

PENERAPAN ALGORITMA C 4.5 DALAM SISTEM PENDUKUNG KEPUTUSAN EVALUASI KINERJA FASILITATOR PAMSIMAS (STUDI KASUS DI KABUPATEN KAMPAR)

Zelvi Gustiana¹

¹Teknologi Informasi / ¹Universitas Dharmawangsa
¹Jl.K.L. Yos Sudarso No. 224 Glugur Kota Medan
Email: zelvigustiana@dharmawangsa.ac.id

Abstrak— Evaluasi kinerja merupakan metode dan proses penilaian dari pelaksanaan tugas seseorang atau sekelompok orang (fasilitator) dalam satu perusahaan atau organisasi. Evaluasi kerja fasilitator telah memiliki standar yang telah ditetapkan. Kualitas kinerja fasilitator diukur berdasarkan kriteria tersebut. Penelitian ini bertujuan membantu dan mempermudah penilai dalam mengevaluasi kinerja fasilitator PAMSIMAS dengan menerapkan *Data Mining* yaitu Algoritma C4.5 sebagai Sistem Pendukung Keputusan. Metode yang digunakan dalam penelitian ini adalah metode algoritma C4.5 dengan menggunakan atribut hasil evaluasi sebagai atribut label tujuan sebanyak 4 buah sub-atribut, yaitu: Sangat baik, baik, cukup dan buruk. Penelitian ini menghasilkan pohon keputusan dari 20 berupa value atribut tujuan hasil evaluasi, algoritma C4.5 dapat merepresentasikan sebanyak 13 pohon keputusan dari value atribut tujuan hasil evaluasi dan persentasi akurasi dari pengujian adalah sebesar 50%. Kesimpulan dari penelitian ini yaitu algoritma C4.5 membantu mengklasifikasikan hasil evaluasi kinerja berdasarkan atribut waktu, kelengkapan isian data, respon perbaikan dan tanggung jawab dalam tim. Penelitian dapat membentuk pohon keputusan yang dapat memberikan informasi hasil evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar.

Kata Kunci— Data Mining, Klasifikasi, Algoritma C4.5, Fasilitator, Evaluasi kinerja.

Abstract— Performance evaluation is a method and process of assessment of the implementation of the duties of a person or group of people (facilitator) in a company or organization. The facilitator's work evaluations have established standards. The performance quality of the facilitator is measured based on these criteria. This study aims to assist and facilitate the assessor in evaluating the performance of PAMSIMAS facilitators by applying Data Mining namely Algorithm C4.5 as Decision Support System. The method used in this research is C4.5 algorithm method by using attribute of result of evaluation as attribute of goal label as many as 4 sub-attributes, that is: Very good, good, enough and bad. This research yields a decision tree of 20 values of attributes of the objective of the evaluation result, the C4.5 algorithm can represent as many as 13 decision trees from the attribute value of the objective of the evaluation result and the accuracy percentage of the test is 50%. The conclusion of this research is C4.5 algorithm help classify performance evaluation result based on time attribute, completeness of data entry, response of improvement and responsibility in team. The research can form a decision tree that can provide information on the performance evaluation results of PAMSIMAS facilitators in Kampar regency.

Keywords— Data Mining, Classification, Algoritma C 4.5, Fasilitators, Performance evaluation.

I. PENDAHULUAN

Pada saat ini, dunia terus menerus mengalami perkembangan dan kemajuan pada berbagai aspek, termasuk kemajuan teknologi informasi berkembang dengan sangat pesat. Kebutuhan akan informasi yang akurat, cermat dan cepat sangat dibutuhkan oleh

masyarakat dalam menjalani kehidupan yang terus dilingkupi dengan perkembangan teknologi yang terus mengalami pembaharuan. Teknologi informasi yang memiliki kemampuan untuk mengumpulkan dan menyimpan berbagai data lebih baik dari kemampuan manual yang berasal dari mengalisis, meringkas dan mengekstrak pengetahuan dari data. Metode

tradisional yang ada pada saat ini, tidak mampu melakukan pengolahan data dalam kapasitas yang besar. Informasi yang berada di gudang data yang sudah dimiliki akan digali, digunakan dan dimanfaatkan oleh para pengambil keputusan untuk membantu dalam perihal pengambilan keputusan. Para ahli yang mengambil keputusan dibantu dengan gudang data yang sudah dimiliki dengan manfaatkannya untuk informasi yang berguna membantu mengambil keputusan, hal ini mendorong munculnya cabang ilmu yang disebut *Data Mining*. Dalam *data mining* terdapat beberapa algoritma untuk menyelesaikan masalah, salah satunya adalah Algoritma C 4.5. Algoritma C 4.5 termasuk salah satu Algoritma yang digunakan untuk pengklasifikasian data. Algoritma C 4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan termasuk salah satu metode klasifikasi dan prediksi yang terkenal, mampu mengubah fakta yang dipresentasikan menjadi aturan pada pohon keputusan yang dapat dimengerti dengan bahasa alami dan dapat diproses dalam bahasa basis data [1].

Sejauh ini, *Data Mining* dan Algoritma C 4.5 sudah digunakan untuk menyelesaikan banyak permasalahan yang menghasilkan solusi atau keputusan. Berdasarkan implementasi *data mining* dengan metode Algoritma C 4.5 mempunyai banyak kesamaan dengan masalah yang akan penulis angkat sebagai judul penelitian. Karena implementasi Algoritma C 4.5, merupakan pohon keputusan yang nantinya dalam kasus yang penulis angkat, dapat untuk menentukan bagaimana evaluasi kinerja tiap fasilitator PAMSIMAS Kabupaten Kampar. Karena untuk menentukan evaluasi kinerja Fasilitator Kabupaten Kampar, selama ini masih berjalan secara manual sehingga sering terjadi kekeliruan dalam penghitungan ataupun keterlambatan waktu karena masih menggunakan cara yang manual.

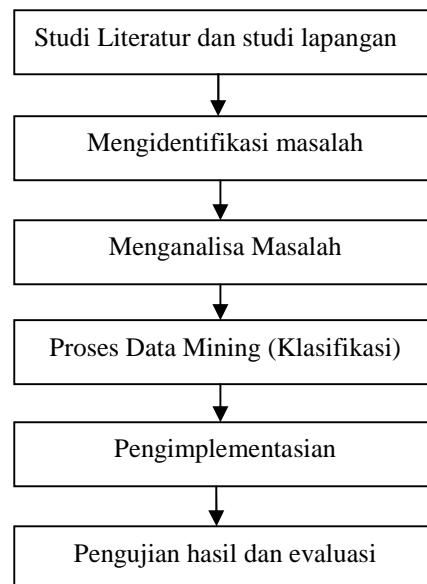
Khusus untuk fungsionalitas *data mining* menggunakan klasifikasi, beberapa penelitian telah dilakukan misalnya oleh Saefudin dan Sri Lestari pada tahun 2015 [2], mengenai Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5 dengan putusan penilaian kinerja yaitu istimewa, baik dan cukup. Dengan kriteria penilaian berdasarkan kemampuan teknis yang terdiri dari kualitas kerja, kuantitas kerja dan penguasaan kerja, kriteria kemampuan HAM terdiri dari kerjasama, kepemimpinan dan tanggung jawab, kriteria kepribadian terdiri dari integritas, disiplin dan semangat kerja. Hasil dari penelitian yang dilakukan yaitu mampu mengatasi permasalahan yang terjadi pada PT. Mulia Spindo Mills khususnya dapat mengubah penilaian kinerja karyawan yang sebelumnya dilakukan secara subyektif menjadi lebih objektif. Penelitian Data mining tentang klasifikasi salah satunya dilakukan oleh Dewi Sartika dan Dana

Indra Sensuse pada tahun 2017 [3] yaitu Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. Kesimpulan yang diperoleh pada penelitian ini adalah dari hasil perbandingan algoritma klasifikasi *nearest neighbour*, *naive bayes* dan *decision tree* yang digunakan pada studi kasus pengambilan keputusan pemilihan pola pakaian menyatakan bahwa algoritma klasifikasi *decision tree* merupakan algoritma klasifikasi yang memiliki tingkat akurasi paling tinggi dibandingkan algoritma klasifikasi *naive bayes* dan *nearest neighbour* yaitu mencapai 75.6% pada pengujian yang dilakukan dengan menggunakan mode pengujian *percentage split*.

II. METODOLOGI PENELITIAN

2.1 Kerangka Penelitian

Kerangka penelitian berguna untuk membuat tahapan-tahapan yang akan dilakukan dalam melakukan penelitian. Kerangka penelitian dapat membantu peneliti ketika peneliti tidak tahu apa yang akan dikerjakan setelah satu tahap pekerjaan sudah diselesaikan. Setiap tahapan dilakukan sesuai dengan yang telah direncanakan. Oleh karena itu semua tahapan pada kerangka penelitian penelitian ini berpengaruh pada tahapan selanjutnya..



Gbr 1. Kerangka Penelitian

Berdasarkan kerangka kerja pada Gambar 1 maka masing-masing langkahnya dapat diuraikan seperti berikut ini :

1. Studi Leteratur dan Studi Lapangan

Studi literatur dalam penelitian ini bersumber dari jurnal-jurnal dan buku yang berhubungan dengan penelitian yang dilakukan.

Sedangkan studi lapangan merupakan penelitian yang langsung penulis teliti pada objek, yaitu dengan cara observasi dan wawancara

2. Mengidentifikasi Masalah

Langkah yang dilakukan pada tahap ini adalah dengan melakukan peninjauan ulang terhadap permasalahan yang akan diteliti. Untuk menemukan tujuan dan rumusan masalah, serta penerapan Algoritma C 4.5 untuk menentukan evaluasi kinerja fasilitator.

3. Menganalisa Permasalahan

Langkah analisa masalah adalah langkah untuk dapat memahami masalah yang akan dijadikan penelitian. Teknik analisa yang dapat dilakukan dengan beberapa tahapan yang yaitu tahap *identify*, tahap *understand*, dan tahap *analyze*.

Pada tahap analisa ini, diusulkan metode baru untuk mengevaluasi kinerja fasilitator dengan cara menganalisis data dalam penentuan parameter atau atribut-atribut yang diperlukan pada saat melakukan perhitungan dengan algoritma C 4.5. Dengan tahapan yang dilakukan yaitu :

a. Masukan atau *input*

Merupakan kumpulan dari *Data set* yang sudah didapatkan pada hasil studi literatur dan studi lapangan. *Data set* berupa parameter atau atribut yang diperlukan untuk menentukan evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar

b. *Data Mining*

Mengolah *Data set* menjadi *Data Training* dan *Data Testing* yang akan diolah dengan metode klasifikasi Algoritma C 4.5.

c. Keluaran atau *Output*

Keluaran atau *output* merupakan hasil dari hasil pengolahan *Data Training*, yang akan menghasilkan pola terbaik atau *rule-rule* untuk penentuan evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar.

4. Proses *Data Mining* (Proses Klasifikasi)

Proses dengan teknik *Data Mining* yang diterapkan adalah :

a. Dengan memanfaatkan *Data set* yang sudah diolah menjadi *Data Training* untuk mendapatkan atribut sebagai pola yang tepat dalam penentuan evaluasi kinerja fasilitator, dengan menggunakan metode Algoritma C 4.5.

b. Pembentukan Pohon Keputusan, pendekatan ini sering menggunakan pohon keputusan berbasis jaringan saraf.

c. *Data Testing* digunakan untuk memperkirakan keakuratan aturan klasifikasi

d. Membandingkan hasil Keputusan pengolahan *Data Training* dengan hasil keputusan pengolahan *Data Training*

5. Pengimplementasian

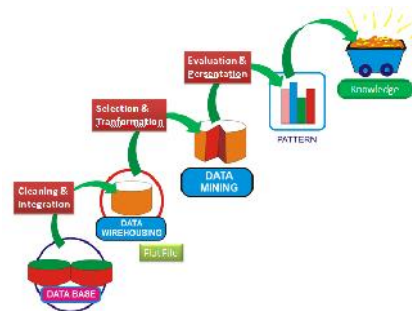
Implementasi dilakukan untuk mendapatkan pola terbaik dalam penelitian tentang penentuan evaluasi kinerja fasilitator, dengan menggunakan *Data Training* dan *Data Testing* yang sudah terbentuk.

6. Menguji Hasil dan Evaluasi

Jika proses implementasi dan pengujian sudah selesai dilakukan, maka setelah itu dilakukan proses evaluasi. Dengan cara menggunakan persentase akurat dan persentase kesalahan melalui perhitungan secara manual. Yang bisa dilihat dari proses menggunakan Aplikasi *Data Mining*, untuk evaluasi dilakukan saat implementasi dan pengujian telah dilakukan.

2.2 Data Mining

Data mining adalah proses mengekstraksi dan mengidentifikasi informasi yang bermanfaat serta pengetahuan yang terkait dari berbagai database besar dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning*. Istilah data mining memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya untuk menemukan, menggali atau menambang mengetahui dari data atau informasi yang sudah ada. *Data mining*, atau sering dikenal dengan nama *Knowledge Discovery in Database* (KDD). KDD merupakan serangkaian kegiatan seperti pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan pada set data berukuran besar.



Gbr 2. Proses Data Mining

Tahapan *data mining* adalah sebagai berikut:

1. Pembersihan data (*data cleaning*)
Pada tahap ini terjadi proses penghapusan atau menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan.
2. Integrasi data (*data integration*)
Pada tahap Integrasi data terjadi penggabungan data dari berbagai *database* ke dalam satu *database* baru.
3. Seleksi data (*data selection*)

Pada tahap ini, dilakukan pengambilan data dari database yang sesuai untuk dianalisis, dan dalam database sering kali tidak semuanya data dipakai, oleh karena itu terjadi penyeleksian.

4. Transformasi data (*data transformation*)

Pada transformasi data, data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

5. Proses *mining*

Proses mining merupakan proses utama yang dilakukan saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.

7. Presentasi pengetahuan (*knowledge presentation*)

Pada presentasi pengetahuan terjadi penggambaran dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

2.3 Algoritma C4.5

Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma yang lainnya. Kelebihan algoritma C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik. Dalam mengkonstruksi pohon, algoritma C4.5 membaca seluruh sampel *data training* dari *storage* dan memuatnya ke memori. Hal inilah yang menjadi salah satu kelemahan algoritma C4.5 dalam kategori "skalabilitas" adalah algoritma ini hanya dapat digunakan jika *data training* dapat disimpan secara keseluruhan dan pada waktu yang bersamaan di memori.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Perhitungan *Entropy* dan *Gain*
2. Pemilihan *Gain* tertinggi sebagai akar (*Node*)
3. Ulangi proses perhitungan *Entropy* dan *Gain* untuk mencari cabang sampai semua kasus pada cabang memiliki kelas yang sama yaitu pada saat semua variabel telah menjadi bagian dari pohon keputusan atau masing – masing variabel telah memiliki daun atau keputusan.
4. Membuat *Rule* berdasarkan pohon keputusan.

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

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

Keterangan:

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan berikut ini:

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

Keterangan:

S : Himpunan Kasus

n : Jumlah partisi S

p_i: Proporsi dari S_i terhadap S

2.4 Kinerja Karyawan

Kinerja karyawan merupakan suatu hal yang sangat penting dalam suatu organisasi untuk mencapai tujuannya, sehingga berbagai kebijakan harus dilakukan organisasi untuk kinerja karyawannya. Salah satu diantaranya adalah melalui penilaian kinerja. Handoko mendefinisikan penilaian prestasi kerja (*performance appraisal*) merupakan proses melalui nama organisasi-organisasi mengevaluasi atau menilai prestasi kerja karyawan. Suatu kinerja dikatakan tinggi jika suatu target kerja dapat diselesaikan pada waktu yang tepat atau tidak melampaui batas waktu yang disediakan. Dan sebaliknya, kinerja menjadi rendah jika diselesaikan melampaui batas waktu yang disediakan atau sama sekali tidak terselesaikan

Penilaian kinerja karyawan yang dilakukan secara objektif, tepat dan didokumentasikan secara baik memiliki manfaat menurunkan potensi penyimpangan yang dilakukan karyawan, sehingga kinerjanya diharapkan harus bertambah baik sesuai dengan kinerja yang dibutuhkan perusahaan. Disamping itu penilaian kinerja karyawan merupakan sarana untuk memperbaiki karyawan yang tidak melakukan tugasnya dengan baik dan membuat karyawan mengetahui posisi dan perannya dalam menciptakan tercapainya tujuan perusahaan. Hal tersebut bisa menambah motivasi dan keinginan karyawan untuk berkinerja lebih baik lagi, karena masing-masing dapat bekerja lebih baik dan benar sesuai dengan tanggung jawabnya.

III. HASIL PENELITIAN DAN PEMBAHASAN

3.1 Data

Data yang akan diambil untuk diteliti adalah data evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar selama bulan Desember 2017. Data tersebut merupakan data yang diambil dari kinerja fasilitator selama sebulan dan akan menghasilkan hasil evaluasi dari fasilitator tersebut. Data yang digunakan untuk pengolahan terdapat 52 buah data yang memiliki atribut lengkap dan telah memenuhi lima atribut yang akan diolah.

TABEL I
DATA FASILITATOR PAMSIMAS KAMPAR

No Absen	NAMA	POSISI	TIM	FM				HASIL EVALUASI
				WAKTU	Kelengkapan isian data	Respon Perbaikan	Tanggung Jawab dalam tim	
1	Ronni Olswan da	WSS	1	3	1	1	0	Cukup
2	M. Irfan	CD	1	2	2	1	1	Baik
3	Yusrizal	CD	1	1	2	1	2	Baik
4	Rastuti	CD	1	1	2	0	1	Cukup
5	Ika Marwardi	CD	1	0	1	1	2	Buruk
...
30	Fara Dube Putri	CD	7	1	1	1	2	Cukup
31	Cupitno Terulis	CD	7	2	1	0	1	Cukup
32	Zuliana	CD	7	3	1	0	2	Baik

TABEL II
ATRIBUT PENILAIAN

KET	INDIKATOR dan SUB INDIKATOR	POINT
WAKTU : (dilihat / mengacu pada hasil pemeriksaan FS dan Verifikasi DC)		
3	Selalu tepat waktu (setiap jum'at untuk kegiatan dalam minggu yang sama)	4
2	Sering tepat waktu (hari jum'at SETELAH 1 MINGGU KEGIATAN)	3
1	Kadang-kadang (hari jumat SETELAH 2 MINGGU KEGIATAN)	1
0	Menyerahkan di akhir bulan / SAAT PELAKSANAAN MISDAY (SETELAH 3 MINGGU KEGIATAN)	0
KELANGKAPAN ISIAN DATA: (dilihat/mengacu pada hasil pemeriksaan FS dan Verifikasi DC)		
2	Data selalu tidak lengkap,	0
1	Data kadang-kadang tidak lengkap,	1
0	Data selalu Lengkap,	3
RESPON PERBAIKAN PERBAIKAN DATA: (dilihat/mengacu pada hasil pemeriksaan FS dan Verifikasi DC)		
1	apabila ada data yg kurang/salah Respon Perbaikan cepat	1
0	apabila ada data yg kurang/salah Respon Perbaikan lambat,	0
TANGGUNG JAWAB DI TIM		
0	lempar tanggung jawab antar anggota tim	0
1	hanya mengurus diri sendiri/tidak membantu anggota tim	1
2	membantu melengkapi data FM lain di Timnya	2

3.2 Processing dan Transformasi Data

Tahan awal dari teknik klasifikasi Algoritma C 4.5 adalah *processing* dan transformasi data agar

data mentah yang akan diolah merupakan data yang memiliki atribut lengkap dan data menghasilkan pohon keputusan. Data yang akan diolah dalam penelitian ini hanya 30 buah data yang akan diproses. Data yang akan diambil tidak semua atribut dalam database evaluasi kinerja fasilitator, seperti atribut no absen, nama, posisi dan tim tidak dibutuhkan dalam penelitian ini, sehingga nantinya akan diolah lagi dan hanya menyisakan lima buah atribut, yaitu waktu, Kelengkapan isian data isi data, Respon Perbaikan, tanggung jawab dalam tim sebagai atribut tujuan. Dan dari kelima data itulah nanti akan didapatkan pohon keputusan dari evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar ini. Untuk Sampel data yang digunakan dan Sampel data transformasi dapat dilihat pada tabel dibawah ini :

TABEL III
SAMPEL DATA EVALUASI KINERJA FASILITATOR

No Absen	NAMA	POSISI	TIM	FM				HASIL EVALUASI
				WAKTU	Kelengkapan isian data	Respon Perbaikan	Tanggung Jawab dalam tim	
1	Ronni Olswan da	WSS	1	3	1	1	0	Cukup
2	M. Irfan	CD	1	2	2	1	1	Baik
3	Yusrizal	CD	1	1	2	1	2	Baik
4	Rastuti	CD	1	1	2	0	1	Cukup
5	Ika Marwardi	CD	1	0	1	1	2	Buruk
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
28	Afizan	CD	4	2	2	0	1	Baik
29	Yosi Harwanda	CD	5	2	1	1	0	Cukup
30	Zaki Riharvan	WSS	5	2	0	1	1	Cukup

Analisa sistem dalam penelitian ini adalah dengan menggunakan metode Algoritma C4.5. Algoritma C4.5 sendiri merupakan pengembangan dari Algoritma ID3. Dari tahap pengolahan data yang sudah di proses didapatkan hasil pengolahan data yang nantinya dapat dipresentasikan dalam bentuk *decision tree* atau pohon keputusan. Algoritma C4.5, dimulai dari proses memilih atribut dengan nilai gain tertinggi sebagai akar pohon, kemudian membuat cabang untuk tiap-tiap nilai, lalu membagi kasus dalam cabang, setelah itu ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

TABEL IV
SAMPEL DATA TRANSFORMASI

No	Waktu	Kelengkapan isian data	Respon Perbaikan	Tanggung jawab dalam tim	Hasil Evaluasi
1	selalu tepat waktu	kadang-kadang lengkap	cepat	lempar tanggung jawab	Cukup
2	sering tepat waktu	selalu lengkap	cepat	mengurus diri sendiri	Baik
3	kadang-kadang	selalu lengkap	cepat	membantu tim	Baik

No	Waktu	Kelengkapan isian data	Respon Perbaikan	Tanggung jawab dalam tim	Hasil Evaluasi
	g				
4	kadang-kadang	selalu lengkap	lambat	mengurus diri sendiri	Cukup
5	menyerahkan di akhir	kadang-kadang lengkap	cepat	membantu tim	Buruk
6	sering tepat waktu	selalu lengkap	cepat	membantu tim	Sangat Baik
7	selalu tepat waktu	kadang-kadang lengkap	lambat	membantu tim	Baik
8	kadang-kadang	kadang-kadang lengkap	cepat	membantu tim	Cukup
9	sering tepat waktu	tidak lengkap	lambat	membantu tim	Cukup
10	sering tepat waktu	kadang-kadang lengkap	lambat	mengurus diri sendiri	Cukup
11	kadang-kadang	tidak lengkap	cepat	membantu tim	Buruk
12	sering tepat waktu	selalu lengkap	lambat	mengurus diri sendiri	Baik
13	selalu tepat waktu	kadang-kadang lengkap	cepat	membantu tim	Baik
14	kadang-kadang	kadang-kadang lengkap	lambat	membantu tim	Buruk
15	kadang-kadang	selalu lengkap	cepat	lempar tanggung jawab	Cukup
16	kadang-kadang	selalu lengkap	cepat	membantu tim	Baik
17	sering tepat waktu	selalu lengkap	cepat	mengurus diri sendiri	Baik
18	sering tepat waktu	kadang-kadang lengkap	cepat	lempar tanggung jawab	Cukup
19	sering tepat waktu	tidak lengkap	cepat	lempar tanggung jawab	Buruk
20	sering tepat waktu	selalu lengkap	lambat	mengurus diri sendiri	Baik
21	sering tepat waktu	kadang-kadang lengkap	lambat	mengurus diri sendiri	Cukup
22	sering tepat waktu	tidak lengkap	cepat	membantu tim	Cukup
23	selalu tepat waktu	kadang-kadang lengkap	lambat	mengurus diri sendiri	Cukup

No	Waktu	Kelengkapan isian data	Respon Perbaikan	Tanggung jawab dalam tim	Hasil Evaluasi
24	sering tepat waktu	kadang-kadang lengkap	cepat	membantu tim	Baik
25	sering tepat waktu	kadang-kadang lengkap	Cepat	membantu tim	Baik
26	sering tepat waktu	kadang-kadang lengkap	Cepat	mengurus diri sendiri	Cukup
27	sering tepat waktu	selalu lengkap	Cepat	membantu tim	Sangat Baik
28	sering tepat waktu	selalu lengkap	lambat	mengurus diri sendiri	Baik
29	sering tepat waktu	kadang-kadang lengkap	Cepat	lempar tanggung jawab	Cukup
30	sering tepat waktu	tidak lengkap	Cepat	mengurus diri sendiri	Cukup

3.3 Pencarian Entropy Total

Proses pencarian entropy total dan gain dilakukan dengan cara mengelompokkan data dengan benar, kemudian menghitung data serta menggunakan rumus pencarian Entropy dan gain pada masing-masing atribut data, seperti rumus berikut ini:

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

Keterangan:

S : Himpunan Kasus

n : Jumlah partisi S

pi: Proporsi dari Si terhadap S

Pencarian entropy total untuk hasil evaluasi yaitu Sangat Baik, Baik, Cukup dan Buruk.

Jumlah Hasil Evaluasi Sangat Baik = 2

Jumlah Hasil Evaluasi Baik = 11

Jumlah Hasil Evaluasi Cukup = 13

Jumlah Hasil Evaluasi Buruk = 4

Entropy (X) total

$$= \left(-\frac{2}{30} * \log_2 \left(\frac{2}{30} \right) \right) + \left(-\frac{11}{30} * \log_2 \left(\frac{11}{30} \right) \right) + \left(-\frac{13}{30} * \log_2 \left(\frac{13}{30} \right) \right) + \left(-\frac{4}{30} * \log_2 \left(\frac{4}{30} \right) \right)$$

Jadi, Entropy(X) total = 1.7016

Lakukan perhitungan untuk jumlah kasus keseluruhan kasus yang ada, kemudian lakukan perhitungan untuk mencari nilai entropy dari setiap nilai atribut dan nilai gain setiap atribut.

3.4 Mencari Entropy Semua Atribut

Pencarian entropy semua atribut dilakukan untuk mendapatkan node awal yang menjadi akar untuk membuat pohon keputusan. Perhitungan dilakukan pada nilai entropy setiap atribut (Waktu, Kelengkapan isian data isian data, Respon Perbaikan dan Tanggung jawab dalam tim) setelah mendapatkan entropy total.

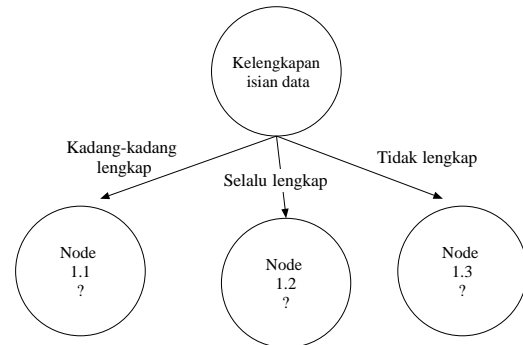
TABEL V
PERHITUNGAN ENTROPY TOTAL DAN GAIN TERTINGGI

Perhitungan Node 1							
Jumlah Kasus		Sangat Baik	Baik	Cukup	Buruk	Entropy	Gain
Total	30	2	11	13	4	1.701575	
Waktu							0.224783
Selalu tepat waktu	4	0	2	2	0	1	
Sering tepat waktu	18	2	7	8	1	1.633731	
Kadang-kadang	7	0	2	3	2	1.556657	
Menyerahkan di akhir	1	0	0	0	1	0	
Kelengkapan isian data							0.416242
Selalu Lengkap	11	2	7	2	0	1.309297	
Kadang-kadang Lengkap	14	0	4	8	2	1.378783	
Tidak lengkap	5	0	0	3	2	0.970951	
Respon Perbaikan							0.046848
Cepat	20	2	7	8	3	1.801609	
Lambat	10	0	4	5	1	1.360964	
Tanggung jawab							0.340662
Membantu tim	14	2	6	3	3	1.877387	
Mengurus diri sendiri	11	0	5	6	0	0.99403	
Lempar Tanggung jawab	5	0	0	4	1	0.721928	

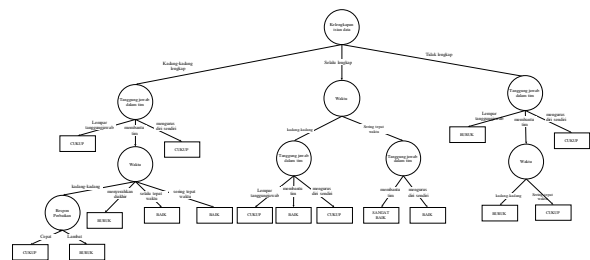
3.5 Menentukan Node akar

Dari hasil perhitungan pada tabel diatas, dapat diketahui bahwa nilai Gain tertinggi ada pada atribut “Kelengkapan isian data isian data” yaitu sebesar 0.416242. Sehingga atribut “Kelengkapan isian data

isian data” menjadi node akar. Pada atribut “Kelengkapan isian data isian data” terdapat 3 nilai atribut yaitu, kadang-kadang lengkap, selalu lengkap dan tidak lengkap, maka perlu dilakukan perhitungan untuk proses selanjutnya. Dari proses yang sudah dilakukan maka dapat dihasilkan pohon sementara seperti dibawah ini



Gbr 3. Node Akar



Gbr 4. Pohon Keputusan akhir

Setelah proses perhitungan selesai dan pohon keputusan didapatkan, maka didapatkan aturan (*rule*) yaitu :

- IF Kelengkapan isian data isian data “Kadang-kadang lengkap” AND Tanggung jawab dalam tim “Lempar Tanggung jawab” THEN Hasil evaluasi “CUKUP”
- IF Kelengkapan isian data isian data “Kadang-kadang lengkap” AND Tanggung jawab dalam tim “membantu tim” AND Waktu “Kadang-Kadang” AND Respon Perbaikan “Cepat” THEN Hasil evaluasi “CUKUP”
- IF Kelengkapan isian data isian data “Kadang-kadang lengkap” AND Tanggung jawab dalam tim “membantu tim” AND Waktu “Kadang-Kadang” AND Respon Perbaikan “Lambat” THEN Hasil evaluasi “BURUK”
- IF Kelengkapan isian data isian data “Kadang-kadang lengkap” AND Tanggung jawab dalam tim “membantu tim” AND Waktu “Menyerahkan di akhir” THEN Hasil evaluasi “BURUK”
- IF Kelengkapan isian data isian data “Kadang-kadang lengkap” AND Tanggung jawab dalam tim “membantu tim” AND Waktu “Selalu tepat waktu” THEN Hasil evaluasi “BAIK”

IF Kelengkapan isian data isian data “ Kadang-kadang lengkap” *AND* Tanggung jawab dalam tim “membantu tim” *AND* Waktu “Sering tepat waktu” *THEN* Hasil evaluasi “BAIK”

IF Kelengkapan isian data isian data “ Kadang-kadang lengkap” *AND* Tanggung jawab dalam tim “mengurus diri sendiri” *THEN* Hasil evaluasi “CUKUP”

IF Kelengkapan isian data isian data “Selalu lengkap” *AND* Waktu “kadang-kadang *AND* Tanggung jawab dalam tim “Lempar tanggung jawab” *THEN* Hasil evaluasi “CUKUP”

IF Kelengkapan isian data isian data “Selalu lengkap” *AND* Waktu “kadang-kadang” *AND* Tanggung jawab dalam tim “membantu tim” *THEN* Hasil evaluasi “BAIK”

IF Kelengkapan isian data isian data “Selalu lengkap” *AND* Waktu “kadang-kadang” *AND* Tanggung jawab dalam tim “mengurus diri sendiri” *THEN* Hasil evaluasi “CUKUP”

IF Kelengkapan isian data isian data “Selalu lengkap” *AND* Waktu “sering tepat waktu” *AND* Tanggung jawab dalam tim “membantu tim” *THEN* Hasil evaluasi “SANGAT BAIK”

IF Kelengkapan isian data isian data “Selalu lengkap” *AND* Waktu “sering tepat waktu” *AND* Tanggung jawab dalam tim “mengurus diri sendiri” *THEN* Hasil evaluasi “BAIK”

IF Kelengkapan isian data isian data “Tidak lengkap” *AND* Tanggung jawab dalam tim “Lempar tanggung jawab” *THEN* Hasil evaluasi “BURUK”

IF Kelengkapan isian data isian data “Tidak lengkap” *AND* Tanggung jawab dalam tim “Membantu tim” *AND* Waktu “Kadang-kadang” *THEN* Hasil evaluasi “BURUK”

IF Kelengkapan isian data isian data “Tidak lengkap” *AND* Tanggung jawab dalam tim “Membantu tim” *AND* Waktu “Sering tepat waktu” *THEN* Hasil evaluasi “BURUK”

IF Kelengkapan isian data isian data “Tidak lengkap” *AND* Tanggung jawab dalam tim “Mengurus diri sendiri” *THEN* Hasil evaluasi “CUKUP”

3.6 Pengujian

Pengujian yang dilakukan pada Pemrograman PHP sesuai dengan *rule* yang dihasilkan oleh perhitungan Algoritma C 4.5.

ATTURAN MATRIK	ATRIBUT	NILAI ATRIBUT	TOTAL KASUS	JUMLAH KASUS BAIK	JUMLAH KASUS BURUK	JUMLAH KASUS CUKUP	JUMLAH KASUS BAIK	PERSYIPKAN
kelengkapan_isian_data	Tidak	0	20	2	11	11	0,1000	0,5500
kelengkapan_isian_data	Kadang-kadang	1	8	2	2	2	0,2500	0,2500
kelengkapan_isian_data	Sering	2	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	3	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Selalu	4	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	5	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	6	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	7	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	8	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	9	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	10	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	11	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	12	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	13	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	14	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	15	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	16	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	17	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	18	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	19	2	2	0	0	1,0000	0,5000
kelengkapan_isian_data	Sangat	20	2	2	0	0	1,0000	0,5000

Gbr 5. Halaman Tampilan Perhitungan C 4.5

Dilihat dari gambar diatas, pada data-data yang sudah dimasukkan ke dalam database, “Kelengkapan_isian_data” merupakan atribut yang akan menjadi node akar dari pohon keputusan yang akan dihasilkan nanti. Dan untuk proses semuanya sudah disesuaikan dengan ketentuan rumus Algoritma C 4.5.

IV. PENUTUP

Berdasarkan hasil implementasi dan pengujian algoritma C4.5 pada evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar, maka dapat dikemukakan beberapa kesimpulan berikut:

1. *Data Mining* teknik klasifikasi menggunakan algoritma C4.5 dapat diterapkan pada evaluasi kinerja fasilitator PAMSIMAS Kabupaten Kampar, dengan mengikuti beberapa tahapan dalam KDD mulai dari *data cleaning*, *data selection*, *data transformation* dan *Data Mining* dengan menggunakan algoritma C4.5.
2. Perancangan pohon keputusan menghasilkan 16 buah *rule* yang dapat diterapkan pada evaluasi kinerja fasilitator yaitu Sangat Baik, Baik, Cukup dan Buruk. Atribut Kelengkapan isian data merupakan atribut akar (*root*) sehingga dapat diketahui bahwa kelengkapan isian data yang dilakukan oleh fasilitator merupakan atribut yang paling menentukan dalam proses pengklasifikasian menggunakan algoritma C4.5, selanjutnya atribut yang menjadi cabang adalah tanggung jawab dalam tim dan diikuti oleh waktu dan respon perbaikan.
3. Pengujian yang dilakukan pada Pemrograman PHP sesuai dengan *rule* yang dihasilkan oleh perhitungan Algoritma C 4.5 sehingga dapat membantu pihak terkait untuk meminimalisir kesalahan dalam pengimputan data evaluasi fasilitator.

UCAPAN TERIMA KASIH

Penulis mengucapkan banyak terima kasih kepada Risanuri Hidayat yang telah meluangkan waktu untuk membuat template ini.

REFERENSI

- [1] Siallagan T.F.P. Pencarian Nasabah dengan Menggunakan Data Mining dan Algoritma C 4.5 Koperasi Maduma. *Jurnal Teknik Informatika dan Sistem Informasi Volume 1 Nomor 3 Desember 2015* Subang
- [2] Dewi Sartika D and Sensuse D.I. Perbandingan Algoritma Klasifikasi *Naive Bayes*, *Nearest Neighbour*, dan *Decision Tree* pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. *Jatiji*, Vol. 1 No. 2 Maret 2017
- [3] Saefudin., Sri Lestari. Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5. *Jurnal PROTEKINFO* Vol.2 September 2015
- [4] Masripah, Siti. Komparasi Algoritma Klasifikasi *Data Mining* untuk Evaluasi Pemberian Kredit *Bina Insani Ict Journal*, Vol.3, No.1, Juni 2016, 187 – 193
- [5] Oktavia, Chaulina Alfianti., Rahmadwati., Purnomo Budi S. Analisis Kinerja Algoritma C4.5 Pada Sistem Pendukung Keputusan Penentuan Jenis Pelatihan. *Jurnal EECCIS* Vol. 9, No. 2, Desember 2015
- [6] Ridwan, Mujib., Hadi Suyono., M. Sarosa. Penerapan *Data Mining* Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma *Naive Bayes Classifier* *Jurnal EECCIS* Vol.7, No. 1, Juni 2013
- [7] Priatna J, Abdillah G, Uriawan. Sistem Pendukung Keputusan Pemberian Bantuan Buku Untuk Perpustakaan Di Kabupaten Bandung Barat Menggunakan Metode AHP Dan SAW. *Seminar Nasional Teknologi Informasi dan Komunikasi 2016 (SENTIKA 2016) Yogyakarta, 18-19 Maret 2016*
- [8] Nasari F. Penerapan Algoritma C4.5 Dalam Pemilihan Bidang Peminatan Program Studi Sistem Informasi Di Stmik Potensi Utama Medan. *Seminar Nasional Informatika 2014*.
- [9] Setiadi B. Penerapan Algoritma *Decision Tree* C4.5 Untuk Penilaian Rumah Tinggal. *INFO TEKNIK Volume 16 No. 2 Desember 2015 (195-206)*
- [10] Sumathi K., Kannan S and Nagarajan K. Data Mining: Analysis of student database using Classification Techniques. *International Journal of Computer Applications (0975 – 8887) Volume 141 – No.8, May 2016*
- [11] Gunadi G and Sensuse D.I. Penerapan Metode *Data Mining Market Basket Analysis* Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan *Frequent Pattern Growth* (Fp-Growth) : Studi Kasus Percetakan Pt. Gramedia. *Jurnal TELEMATIKA MKOM Vol.4 No.1, Maret 2012*
- [12] Santoso H., Hariyadi P.I and Prayitno. Data Mining Analisa Pola Pembelian Produk Dengan Menggunakan Metode Algoritma Apriori. *Seminar Nasional Teknologi Informasi Dan Multimedia 2016 Stmik Amikom Yogyakarta, 6-7 Februari 2016*
- [13] Jiandy R, Wiratama W. Y and Hansun S. Implementation of C4.5 and PAPI Kostick to Predict Students Potential as Organizatio Caretaker. *International Journal of Information Technology, Computer Science and Open Source* Vol.1, No.1, 2017
- [14] Priyadharsini and Thanamani A.S. An Overview of Knowledge Discovery Database and Data mining Techniques. *International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014*
- [15] Rikhi N. Data Mining and Knowledge Discovery in Database. *International Journal of Trends and Technology (IJETT) – Volume23 Number 2- May 2015*
- [16] Zulkifli A. Metode C45 Untuk Mengklarifikasi Pelanggan Perusahaan Telekomunikasi *Seluler. Riau Journal Of Computer Science* Vol.2/No.1/2016 : 65 – 76
- [17] Nurjannah N., Arifin Z., and Khairina D.M. Sistem Pendukung Keputusan Pembelian Sepeda Motor Dengan Metode *Weighted Product*. *Jurnal Informatika Mulawarman* Vol. 10 No. 2 September 2015
- [18] Situmorang H. Sistem Pendukung Keputusan Pemilihan Calon Peserta Olimpiade Sains Tingkat Kabupaten Langkat Pada Madrasah Aliyah Negeri (MAN) 2 Tanjung Pura Dengan menggunakan Metode Simple Additive Weighting (SAW). *Jurnal Times* , Vol. Iv No 2 : 24-30 , 2015