



ANALISA KEPUASAN PELANGGAN MENGGUNAKAN ALGORITMA C4.5 PADA PT. CAHAYA INDOTAMA ENGINEERING

Candra Naya¹, Arif Siswandi²

Program Studi Teknik Informatika Fakultas Teknik Universitas Pelita Bangsa

¹candranaya@pelitabangsa.ac.id, ²arifsiswandi@pelitabangsa.ac.id

Abstrak

Dalam penelitian ini kepuasan pelanggan menjadi salah satu penilaian terhadap produk dan jasa untuk memenuhi kebutuhan pelanggan. Salah satu hal yang penting dalam pengelolaan pelanggan adalah bagaimana suatu perusahaan dapat mempertahankan pelanggan yang dimilikinya. Kepuasan pelanggan merupakan perasaan senang atau kecewa seorang yang merupakan hasil dari analisa kepuasan pelanggan melalui kuesioner para pelanggan. Maka, metode klasifikasi mampu menemukan model yang mampu memprediksi dan menentukan kelas dari kepuasan pelanggan dengan data training yang telah diperoleh, Algoritma C4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Kelebihan ini misalnya dapat mengolah data *numeric* dan *diskret*, dapat menangani nilai atribut yang hilang, yang menghasilkan aturan-aturan mudah diinterpretasikan dan tercepat diantara algoritma yang lainnya untuk memprediksi kepuasan pelanggan dengan hasil penelitian menggunakan metode C4.5 termasuk algoritma yang akurat untuk memprediksi karena hasil akurasi menggunakan *RapidMiner* menyatakan tingkat *accuracy*, *precision*, dan *recall* C4.5 memiliki tingkat *accuracy* 86,11%, *precision* 89,02% dan *recall* 92,40%.

Kata kunci : *Data mining*, Algoritma C4.5, Klasifikasi, Kepuasan pelanggan

Abstract

In this research customer satisfaction becomes one of the assessments of products and services to meet customers is how a company can retain its customer need. One of important things in managing customers is how a company can retain its customers. Customer satisfaction is a feeling of pleasure or disappointment for a person which is the result of customer satisfaction analysis through a customer questionnaire. So, the classification method is able to find a model that is able to predict and determine the class of customer satisfaction with the training data that has been obtained. The C4.5 Algorithm is a data classification algorithm with a well-known and preferred decision tree technique because it has advantages. These advantages, for example, can process numeric and discrete data, can handle missing attribute values, which results in rules that are easy to interpret and the fastest among other algorithms. to predict customer satisfaction with the results of the study using the C4.5 method, including an accurate algorithm for predicting because the accuracy results using RapidMiner states the level of accuracy,

precision, and recall C4.5 has accuracy level of 86,11 %, precision 89,02% and recall 92,40%.

Keywords : Data mining, C4.5 Algorithm, Classification, customer satisfaction.

1. Pendahuluan

PT. Cahaya Indotama Engineering merupakan sebuah perusahaan swasta yang bergerak di bidang jasa penyewaan alat berat yang beralamat di Jl. Jababeka V Blok F No. 4, Kawasan Industri Jababeka 1, Cikarang, Bekasi 17550, Jawa Barat yang berfokus pada alat jenis *Tower Crane, Passenger & Material, Concrete Placing Boom* dan *Concrete Pump*. Pada awal tahun 2000 kontraktor yang ada di Indonesia lebih cenderung untuk menyewa alat untuk digunakan di proyeknya daripada membeli alat karena tingkat efisiensinya. Dengan menyewa alat, kontraktor tidak harus mengeluarkan biaya investasi yang besar untuk modal awal pembelian alat mereka dan juga mereka tidak harus mengeluarkan biaya untuk perawatan alat sehingga menghemat biaya pengeluarannya. PT. Cahaya Indotama Engineering menyediakan layanan peralatan lengkap mulai dari mobilisasi, instalasi, operasi, pemeliharaan,

pembongkaran dan berakhir dengan demobilisasi peralatan. PT. Cahaya Indotama Engineering sampai saat ini adalah bergerak dalam bidang sewa alat berat.

Populasi manusia dan kemajuan teknologi pembangunan infrastruktur di Indonesia yang semakin berkembang. Pembangunan pada berbagai sektor seperti gedung apartemen, rumah sakit, hotel, jembatan, pembangkit listrik, pabrik dan pembangunan lainnya terus berkembang pada berbagai wilayah di Indonesia.

Industri jasa sewa alat berat semakin banyak di Indonesia, PT. Cahaya Indotama Engineering sebagai suatu perusahaan swasta tentu saja berupaya untuk mendapatkan keuntungan serta terus berusaha memberikan pelayanan yang memuaskan bagi para pelanggannya. Namun tidak hanya PT. Cahaya Indotama Engineering sendiri yang bergerak di bidang jasa sewa alat berat maka hal ini akan menimbulkan persaingan antar perusahaan yang bergerak dibidang jasa rental alat berat ini.

Semakin pesatnya kemajuan dalam hal pendidikan dan perekonomian yang lebih sejahtera serta perkembangan ilmu pengetahuan dan teknologi, maka kesadaran masyarakat akan unsur-unsur pelayanan yang dapat diberikan oleh perusahaan semakin meningkat. Pada saat ini kesadaran konsumen akan pentingnya nilai kulaitas pelayanan (*Service Quality*) yang diberikan oleh perusahaan baik dalam bentuk jasa maupun dalam bentuk barang semakin meningkat. Dari beberapa pengalaman menunjukkan bahwa atas pemberian suatu kualitas jasa / pelayanan tertentu akan menimbulkan penilaian yang berbeda dari setiap konsumen, karena tergantung dari bagaimana konsumen mengharapkan kualitas jasa / pelayanan tersebut.

Perkembangan dunia usaha yang semakin pesat ini menyebabkan perusahaan harus menghadapi persaingan yang ketat perusahaan pada umumnya menginginkan pelanggan yang dimilikinya dapat dipertahankan selama-lamanya. Untuk mewujudkan hal itu bukanlah sesuatu yang mudah diiklim persaingan usaha yang sangat ketat saat ini mengingat adanya perubahan-perubahan yang cepat dapat terjadi setiap saat seperti perubahan pada diri pelanggan, pesaing maupun perubahan kondisi secara luas yang selalu dinamis. Hal ini menuntut para pengambil kebijakan untuk mengembangkan suatu strategi yang mampu mencapai sasaran pertumbuhan penjualan, peningkatan porsi pasar perusahaan, serta pencapaian kemampuan sebagai dasar pertumbuhan berkelanjutan.

Seorang konsumen memungkinkan mengalami berbagai tingkat kepuasan yaitu bila produk tidak sesuai dengan harapan setelah dikonsumsi, maka konsumen tersebut akan merasa tidak puas. Namun bila terjadi sebaliknya yaitu produk sesuai dengan harapannya, maka konsumen merasakan puas sehingga suatu saat mengkonsumsi kembali produk tersebut. Konsumen yang puas adalah konsumen yang akan berbagai kepuasan dengan produsen.

Dalam penelitian ini hasil survey kepuasan pelanggan tersebut juga akan dianalisis menggunakan pendekatan *data mining* dengan algoritma C4.5 dikarenakan algoritma ini secara umum lebih cepat, mudah diinterpretasikan dan dipahami pengguna. Disamping itu, berdasarkan beberapa penelitian

menunjukkan bahwa performansi lebih konsisten dibandingkan beberapa algoritma klasifikasi jenis pohon keputusan lainnya.

Algoritma C4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Kelebihan ini misalnya dapat mengolah data *numeric* dan *diskret*, dapat menangani nilai atribut yang hilang, yang menghasilkan aturan-aturan mudah diinterpretasikan dan tercepat diantara algoritma yang lainnya. Keakuratan prediksi yaitu kemampuan model untuk dapat memprediksi label kelas terhadap data-data yang baru ataupun data yang belum diketahui sebelumnya dengan baik.

Penelitian ini merupakan replikasi dari penelitian terdahulu yaitu penelitian yang dilakukan oleh [3] Melaporkan bahwa penelitiannya yang berjudul “Penerapan Algoritma C4.5 Pada Kepuasan Pelanggan Perum DAMRI” Dari hasil pengujian algoritma C4.5 dalam memprediksi kepuasan pelanggan perum DAMRI atas 90 sample data pelanggan yang diuji dalam penelitian ini, menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi yang cukup tinggi yaitu sebesar 93%. Selanjutnya [4] melaporkan bahwa penelitiannya menggunakan Algoritma C4.5 mendapatkan hasil yang bagus. Dari hasil uji coba terhadap 100 data, peneliti mendapatkan tingkat akurasi dari hasil prediksi kelulusan terhadap data testing sebesar 87.5%.

2. Landasan Pemikiran

2.1. Tinjauan Pustaka

Pada penelitian yang berjudul “Penerapan Algoritma C4.5 Pada Kepuasan Pelanggan Perum DAMRI” Dari hasil pengujian algoritma C4.5 dalam memprediksi kepuasan pelanggan perum DAMRI atas 90 sample data pelanggan yang diuji dalam penelitian ini, menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi yang cukup tinggi yaitu sebesar 93%. [3]

Penelitian lain dengan judul “Implementasi Data mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa” Dari hasil uji coba terhadap 100 data, peneliti mendapatkan tingkat akurasi dari hasil prediksi kelulusan terhadap data testing sebesar 87.5%.

Selanjutnya penelitian yang berjudul “Penerapan Metode Klasifikasi Data mining Untuk Prediksi Kelulusan Tepat Waktu”. Metode yang digunakan dalam penelitian ini adalah algoritma C4.5, Naive Bayes, dan Neural Network untuk memperkirakan kelulusan tepat waktu mahasiswa dengan melihat pengaruh dari IMK dan IPK. Dari penelitian ini menunjukkan bahwa algoritma terbaik adalah algoritma yang paling tinggi tingkat *accuracy* pada model klasifikasi yaitu C4.5 dan *Neural Network* dengan tingkat *accuracy* 100% sedangkan *Naive Bayes* 99.8878%. Hasil *data mining* dari algoritma terpilih dalam penelitian ini menggunakan C4.5, *interface* dirancang menggunakan *java engine* yang dapat menampilkan prediksi kelulusan tepat waktu beserta jumlah kelulusan tepat waktu setiap Program Studi.

Jurnal dengan judul “Penerapan *Data mining* Klasifikasi Pola Nasabah Menggunakan Algoritma

C4.5 Pada Bank BRI Batang”. Metode yang digunakan dalam penelitian ini adalah Decision Tree untuk melakukan klasifikasi data nasabah dari Bank BRI Batang yang menjadi bahan acuan untuk menganalisa pola nasabah pemohon kredit. Pemohon kredit termasuk dalam kategori lancar atau macet. Dalam penelitian ini, decision tree menggunakan bahasa pemrograman java. Kemudian hasil akurasi dari aplikasi yang telah diimplementasikan akan dibandingkan dengan hasil yang menggunakan software rapidminer. Sehingga diperoleh akurasi dengan decision tree sebesar 89,5%.

Penelitian yang berjudul “Penerapan Algoritma Klasifikasi Data mining C4.5 Pada Dataset Cuaca Wilayah Bekasi”. Metode yang digunakan adalah C4.5 yang bertujuan untuk mendapatkan pola klasifikasi cuaca. Berdasarkan hasil pengujian algoritma C4.5 yang menggunakan 10-fold validation dan dibuktikan dengan pembuatan aplikasi web untuk pengujian sehingga menghasilkan nilai yang akurasinya sebesar 88,89%.

2.2. Landasan Teori

2.2.1. Kualitas Pelayanan

Pelayanan merupakan salah satu unsur yang sangat penting dalam menciptakan kepuasan konsumen. Agar harapan konsumen terpenuhi, perusahaan harus memberikan pelayanan yang berkualitas. Kualitas dapat diartikan sebagai pengukuran seberapa baik tingkat pelayanan yang diberikan dan sesuai dengan harapan konsumen, jadi dengan kata lain memberikan pelayanan berkualitas berarti menyesuaikan diri dengan harapan konsumen. Ini merupakan salah satu faktor keberhasilan dalam persaingan yang makin ketat. Pelayanan yang berkualitas adalah orientasi semua sumber daya manusia dalam suatu perusahaan terhadap kepuasan pelanggan [8].

Definisi kualitas jasa adalah:

“Suatu sifat atau ciri yang membedakan nilai dari suatu barang atau jasa dengan nilai dari barang atau jasa yang lain yang sejenis” [8].

2.2.2. Kepuasan Konsumen

Kepuasan konsumen merupakan hal yang sangat penting dalam industri jasa. Karena dalam industri jasa, pelayanan yang dapat memuaskan konsumen akan memberikan imbalan yang menguntungkan, serta meningkatkan daya saing perusahaan. [9]

Definisi kepuasan pelanggan adalah:

“*Satisfaction is a person's feelings of pleasure or isappointment resulting from comparing a product's perceived performance (or outcome) in relation to his or her expectations*”. [9]

Secara umum kepuasan konsumen dan ketidakpuasan konsumen merupakan hasil dari perbedaan antara harapan dengan kinerja yang dirasakan oleh konsumen, Atau dengan kata lain ada dua kemungkinan yang akan terjadi, yaitu:

1. Kinerja yang dirasakan konsumen lebih besar dari yang diharapkan, artinya konsumen merasa puas dengan kualitas pelayanan yang diberikan oleh perusahaan

2. Kinerja yang dirasakan konsumen lebih kecil dari yang diharapkan, artinya konsumen tidak puas dengan kualitas pelayanan yang diberikan perusahaan.

2.2.3. Data mining

Data mining merupakan proses penggalian dan pertambangan pengetahuan dari sejumlah data yang besar, *database* atau *repository database* lainnya. Tujuan utama penambangan data ini untuk menemukan pengetahuan baru yang tersembunyi dari database tersebut [15].

Data mining adalah suatu rangkaian dari proses kemudian dapat dipisah-pisah menjadi beberapa tahapan. Tahapan-tahapan yang ada dalam *data mining* bersifat interaktif terhadap pengguna yang terlibat langsung dengan perantara knowledge base. Tahap-tahap dalam *data mining* antara lain :

1. Pembersihan Data

Tahap pembersihan data dilakukan untuk membuang data yang tidak konsisten dan *noise*. Selain itu, terdapat atribut data yang tidak sesuai dengan hipotesis *data mining* yang ada. Pembersihan data dapat mempengaruhi kinerja dari sistem *data mining* karena data yang diolah akan berkurang jumlah dan kompleksitasnya.

2. Integrasi Data

Integrasi data digunakan untuk menggabungkan data dari beberapa sumber karena dapat terjadi data yang dibutuhkan dapat berasal dari beberapa database atau file task. Tahap ini dilakukan pada atribut-atribut yang unik seperti nama, jenis produk, dan nomor pelanggan. Untuk menghasilkan data yang tepat dan tidak menyimpang maka harus dilakukan dengan cermat pada tahap ini.

3. Transformasi Data

Transformasi data dilakukan dengan mengubah data menjadi bentuk atau format yang sesuai. Sebagai contoh beberapa teknik dasar seperti analisis asosiasi dan klustering hanya dapat menerima input data kategorikal. Karena data yang berupa angka numerik perlu dipecah menjadi beberapa interval. Proses tersebut yang dinamakan binning. Transformasi dan pemilihan data ini menentukan ketepatan hasil dari *data mining* karena ada beberapa karakteristik dari teknik-teknik yang ada pada *data mining* tertentu bergantung dengan tahap ini.

4. Aplikasi Teknik *Data mining*

Tahapan aplikasi teknik *data mining* adalah bagian dari salah satu proses *data mining*. Sehingga harus diperhatikan bahwa teknik-teknik yang ada tidak selamanya dapat mencukupi untuk melaksanakan *data mining* tertentu.

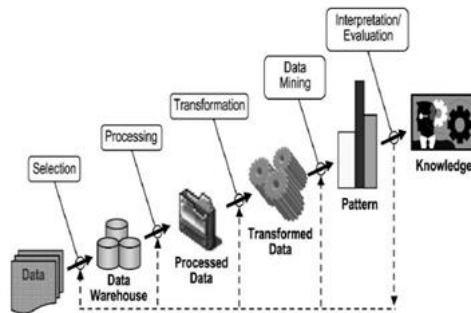
5. Evaluasi Pola yang Ditemukan

Tahap evaluasi pola yang ditemukan digunakan untuk menemukan pola-pola yang dengan ciri khas maupun prediksi yang bernilai. Apabila hasil yang ada tidak cocok dengan hipotesis yang ada maka terdapat cara lain yang dapat dilakukan.

6. Presentasi Pola yang Ditemukan

Selanjutnya tahap presentasi pola yang ditemukan digunakan untuk menghasilkan tindakan atau

langkah yang harus dilakukan dari analisis yang diperoleh dengan bentuk pengetahuan yang dapat dipahami semua orang. Dalam presentasi ini visualisasi membantu menampilkan hasil *data mining*.



Gambar1. Proses *Data Mining*

2.2.4. Metode CRISP-DM

CRISP-DM (Cross-Industry Standard Process Model for Data Mining) adalah standarisasi *data mining* yang disusun oleh tiga penggagas *data mining* market. Yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Kemudian dikembangkan pada berbagai workshops (antara 1997 – 1999). Lebih dari 300 organisasi yang berkontribusi dalam proses modeling ini dan akhirnya *CRISP-DM 1.0* dipublikasikan pada 1999. Proses *data mining* berdasarkan *CRISP-DM* terdiri dari enam fase, yaitu : [10]

3. Fase Pemahaman Bisnis (*Business understanding Phase*)
 - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
 - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.
4. Fase Pemahaman Data (*Data Understanding Phase*)
 - a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.
5. Fase Pengolahan Data (*Data Preparation Phase*)
 - a. Siapkan dari data awal, kumpulan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
 - b. Pilih kasus dan variabel yang ingin di analisis dan yang sesuai analisa yang akan dilakukan.
 - c. Siapkan data awal sehingga siap untuk perangkat pemodelan.
6. Fase Pemodelan (*Modeling Phase*)
 - a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
 - b. Kalibrasi aturan model untuk mengoptimalkan hasil.

- c. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *data mining* yang sama.
- d. Jika diperlukan proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.

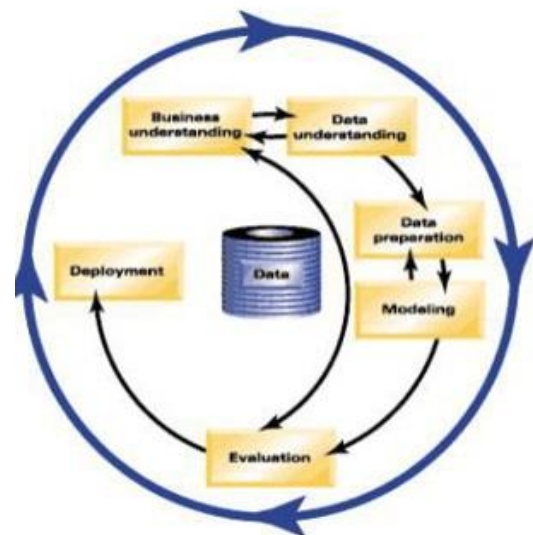
7. Fase Evaluasi (*Evaluation Phase*)

- a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarakan untuk digunakan.
- b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
- c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan berkaitan dengan penggunaan hasil dari *data mining*.

2. Fase Penyebaran (*Deployment Phase*)

3. Metode Penelitian

Metode yang digunakan untuk menyelesaikan penelitian ini adalah Penelitian eksperimen. Desain Eksperimen yang digunakan adalah *Cross Industry Standar Process Model for Data mining (CRISP-DM)* yaitu melakukan pengujian tingkat akurasi, precision dan recall dengan algoritma C4.5 dalam analisa kepuasan pelanggan PT. Cahaya Indotama Engineering dengan atribut nama, harga, waktu, mutu, pelayanan dan K3.



Gambar2. Proses *CRISP-DM*

3.1. *Business Understanding*

Pemahaman bisnis (*business understanding*), tahap pertama dalam proses *CRISP-DM* yang juga dapat disebut sebagai tahap pemahaman bisnis, berfokus pada pemahaman tujuan kebutuhan berdasarkan penilaian bisnis. Kemudian pemahaman tersebut diubah menjadi sebuah rencana awal *data mining* yang dirancang untuk mencapai tujuan, pada tahap ini diperlukan pemahaman tentang latar belakang dan tujuan pada proses bisnis untuk mendapatkan pelanggan.

3.2. Data Understanding

Pemahaman Data (*Data Understanding*) adalah tahap pengumpulan data. Data yang telah dikumpulkan dalam penelitian ini adalah data form kepuasan pelanggan yang didapat dari bagian staff Marketing PT. Cahaya Indotama Engineering berupa dokumen *Excel* sejumlah 108 record.

3.3. Data Reparation

Persiapan Data (*Data Reparation*) mencakup semua kegiatan untuk membangun dataset kuesioner kepuasan pelanggan yang akan diterapkan ke dalam alat pemodelan, dari data mentah awal berupa kuesioner dataset kepuasan pelanggan, pemilihan atribut dan selanjutnya akan melakukan proses *data mining*

3.4. Modelling

Pemodelan adalah fase yang secara langsung melibatkan teknik *data mining* yaitu dengan melakukan teknik *data mining* dan menentukan algoritma yang digunakan, *tools* yang digunakan dalam penelitian ini adalah *RapidMiner* dan menggunakan algoritma C4.5.

3.5. Evaluasi

Evaluasi adalah fase lanjutan terhadap tujuan *data mining*. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap pemodelan sesuai dengan sasaran yang ingin dicapai dalam tahap *business understanding*.

3.6. Deployment

Pada tahap ini pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna untuk mencapai *business understanding*

3.7. Metode Pengumpulan Data

Penelitian menggunakan beberapa metode pendukung guna mendapatkan data pada objek penelitian. Metode yang digunakan adalah:

a. Wawancara

Wawancara adalah teknik pengumpulan data yang dilakukan dengan menemui dan tanya jawab langsung antara peneliti dan narasumber. Pada metode ini peneliti langsung menanyakan beberapa hal tentang kepuasan pelanggan.

b. Kuesioner

Kuesioner yaitu teknik pengumpulan data yang dilakukan dengan cara memberikan seperangkat pertanyaan atau pernyataan kepada orang lain yang dijadikan respon dan untuk dijawabnya. Kuesioner ini dibuat dan dikumpulkan dari tahun 2017 sampai tahun 2020 di Cikarang, menyebarkan kuesioner untuk dibagikan kepada para penyewa alat berat yang ada di berbagai lokasi proyek. berupa pertanyaan-pertanyaan terbuka dan tertutup. Kuesioner berupa pertanyaan terbuka yaitu pertanyaan yang bersifat bebas dan digunakan untuk menyatakan alasan dan tanggapan atas pertanyaan tertutup. Sedangkan pertanyaan tertutup yaitu pertanyaan yang digunakan untuk mendapatkan data dari responden dalam objek penelitian dengan alternative

jawaban-jawaban yang disediakan oleh peneliti. Adapun atribut dalam kuisisioner tersebut meliputi harga, waktu, mutu, pelayanan dan K3.

c. Studi Pustaka

Studi pustaka merupakan kegiatan untuk mencari informasi yang relevan dengan topik atau masalah yang menjadi objek penelitian. Informasi tersebut dapat diperoleh dari buku-buku, karya ilmiah dan sumber-sumber lain yang berkaitan dengan kualitas pelayanan, kepuasan konsumen, *data mining*, dan algoritma C4.5.

4. Pembahasan

4.1. Langkah Perhitungan Algoritma

Data yang digunakan dalam penelitian ini dilakukan pada survey kepuasan pelanggan dari tahun 2017 sampai dengan tahun 2020 yang penulis peroleh dari bagian staff marketing PT. Cahaya Indotama Engineering dengan attribute harga, waktu, mutu, pelayanan dan K3. Form dari kuisisioner tersebut seperti gambar yang terlihat di bawah ini: dapun keterangan dari kuisisioner diatas yang di jelaskan pada tabel di bawah ini:

Gambar 3. Keterangan Kuesioner

Gambar 4. Form kuesioner kepuasan pelanggan

Selanjutnya adalah pemilihan atribut yang digunakan :

Tabel 1. Atribut Kuesioner

Atribut	Detail Penggunaan	
Nama	√	Variabel Bebas
Nama Perusahaan	X	No
Nama Proyek	X	No
Alamat	X	No
Spesifikasi Alat	X	No
Harga	√	Variabel Bebas
Waktu	√	Variabel Bebas
Mutu	√	Variabel Bebas
Pelayanan	√	Variabel Bebas
K3	√	Variabel Bebas
Kepuasan	√	Label Target

Tabel 2. Sample Dataset yang digunakan

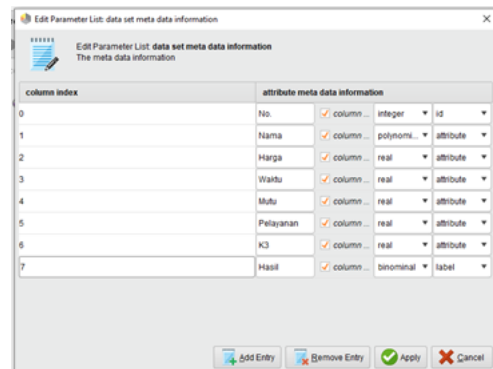
No.	Nama	Harga	Waktu	Mutu	Pelayanan	K3	Hasil
1	Agus	3	4	4	3,5	3	Puas
2	Andrie	3,5	4	4	4	4	Tidak Puas
3	Tony Suryana	3	3	4	3,25	3	Tidak Puas
4	Heri	2,5	3	3	4	4	Puas
...
106	Kiki	4,5	5	4,5	4,75	4	Puas
107	Ligya Moy	4,5	5	4,5	4,75	4	Puas
108	Antoine De Winter	2,5	3,25	4,5	3	3	Tidak Puas

Setelah kuesioner tersebut sudah terkumpul, kemudian mengolah data-data tersebut dengan menggunakan *Microsoft Office Excel* seperti yang terlihat dibawah :

4.2. Proses Data mining

Pada tahap selanjutnya data akan di proses menggunakan bantuan *tools RapidMiner* dan untuk pengujiannya menggunakan *cross validation* yang ada di dalam *RapidMiner*, maka data akan dilihat nilai

akurasi, pohon keputusan, dan rulenya. Berikut adalah Tabel Data yang akan di *import* ke dalam *RapidMiner*.



Gambar 5. Data yang Akan diimport kedalam *Rapidminer*

Pohon keputusan diatas menghasilkan rule yang akan diimplementasikan pada program. Rule yang dihasilkan sebagai berikut :

Tabel 4. 1. Rule dari Pohon Keputusan

Rules untuk hasil = Puas	
1.	IF nilai Harga>3.250 and nilai Mutu>3.250 and nilai K3>3.375 THEN hasil = Puas
3.	IF nilai Harga>3.250 and nilai Mutu>3.250 and nilai K3>3.375 and nilai Waktu ≤ 4.125 THEN hasil = Puas
4.	IF nilai Harga≤3.250 and nilai Pelayanan≤3.750 and nilai Waktu>3.875 THEN hasil = Puas IF nilai Harga≤3.250 and nilai Pelayanan>3.750 THEN hasil = Puas
Rules untuk hasil = Tidak Puas	
1.	IF nilai Harga>3.250 and nilai Mutu>3.250 and nilai K3>3.375 and nilai Waktu > 4.125 THEN hasil = Tidak Puas
2.	IF nilai Harga>3.250 and nilai Mutu≤3.250 THEN hasil = Tidak Puas
3.	IF nilai Harga≤3.250 and nilai Pelayanan≤3.750 and nilai Waktu≤3.875 THEN hasil = Tidak Puas

4.3. Hasil Pengujian

Pengujian ini dilakukan untuk mengetahui kinerja dari metode algoritma C4.5 dalam melakukan klasifikasi terhadap kelas yang telah ditentukan dalam uji coba ini. Pengujian ini dilakukan sebanyak tiga kali dengan *K-Fold Validation* yang berbeda pada operator *cross validation* berfungsi untuk membagi jumlah data *training* dan data *testing* pada data yang di uji. Pengujian terdiri dari :

1. *K-Fold Validation 10*
2. *K-Fold Validation 5*
3. *K-Fold Validation 3*

Algoritma C4.5 melakukan *training* terhadap data – data yang telah dibagi oleh *cross validation* menjadi dua kotak yaitu *training* dan *testing*. *Training*

terdiri dari *decision tree*, sedangkan *testing* terdiri dari *apply model* dan *performance*.

Berikut dijelaskan mengenai parameter operator yang digunakan pada model algoritma C4.5 sebagai berikut :

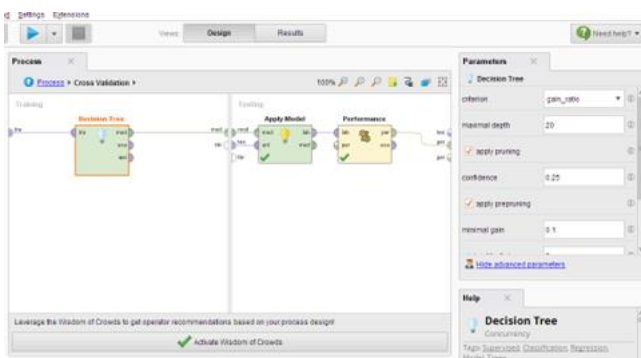
1. *Retreiving Data* adalah operator yang digunakan untuk mengimport dataset yang akan digunakan, pada penelitian ini data diimport dari file *excel*.
2. *Validation* operator yang digunakan dalam penelitian ini adalah *split validation*, *validation* ini hanya membagi total dari keseluruhan dataset menjadi data *training* dan data *testing*.
3. *Decision tree* adalah metode klasifikasi yang digunakan dalam penelitian ini.
4. *Apply model* adalah operator yang digunakan dalam penelitian ini yaitu algoritma C4.5
5. *Performance* adalah operator yang digunakan untuk mengukur performa akurasi model.

4.3.1. Pengujian Pertama

Pengujian pertama dilakukan dengan membagi sebanyak sepuluh bagian pada 108 data yang akan diuji. sepuluh bagian tersebut terdiri dari sembilan bagian data *training* dan satu bagian data *testing*. Dan dalam pengujian ini akan menghasilkan nilai akurasi, *precision*, *recall* dan *AUC (Area Under Curve)*.

Pada gambar dataset kepuasan pelanggan PT. Cahaya Indotama Engineering dihubungkan dengan operator *cross validation* yang didalamnya terdapat proses seperti gambar dibawah ini.

Algoritma C4.5 melakukan training terhadap data – data yang telah dibagi oleh *cross validation* menjadi dua kotak yaitu *training* dan *testing*. *Training* terdiri dari *decision tree*, sedangkan *testing* terdiri dari *apply model* dan *performance*. Hasil akurasi *confusion matrix* pada pengujian pertama sebagai berikut.



Gambar 6. Pengujian Algoritma C4.5 pada *RapidMiner*

Tabel 4. 2. *Confusion Matrix* Pengujian Pertama

	Actual	
	True Puas	True Tidak Puas
Pred. Puas	73	9
Pred. Tidak Puas	6	20

Jumlah *true positive (tp)* sebanyak 73 *record*, *false positive (fp)* sebanyak 6 *record*, jumlah *true negative (tn)* sebanyak 20 *record* dan jumlah *false negative (fn)* sebanyak 6 *record*. Akurasi yang dihasilkan dari pengujian pertama pada *RapidMiner* sebesar 86,11%

Berikut adalah perhitungan akurasi :

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \\
 &= \frac{73 + 20}{75 + 20 + 6 + 9} \times 100\% \\
 &= \frac{93}{108} \times 100\% \\
 &= 0,8611 \times 100\% \\
 &= 86,11\%
 \end{aligned}$$

Nilai *precision* dihitung dengan cara membagi jumlah data benar yang berniali positif (*True Positive*) dibagi dengan jumlah data benar yang bernilai positive (*True Positive*) dan data yang salah yang bernilai negative (*False Negative*)

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FN} \times 100\% \\
 &= \frac{73}{73 + 9} \times 100\% \\
 &= \frac{73}{82} \times 100\% \\
 &= 0,8902 \times 100\% \\
 &= 89,02\%
 \end{aligned}$$

Nilai *recall* dihitung dengan cara membagi jumlah data benar yang bernilai positif (*True Positive*) dibagi dengan jumlah data benar yang bernilai positif (*True Positive*) dan data salah yang bernilai positif (*False Positive*)

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP + FP} \times 100\% \\
 &= \frac{73}{73 + 6} \times 100\% \\
 &= \frac{73}{79} \times 100\% \\
 &= 0,9240 \times 100\% \\
 &= 92,40\%
 \end{aligned}$$

Hasil *AUC (Area Under Curve)* pada pengujian pertama sebagai berikut

Hasil *AUC* adalah 0,760 yang berarti pengujian pertama termasuk *fair classification*



Gambar 7. *AUC (Area Under Curve)* Pengujian Pertama

Tabel 4. 3. Evaluasi dan Validasi Pertama

K-Fold Validasi	Akurasi	Precision	Recall	Klasifikasi
10	86,11%	89,02%	92,40%	Menggunakan Algoritma C4.5

5. Penutup

Setelah melakukan analisis dan pengujian dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan bahwa :

1. Kepuasan pelanggan PT. Cahaya Indotama Engineering dapat diprediksi dan dievaluasi dengan memanfaatkan teknik *data mining* menggunakan algoritma *decision tree C4.5* untuk memprediksi (menentukan kelas) dari kepuasan pelanggan dengan data training yang telah diperoleh.
2. Dari tiga kali percobaan dan pengujian prediksi kepuasan pelanggan PT. Cahaya Indotama Engineering menggunakan aplikasi *rapidMiner* menggunakan metode *decision tree C4.5* diperoleh akurasi sebesar 87,96% dengan kriteria *good classification* menggunakan *confusion matrix*.

Daftar Pustaka

- [1] Resty Mardiana, "Faktor – Faktor Yang Mempengaruhi Kepuasan Pengguna Layanan Jasa Taksi Blue Bird," Jakarta, Universitas Gunadarma, 2009, pp. 3-6.
- [2] V. Mandasari and B. Tama, "Analisis Kepuasan Konsumen Terhadap Restoran Cepat Saji Melalui Pendekatan Data Mining: Studi Kasus XYZ," *J. Generic*, vol. 6, no. 1, pp. 25–28, Jan. 2011.
- [3] Ibnu Fathur Rochman, "Penerapan Algoritma C4.5 Pada Kepuasan Pelanggan Perum DAMRI," *Teknik Informatika*, Semarang, Juni 2015.
- [4] David Hartanto Kamagi, Seng Hansun, "Implementasi *Data mining* dengan Algoritma C4.5 untuk memprediksi Tingkat Kelulusan Mahasiswa," *Ultimatics*, vol. 6, no. 1, 2014.
- [5] Saefulloh & Moedjiono, 2013 "Penerapan Metode Klasifikasi *Data mining* Untuk Prediksi Kelulusan Tepat Waktu," *InfoSys Journal*, vol. 2, no. 1, pp. 41-54, Feb. 2013.
- [6] Ayu Rizqi Oktaviana, "Penerapan *Data Mining* Klasifikasi Pola Nasabah Menggunakan Algoritma C4.5 Pada Bank BRI Batang," *PSI Udinus*, Semarang, 2016.
- [7] Novandya, Oktria, "Penerapan Algoritma Klasifikasi *Data mining* C4.5 Pada Dataset Cuaca Wilayah Bekasi," *Format*, vol. 6 no. 2, pp. 98-106, 2017
- [8] Wirasmita, Rivai , Sitorus, Roland, Manurung, Brenggan, " Kamus Lengkap Ekonomi," *Bandung : Pionir Jaya* , 1999.
- [9] Kotler, Philip, " *Marketing Management: Analysis. Planning. Implementation. And Control, Ninth Edition*," *New Jersey : Prentice-Hall .Inc*, 1997.
- [10] J. Ardhyanti, M. Nugraha, Y. Kusumawati, S. Informasi, F. I. Komputer, and U. D. Nuswantoro, "Data Mining Dengan Metode *Clustering* Untuk Pengolahan Informasi Persediaan Obat Pada Puskesmas Pandanaran Semarang," *UDiNus Repos.*, vol. 1, no. 1, p. 1, 2014.
- [11] Gorunescu F, " *Data mining Concept Model Technique*," *Craiova, Romania: Springer*, vol. 12, 2011.
- [12] Kusrini, Emha Taufiq Luthfi, "Algoritma *Data mining*," *Yogyakarta: Andi Offset*, 2009.
- [13] Rafik Khairul Amin, Dra.Indwiarti, M.Si., Yuliant Sibaroni,S.Si., M.T., "Implementasi Klasifikasi *Decision Tree* Dengan Algoritma C4.5 Dalam Pengambilan Keputusan Permohonan Kredit Oleh Debitur," *e-Proceeding of Engineering*, vol.2, no.1, pp. 1768-1778, April 2015.
- [14] Lis Saumi Ramdhani, "Penerapan *Particle Swarm Optimization (PSO)* Untuk Seleksi Atribut Dalam Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis dengan Metode Algoritma C4.5," *Swabumi*, vol. 4, no 1, pp. 7-9, Maret 2016.
- [15] Erlin Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam *Data Mining* untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Konstruksi PT. Arupadhatu Adisesanti," *JOIN*, vol. 2, no. 1, pp. 37-40, Juni 2017.