.
3. Perhitungan dilanjutkan dengan Menghitung jumlah kasus, jumlah kasus untuk keputusan *Yes* dan jumlah kasus untuk keputusan *No*, dan *entropy* dari semua kasus dan kasus yang dibagi ke dalam *Status Pelanggan*, *Jumlah Amount*, *Jenis Pesanan*, dan *Status Transaksi* yang dapat menjadi *node akar* dari nilai atribut *Sekali*. Setelah itu, dilakukan perhitungan *gain* untuk tiap-tiap atribut hasil perhitungan selanjutnya.

4. Kemudian dari hasil perhitungan dapat diketahui bahwa atribut dengan *gain* tertinggi yaitu *Status Transaksi* yaitu sebesar 0,3681. Dengan demikian *Status Transaksi* dapat menjadi *node cabang* dari atribut *Sekali*. Dari atribut *Status Transaksi* terdapat 2 atribut yaitu *Lancar* dan *Tidak Lancar*. *Lancar* sudah mengklasifikasikan dirinya menjadi *Yes* sedangkan *Tidak Lancar* belum terklasifikasi sehingga masih memerlukan perhitungan selanjutnya.
5. Perhitungan dilanjutkan dan didapatkan informasi bahwa atribut dengan *gain* tertinggi yaitu *Status Pelanggan* yaitu sebesar 0.518791. Dengan demikian *Status Pelanggan* dapat menjadi *node cabang* dari nilai atribut *Tidak Lancar*. Dari atribut *Status Pelanggan* terdapat 2 atribut yaitu *Personal* dan *Corporate*. *Personal* sudah mengklasifikasikan dirinya menjadi *Yes* sedangkan *Corporate* belum terklasifikasi sehingga masih memerlukan perhitungan selanjutnya.
6. Berdasarkan perhitungan lanjutan yang dilakukan dapat diketahui bahwa atribut dengan *gain* tertinggi yaitu *Jumlah Amount* yaitu sebesar 0.666578. Dengan demikian *Jumlah Amount* dapat menjadi *node cabang* dari nilai atribut *Corporate*. Dari atribut *Jumlah Amount* terdapat 3 atribut yaitu *Small*, *Medium*, dan *Large*. *Medium* dan *Large* sudah mengklasifikasikan dirinya menjadi *Yes* sedangkan *Small* belum terklasifikasi. Namun, karena tersisa 2 atribut terakhir, maka tidak lagi diperlukan perhitungan lebih lanjut. Oleh karena itu, *Jenis Pesanan* menjadi *node cabang* dari nilai atribut *Small* dengan *Berkala* terklasifikasi *Yes* dan *Tidak Berkala* terklasifikasi *No*. Dari hasil perhitungan tersebut dapat digambarkan *decision tree* seperti gambar di bawah ini.



Gambar 7. Hasil Akhir *Decision Tree*

Setelah *decision tree* terbentuk maka akan dihasilkan sejumlah *rule* dalam *tree* tersebut. *Rule* yang dapat terbentuk dari *tree* adalah sebagai berikut :

1. “*IF Frekuensi Transaksi = Jarang THEN Yes/No = Yes*”
2. “*IF Frekuensi Transaksi = Sering THEN Yes/No = Yes*”
3. “*IF Frekuensi Transaksi = Sekali and Status Transaksi = Lancar THEN Yes/No = Yes*”
4. “*IF Frekuensi Transaksi = Sekali and Status Transaksi =Tidak Lancar and Status Pelanggan = Personal THEN Yes/No = Yes*”
5. “*IF Frekuensi Transaksi = Sekali and Status Transaksi =Tidak Lancar and Status Pelanggan = Corporate and Jumlah Amount = Large THEN Yes/No = Yes*”
6. “*IF Frekuensi Transaksi = Sekali and Status Transaksi =Tidak Lancar and Status Pelanggan = Corporate and Jumlah Amount = Medium THEN Yes/No = Yes*”
7. “*IF Frekuensi Transaksi = Sekali and Status Transaksi =Tidak Lancar and Status Pelanggan = Corporate and Jumlah Amount = Small and Jenis Pesanan = Berkala THEN Yes/No = Yes*”
8. “*IF Frekuensi Transaksi = Sekali and Status Transaksi =Tidak Lancar and Status Pelanggan = Corporate and Jumlah Amount = Small and Jenis Pesanan = Tidak Berkala THEN Yes/No = No*”

4.3. Implikasi Penelitian

Penelitian ini memiliki implikasi yang berdasarkan dari perhitungan *gain* dan *entropy* yang diperoleh dari hasil pengolahan data, pihak manajemen atau perusahaan sudah bisa menentukan pelanggan mana saja yang memiliki status pelanggan aktif ataupun tidak aktif. Setelah perusahaan mengetahui pelanggan mana yang aktif dan tidak aktif, ada beberapa tindakan yang mungkin dapat membantu perusahaan untuk mempermudah penggolongan pelanggan sesuai status pelanggan dan bahkan dapat mempertahankan hubungan kerja sama antara pelanggan dan perusahaan. Langkah – langkah tersebut sebagai berikut :

1. Melakukan pemilahan terlebih dahulu mana saja pelanggan yang memiliki status pelanggan aktif dan tidak aktif.
2. Untuk pelanggan yang memiliki status aktif, perusahaan harus menjaga hubungan baik dengan pelanggan sampai proyek yang

berlangsung selesai seperti konsistensi jadwal proyek.

3. Untuk pelanggan yang memiliki status tidak aktif, perusahaan melakukan konfirmasi terhadap pelanggan tersebut apakah pelanggan ingin tetap melakukan kerja sama dengan perusahaan atau tidak. Jika masih ingin tetap melakukan kerja sama atau ingin melakukan proyek baru, maka status pelanggan tersebut akan dirubah menjadi aktif. Namun, jika tidak ingin melakukan proyek baru, maka pelanggan tersebut membership-nya dinonaktifkan.

5. SIMPULAN

Berdasarkan hasil yang didapatkan dari tahap analisis, maka ada beberapa hal yang menjadi simpulan dalam penelitian ini, antara lain :

1. Algoritma C4.5 memiliki keakurasian yang cukup tinggi daripada RapidMiner dalam menyelesaikan proses perhitungan untuk menghasilkan *decision tree*.
2. RapidMiner lebih cepat dalam menghasilkan *decision tree* dibandingkan algoritma C4.5 karena algoritma C4.5 harus melakukan perhitungan secara manual. Namun, RapidMiner tidak seakurat algoritma C4.5 karena data yang dianalisis kurang bisa diterima tergantung pada jumlah data yang diinput.
3. Waktu yang diperlukan dalam proses analisis sangat bervariasi tergantung dari banyaknya data yang harus diolah terlebih dahulu.
4. Kesuksesan dalam melakukan pengujian data dengan menghasilkan *decision tree* dapat membantu perusahaan atau supervisor dalam memilih keputusan dengan resiko yang minimum.

6. REKOMENDASI

Pada penelitian ini, analisis yang telah dilakukan masih belum sempurna, masih terdapat beberapa kekurangan sehingga memerlukan penyempurnaan lebih lanjut. Oleh karena itu, dalam rangka penyempurnaan yang lebih baik, penulis memberikan beberapa saran sebagai berikut :

1. Sebelum melakukan analisis, pihak yang akan melakukan analisis terlebih dahulu mengenal, mempelajari, memahami dan mengerti metode-metode analisis yang ada, sehingga hasil analisis yang diperoleh dapat

- memberikan manfaat secara optimal dalam jangka waktu yang lebih lama.
2. Diharapkan pada pengembangan selanjutnya dapat diuji dengan algoritma lain untuk diketahui algoritma mana yang lebih efektif.
 3. Pada pengembangan selanjutnya diharapkan menggunakan perangkat lunak yang lebih modern.

6. DAFTAR PUSTAKA

- [1] Dua, S. & Xian Du (2011), *Data Mining and Machine Learning in Cybersecurity. USA : Taylor & Francis Group*. ISBN-13: 978-1-4398-3943-0.
- [2] Gelar Nurcahya (2013), *Analisis Dan Implementasi Decision Tree Untuk Klasifikasi Data Konsumen Telemarketing Untuk Deposito Pada Bank Menggunakan AlgoritmaC4.5*
- [4] Gorunescu, Florin (2011), *Concept, Model and Technique*, Springer-Verlag Heidelberg
- [5] Han, Jiawei and Kamber, Micheline (2011). *Data Mining Concepts and Techniques, 3rd edition*. Morgan Kaufman.
- [6] Kusriani dan Luthfi. E. Taufiq (2009). *Algoritma Data Mining*. Yogyakarta: Andi.
- [7] Larose, Daniel (2006), *Data Mining Methods and Models*, John Willey & Son, Inc
- [8] Laudon, Kenneth C., Jane P.Laudon, and Ahmed Elragal (2012), *Management Information Systems : Managing The Digital Firm*, Edisi ke-12, Prentice Hall.
- [9] MacLennan, J. dkk. (2009). *Data Mining with Microsoft SQL SERVER 2008*. Wiley: USA.
- [10] Maimon, Oded, Lior Rokarch (2010), *Data Mining and Knowledge Discovery Handbook*, Second Edition, Springer
- [11] McLeod Jr., Raymond, dan George Schell (2009), *Management Information Systems*, Edisi ke-9, New Jersey: Pearson Prentice Hall.
- [12] O'Brien, James A (2010), *Introduction to Information Systems: Essentials for the e-Business Enterprise*, Edisi ke-15, MCGraw-Hill/Irwin.
- [13] RapidMiner Reference
<http://en.wikipedia.org/wiki/RapidMiner>.
- [14] Turban, E., Aronson, J., E., and Liang, T., (2009). *Decision Support Systems and Intelligent Systems, 7th Ed, jilid 1*, Andi Offset, Yogyakarta.
- [15] Vercellis, Carlo (2009), *Business Intelligence Data Mining and Optimization for Decision Making*. (1st edition), United Kingdom : John Wiley & Sons Ltd.
- [16] Turban, Efraim., R. Kelly Rainer, Jr, dan Richard E. Potter (2010), *Introduction to Information Technology*, Edisi ke-3, New 17] Witten, Ian H, Eibe Frank, Mark A. Hall (2011), *Data Mining Practical Machine Learning Tools and Techniques*, Elsevier
- [17] Witten, I. H., Frank, E., Hall, M. A. (2011), *Data Mining Practical Machine Learning Tools and Techniques (3rd ed)*, USA : Elsevier