

# Analisis Kinerja Algoritma C4.5 Pada Sistem Pendukung Keputusan Penentuan Jenis Pelatihan

Chaulina Alfianti Oktavia, Rahmadwati, and Purnomo Budi S

**Abstract**—This study describes the application of the algorithm C4.5 on decision support systems to support trainees in PPTIK STIKI Malang in choosing the appropriate type of training. Decision support system based on several criteria derived from the data filled out by participants prior to register as a participant. Further analysis using an algorithm that is used to form a C4.5 decision tree. The decision tree is a method of classification and prediction that represent rules, the rule is then developed using RGFDT (Rule Generation From Decision Tree). Results of testing done by comparing the system with Weka and showed an accuracy of 86%.

**Index Terms**—Algorithm C4.5, Data Mining, Decision Support System, RGFDT

**Abstrak**—Penelitian ini menjelaskan penerapan algoritma C4.5 dan pengembangannya pada sistem pendukung keputusan di PPTIK STIKI Malang. Analisis lebih lanjut menggunakan algoritma C4.5 yang digunakan untuk membentuk pohon keputusan. Pohon keputusan adalah metode klasifikasi dan prediksi yang mewakili aturan. aturan tersebut kemudian dikembangkan dengan menggunakan RGFDT (*Rule Generation From Decision Tree*). Hasil pengujian dilakukan dengan membandingkan hasil sistem dengan Weka dan menunjukkan akurasi 86%.

**Kata Kunci**—Algorithm C4.5, Data Mining, Decision Support System, RGFDT

## I. PENDAHULUAN

UNTUK memenuhi kebutuhan dalam menguasai komputer, masyarakat menggunakan media dan jasa pelatihan yang disediakan oleh lembaga atau universitas tertentu. Salah satunya di Pusat Pelatihan Teknologi Informasi dan Komputer (PPTIK) Sekolah Tinggi Informatika dan Komputer Indonesia (STIKI) Malang.

Pusat Pelatihan Teknologi Informasi dan Komputer (PPTIK) merupakan unit penunjang yang dimiliki oleh STIKI Malang. PPTIK dikembangkan untuk memenuhi kebutuhan akan pendidikan serta pelatihan praktis bidang Teknologi Informasi dan Komputer yang dilaksanakan

Chaulina Alfianti Oktavia adalah Mahasiswa Program Magister Teknik Elektro, Fakultas Teknik, Universitas Brawijaya (UB) Malang-Indonesia (email: chaulina.1216@gmail.com).

Rahmadwati adalah dosen Jurusan Teknik Elektro Fakultas Teknik Universitas Brawijaya (korespondensi dengan penulis melalui email: rahma@ub.ac.id)

Purnomo Budi Santoso adalah Dosen Teknik Industri, Fakultas Teknik, Universitas Brawijaya, Malang-Indonesia (email: budiakademika@yahoo.com).

sesuai dengan standar mutu pelaksanaan pembelajaran STIKI Malang. *Join Program*, *In House Training* dan *Short Course* merupakan beberapa jenis pelatihan yang terdapat di PPTIK STIKI Malang. *Join program* merupakan pelatihan yang bekerjasama dengan sekolah menengah. Siswa dapat mengikuti jenis pelatihan yang sesuai dengan bidangnya. *In House Training* merupakan pelatihan yang bekerjasama dengan perusahaan/instansi negeri maupun swasta yang berguna untuk meningkatkan kemampuan pegawai dalam bidang teknologi informasi, sedangkan untuk *Short Course* merupakan pelatihan yang yang dirancang dengan waktu yang fleksibel.

Dalam sistem pendukung keputusan terdapat informasi yang perlu ditelaah lebih dalam lagi sehingga informasi yang dibutuhkan dalam pemecahan sebuah masalah dapat diselesaikan dengan baik[1]. Salah satu metode yang digunakan dalam sistem pendukung keputusan adalah metode klasifikasi. Metode klasifikasi merupakan sebuah proses untuk menemukan model yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui[2]. Beberapa metode klasifikasi diantaranya adalah pohon keputusan, Bayes dan Rules. Salah satu metode klasifikasi dalam penentuan keputusan adalah algoritma C4.5.

Algoritma C4.5 merupakan metode berbasis pohon keputusan. Dalam algoritma C4.5 pemilihan atribut dilakukan dengan menggunakan *Gain*, *Ratio*, dengan mencari nilai *Entropy*. Algoritma C4.5 sendiri menggunakan pendekatan induksi dimana dalam pendekatan ini, algoritma C4.5[3] membagi data berdasarkan kriteria yang dipilih untuk membuat sebuah pohon keputusan yang menggunakan pendekatan secara *top-down*. Berdasarkan analisa yang dilakukan Jose Augusto dalam penelitiannya, algoritma C4.5 mampu memberikan hasil yang efektif dalam mendukung suatu keputusan dengan kriteria yang dibuat secara *random*. Selain itu, alasan pemilihan algoritma C4.5 adalah, algoritma tersebut mampu menghasilkan sub sistem *model base* yang dapat digunakan untuk menunjang sistem pendukung keputusan. Untuk mendukung pengembangan algoritma C4.5, digunakan metode RGFDT (*Rules Generation From the Decision Tree*) untuk membangun *general rules* dari *rule set* yang dihasilkan dari algoritma C4.5[4].

Berdasarkan uraian di atas maka dalam penelitian ini digunakan teknik pengambilan keputusan menggunakan algoritma C4.5 sebagai sistem pendukung keputusan

pemilihan materi pelatihan di PPTIK STIKI Malang berdasarkan parameter yang telah ditentukan dalam menunjang suatu sistem pendukung keputusan. Parameter yang digunakan menjadi kriteria yang akan diimplementasikan menggunakan algoritma C4.5.

## II. LANDASAN TEORI

### A. Sistem Pendukung Keputusan

Menurut Mat dan Watson, Sistem Penunjang Keputusan (SPK) merupakan suatu sistem interaktif yang membantu pengambilan keputusan melalui penggunaan data dan model-model keputusan untuk memecahkan masalah-masalah yang sifatnya semi terstruktur dan tidak terstruktur. Sedangkan menurut Moore dan Chang, SPK adalah sistem yang dapat dikembangkan, mampu mendukung analisis data dan pemodelan keputusan, berorientasi pada perencanaan masa mendatang, serta tidak bisa direncanakan *interval* (periode) waktu pemakaiannya. Bonezek, Hosapple dan Whinston mendefinisikan SPK sebagai suatu sistem yang berbasis komputer yang terdiri dari 3 komponen yang berinteraksi satu dengan yang lainnya, yaitu[5]:

- *Language system*, adalah suatu mekanisme untuk menjembatani (*interface*) pemakai dan komponen lainnya.
- *Knowledge system*, adalah repositori pengetahuan yang berhubungan dengan masalah tertentu baik berupa data maupun prosedur.
- *Problem processing system*, adalah sebagai penghubung kedua komponen lainnya, berisi satu atau beberapa kemampuan manipulasi atau menyediakan masalah secara umum, yang diperlukan dalam pengambilan keputusan.

Karakteristik dari Sistem Pendukung Keputusan yang membedakan dari sistem informasi lainnya adalah[6]:

- SPK dirancang untuk membantu pengambil keputusan dalam memecahkan masalah yang sifatnya semi terstruktur ataupun tidak terstruktur.
- Dalam proses pengolahannya, SPK mengkombinasikan penggunaan model-model/teknik-teknik analisis dengan teknik pemasukan data konvensional serta fungsi-fungsi pencari/interogasi informasi.
- SPK dirancang sedemikian rupa sehingga dapat digunakan/dioperasikan dengan mudah oleh orang-orang yang tidak memiliki dasar kemampuan pengoperasian komputer yang tinggi. Oleh karena itu pendekatan yang digunakan biasanya model interaktif.
- SPK dirancang dengan menekankan pada aspek fleksibilitas serta kemampuan adaptasi yang tinggi. Sehingga mudah disesuaikan dengan berbagai perubahan lingkungan yang terjadi dan kebutuhan pemakai.

Sistem pendukung keputusan dapat memberikan berbagai manfaat atau keuntungan bagi pemakainya, antara lain:

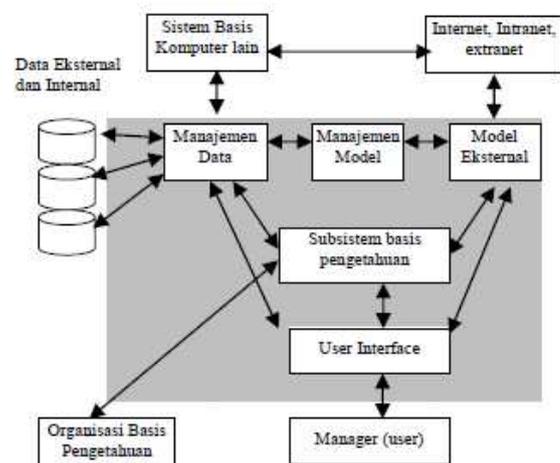
- Memperluas kemampuan pengambilan keputusan dalam memproses data/informasi bagi pemakainya.
- Membantu pengambilan keputusan dalam hal penghematan waktu yang dibutuhkan untuk

memecahkan masalah terutama berbagai masalah yang sangat kompleks dan tidak terstruktur.

- Dapat menghasilkan solusi dengan lebih cepat serta hasilnya dapat diandalkan.
- Walaupun suatu Sistem Pendukung Keputusan, mungkin saja tidak mampu memecahkan masalah yang dihadapi oleh pengambil keputusan, namun dapat menjadi stimulan bagi pengambil keputusan dalam memahami persoalannya, karena sistem pendukung keputusan mampu menyajikan berbagai alternatif.
- Dapat menyediakan bukti tambahan untuk memberikan bukti tambahan untuk memberikan pembenaran sehingga posisi pengambil keputusan

Efraim Turban, dalam bukunya *Decision support system and Intelligent System*, aplikasi sistem pendukung keputusan bisa dikomposisikan dengan subsistem berikut ini[7]:

Komponen-Komponen Sistem Pendukung Keputusan Menurut Turban [2], Sistem Pendukung Keputusan terdiri dari empat subsistem, yaitu:



Gambar 1 Komponen SPK

#### a. Subsistem Manajemen Data

Subsistem manajemen data terdiri dari elemen berikut ini:

- *DSSdatabase* adalah kumpulan data yang saling terkait yang diorganisasi untuk memenuhi kebutuhan dan struktur sebuah organisasi dan dapat digunakan oleh lebih dari satu orang untuk lebih dari satu aplikasi.
- Sistem Manajemen basis data adalah pembuatan pengaksesan, dan pembaharuan (*update*) oleh DBMS yang mempunyai fungsi utama sebagai tempat penyimpanan, mendapatkan kembali (*retrieval*) dan pengontrolan.
- Direktori merupakan sebuah katalog dari semua data di dalam basis data.
- *Query Facility*, yang menyediakan fasilitas akses data. Fungsi utamanya adalah untuk operasi seleksi dan manipulasi data dengan menggunakan model-model yang sesuai dari *model management*.

#### b. Subsistem Manajemen Model

Subsistem manajemen model terdiri atas elemen-elemen berisikan model-model seperti manajemen keuangan, statistik, ilmu manajemen yang bersifat kuantitatif yang memberikan kapabilitas analisis pada sebuah SPK. Model

Strategis digunakan untuk mendukung manajemen puncak untuk menjalankan tanggungjawab perencanaan strategis. Ketika model ini digunakan dalam suatu SPK, maka yang menjadi tujuan di dalamnya adalah untuk membantu pengambilan keputusan strategis yang sifatnya jangka panjang. Model Taktis digunakan terutama oleh manajemen madya untuk membantu mengalokasikan dan mengontrol sumber daya organisasi. Model Operasional digunakan untuk mendukung aktifitas kerja harian pada organisasi. Model Analitik digunakan untuk menganalisis data (untuk aplikasi sendiri), sebagai komponen dari model yang lebih besar, dan digunakan untuk menentukan variabel dan parameter model.

Sistem Manajemen Basis Model merupakan sistem *software* yang fungsi utamanya untuk membuat model dengan menggunakan bahasa pemrograman, alat SPK dan atau subrutin, dan balok pembangun lainnya; membangkitkan rutin baru dan laporan; pembaruan dan perubahan model; dan manipulasi model[8].

Direktori Model Peran direktori model sama dengan direktori basis data. Direktori model adalah katalog dari semua model dan perangkat lunak lainnya pada basis model. Direktori model tersebut berisi definisi model dan fungsi utamanya adalah menjawab pertanyaan tentang ketersediaan dan kapabilitas model.

Eksekusi Model, Integrasi, dan Prosesor Perintah Eksekusi model adalah proses mengontrol jalannya model saat ini. Integrasi model mencakup gabungan operasi beberapa model saat diperlukan atau menintegrasikan SPK dengan aplikasi lain. Sedangkan prosesor model digunakan untuk menerima dan menginterpretasikan instruksi-instruksi pemodelan.

c. Subsistem Dialog

Komponen dialog SPK adalah perangkat lunak dan perangkat keras yang menyediakan antarmuka untuk SPK. Istilah antarmuka pengguna mencakup semua aspek komunikasi antara satu pengguna dan SPK[9]. Cakupannya tidak hanya perangkat lunak dan perangkat keras, tapi juga faktor-faktor yang berkaitan dengan kemudahan pengguna, kemampuan untuk dapat diakses, dan interaksi manusia-mesin.

d. Subsistem Manajemen Knowledge

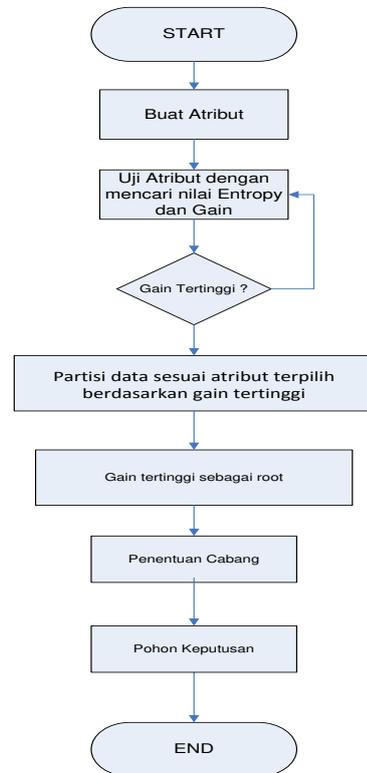
Banyak masalah tak terstruktur dan bahkan semi terstruktur yang sangat kompleks sehingga solusinya memerlukan keahlian. Oleh karena itu banyak SPK canggih yang dilengkapi dengan komponen manajemen *knowledge*. Komponen ini menyediakan keahlian untuk memecahkan beberapa aspek masalah dan memberikan pengetahuan yang dapat meningkatkan operasi komponen SPK lainnya. Berikut adalah modul-modul Sistem Pendukung Keputusan dengan menggunakan beberapa teknik penyelesaian yang ada bisa digunakan untuk memahami bagaimana sebuah Sistem Pendukung Keputusan diaplikasikan.

B. Algoritma C4.5

Algoritma C4.5 adalah pengembangan dari algoritma ID3. Oleh karena pengembangan tersebut algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan algoritma ID3. Hanya saja dalam algoritma C4.5 pemilihan atribut dilakukan dengan menggunakan *Gain Ratio* dengan persamaan[9]:

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

Dimana (S) adalah himpunan kasus, A adalah Atribut, n adalah Jumlah partisi atribut A, |Si| adalah Jumlah kasus pada partisi ke-i dan|S| adalah jumlah kasus dalam S.



Gambar 2.Flowchart Algoritma C4.5

Berdasarkan *flowchart* Gambar 2 dapat diketahui alur algoritma C4.5 yang digunakan. Pada persiapan awal ditentukan atribut yang digunakan kemudian melakukan uji atribut dengan mencari nilai *Gain* tertinggi berdasarkan perhitungan *entropy* dari masing-masing atribut[10]. Apabila ditemukan gain tertinggi maka gain tersebut akan menjadi *root* awal. Selanjutnya dilakukan penentuan cabang dengan cara yang sama dengan melihat *gain* tertinggi dari tiap hasil partisi.

III. METODE PENELITIAN

A. Analisis Masalah

Analisis permasalahan adalah pembahasan yang lebih detil atau pendalaman dari sub bab latar belakang didepan yang diperlukan untuk menemukan konsep solusi. Analisis permasalahan yang dibahas adalah mengenai penentuan jenis pelatihan yang tepat bagi peserta pelatihan di PPTIK STIKI Malang. Saat ini penentuan jenis pelatihan bagi peserta pelatihan di PPTIK STIKI Malang masih menggunakan cara manual, yaitu hanya dilihat dari beberapa parameter berdasarkan biodata pendaftaran peserta terutama berdasarkan keinginan atau minat peserta. Seringkali beberapa peserta pelatihan membutuhkan suatu saran dalam menentukan jenis pelatihan yang tepat agar peserta dapat dengan mudah mengikuti pelatihan dan menguasai bidang tersebut.

### B. Identifikasi Variabel

Dari Analisis permasalahan diatas dapat diidentifikasi variabel-variabel yang akan terlibat dalam solusi sebagai berikut:

- Variabel himpunan kasus
- Atribut
- Jumlah partisi atribut
- Jumlah kasus
- Nilai *entropy*
- *Gain*

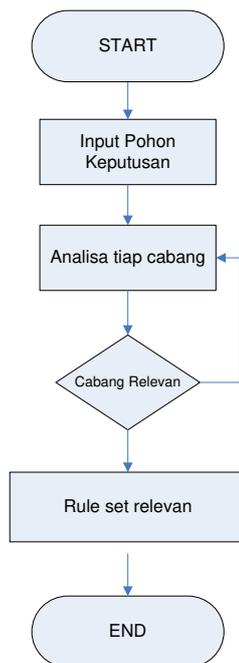
### C. Pengembangan Metode

Berikut ini merupakan metode yang relevan beserta sumbangannya terhadap pemecahan masalah.

- RGFDT (*Rules Generation From the Decision Tree*)

Merupakan sebuah algoritma yang efisien untuk mengkonversi pohon keputusan untuk seperangkat aturan tanpa kondisi yang tidak relevan.

- Algoritma RGFDT
  - **Input** : Pohon keputusan
  - *Let Branch = {Branch1, ..., BranchN};*
  - *For each branch Branch' in Branch Do*
  - **Output**: Rules tanpa kondisi yang tidak relevan;



Gambar 3. Flowchart RGFDT

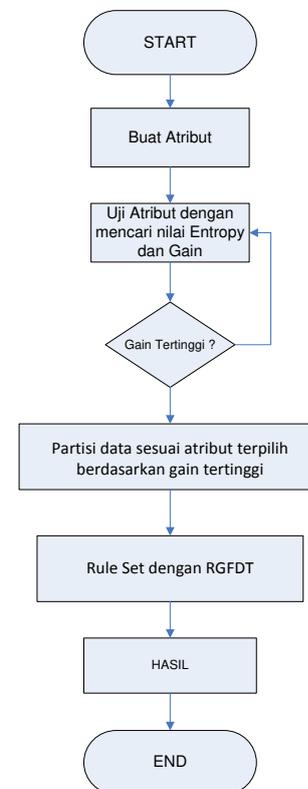
Flowchartgambar 3 merupakan alur dari metode RGFDT. Input berupa pohon keputusan selanjutnya dilakukan analisa tiap cabang, apabila terdapat kondisi yang tidak relevan maka dilakukan pemangkasan cabang. Sehingga menghasilkan *rules* baru tanpa kondisi yang tidak relevan.

### D. Konsep Solusi Terintegrasi

Setelah mengetahui kontribusi masing-masing metode yang relevan, maka untuk mengintegrasikan semua metode diatas menjadi solusi yang tepat dan terarah dalam memilih jenis pelatihan yang sesuai. Konsep solusi dijabarkan lebih lanjut menjadi konsep algoritma sebagai berikut.

Berdasarkan flowchart Gambar 4 dapat diketahui alur

algoritma C4.5 yang digunakan. Pada persiapan awal ditentukan atribut yang digunakan kemudian melakukan uji atribut dengan mencari nilai *Gain* tertinggi berdasarkan perhitungan *entropy* dari masing-masing atribut. Apabila ditemukan *gain* tertinggi maka selanjutnya dilakukan tahap partisi data sesuai atribut terpilih berdasarkan *gain* tertinggi, sebaliknya jika tidak ditemukan *gain* tertinggi maka proses diulangi sampai menemukan nilai *gain* tertinggi.



Gambar 4. Flowchart Algoritma C4.5 dengan metode RGFDT

Atribut yang digunakan adalah beberapa jenis kriteria data yang mempengaruhi hasil penentuan jenis pelatihan. Pada penelitian ini digunakan algoritma C4.5 untuk mendukung hasil penentuan jenis pelatihan berdasarkan data yang telah ada dan selanjutnya diolah untuk diperoleh nilai *entropy* dan *gain*. Selanjutnya dilakukan partisi data berdasarkan *gain* tertinggi. Dari nilai yang dihasilkan akan dibentuk suatu hasil pohon keputusan yang dapat digunakan sebagai acuan dalam mendukung sistem dalam menentukan jenis pelatihan yang tepat bagi peserta pelatihan di PPTIK STIKI Malang. Pohon keputusan tersebut kemudian dikonversi menjadi *rule set* menggunakan metode RGFDT (*Rule Generation From Decision Tree*). RGFDT merupakan metode pengembangan dari *decision tree* yang berfungsi untuk menyederhanakan rule yang dihasilkan dengan mengabaikan kondisi yang tidak relevan dan menghindari masalah kehilangan cabang.

## IV. HASIL DAN PEMBAHASAN

### A. Pengujian Algoritma C4.5

Pengujian dari sisi verifikasi dan validasi aplikasi sistem pendukung keputusan, pengujian dan pembahasan algoritma C4.5 diuraikan berikut ini.

Dari Tabel I diperoleh hasil sesuai sebanyak 38 pengujian dan 12 pengujian tidak sesuai. Dari pengujian pengukuran

validasi sistem dengan *tools* Weka diperoleh hasil validasi sebesar 76%.

TABLE I  
PEBANDINGAN HASIL SISTEM DENGAN PAKAR MENGGUNAKAN ALGORITMA C4.5

No	Skenario	Keterangan
1	Peserta 1	Tidak Sesuai
2	Peserta 2	Sesuai
3	Peserta 3	Sesuai
4	Peserta 4	Sesuai
5	Peserta 5	Sesuai
6	Peserta 6	Tidak Sesuai
7	Peserta 7	Sesuai
8	Peserta 8	Sesuai
9	Peserta 9	Sesuai
10	Peserta 10	Sesuai
11	Peserta 11	Sesuai
12	Peserta 12	Tidak Sesuai
13	Peserta 13	Sesuai
14	Peserta 14	Sesuai
15	Peserta 15	Sesuai
16	Peserta 16	Sesuai
17	Peserta 17	Tidak Sesuai
18	Peserta 18	Sesuai
19	Peserta 19	Sesuai
20	Peserta 20	Sesuai
21	Peserta 21	Tidak Sesuai
22	Peserta 22	Sesuai
23	Peserta 23	Sesuai
24	Peserta 24	Tidak Sesuai
25	Peserta 25	Sesuai
26	Peserta 26	Tidak Sesuai
27	Peserta 27	Sesuai
28	Peserta 28	Sesuai
29	Peserta 29	Sesuai
30	Peserta 30	Sesuai
31	Peserta 31	Tidak Sesuai
32	Peserta 32	Sesuai
33	Peserta 33	Sesuai
34	Peserta 34	Sesuai
35	Peserta 35	Sesuai
36	Peserta 36	Sesuai
37	Peserta 37	Sesuai
38	Peserta 38	Tidak Sesuai
39	Peserta 39	Sesuai
40	Peserta 40	Sesuai
41	Peserta 41	Sesuai
42	Peserta 42	Sesuai
43	Peserta 43	Tidak Sesuai
44	Peserta 44	Sesuai
45	Peserta 45	Sesuai
46	Peserta 46	Tidak Sesuai
47	Peserta 47	Sesuai
48	Peserta 48	Tidak Sesuai
49	Peserta 49	Sesuai
50	Peserta 50	Sesuai

```

Classifier output
=== Summary ===

Correctly Classified Instances      38      76  %
Incorrectly Classified Instances    12      24  %
Kappa statistic                    0.6716
Mean absolute error                 0.1455
Root mean squared error             0.2698
Relative absolute error             47.874 %
Root relative squared error         69.3571 %
Total Number of Instances          50
    
```

Gambar 5. Hasil Output Weka Classifier Pengujian 1

Pada Gambar 5 dapat disimpulkan bahwa data 50 peserta yang diuji menggunakan *tools* Weka berdasarkan algoritma c4.5, menghasilkan presentase sebesar 76% dimana 38 data dari 50 data terklasifikasi dengan benar.

B. Pengujian Algoritma C4.5 dengan metode RGFDT

Selanjutnya dilakukan pengujian algoritma dengan pengembangan menggunakan metode RGFDT ditampilkan pada tabel berikut :

TABLE II  
PEBANDINGAN HASIL SISTEM DENGAN PAKAR MENGGUNAKAN ALGORITMA C4.5 DAN METODE RGFDT

No	Skenario	Keterangan
1	Peserta 1	Sesuai
2	Peserta 2	Sesuai
3	Peserta 3	Sesuai
4	Peserta 4	Sesuai
5	Peserta 5	Sesuai
6	Peserta 6	Tidak Sesuai
7	Peserta 7	Sesuai
8	Peserta 8	Sesuai
9	Peserta 9	Sesuai
10	Peserta 10	Sesuai
11	Peserta 11	Sesuai
12	Peserta 12	Sesuai
13	Peserta 13	Sesuai
14	Peserta 14	Sesuai
15	Peserta 15	Sesuai
16	Peserta 16	Sesuai
17	Peserta 17	Tidak Sesuai
18	Peserta 18	Sesuai
19	Peserta 19	Sesuai
20	Peserta 20	Sesuai
21	Peserta 21	Sesuai
22	Peserta 22	Sesuai
23	Peserta 23	Sesuai
24	Peserta 24	Sesuai
25	Peserta 25	Sesuai
26	Peserta 26	Tidak Sesuai
27	Peserta 27	Sesuai
28	Peserta 28	Sesuai
29	Peserta 29	Sesuai
30	Peserta 30	Sesuai
31	Peserta 31	Tidak Sesuai
32	Peserta 32	Sesuai
33	Peserta 33	Sesuai
34	Peserta 34	Sesuai
35	Peserta 35	Sesuai
36	Peserta 36	Sesuai
37	Peserta 37	Sesuai
38	Peserta 38	Tidak Sesuai
39	Peserta 39	Sesuai
40	Peserta 40	Sesuai
41	Peserta 41	Sesuai
42	Peserta 42	Sesuai
43	Peserta 43	Sesuai
44	Peserta 23	Sesuai
45	Peserta 45	Sesuai
46	Peserta 46	Sesuai
47	Peserta 47	Sesuai
48	Peserta 48	Sesuai
49	Peserta 49	Tidak Sesuai
50	Peserta 50	Sesuai

```

=== Summary ===

Correctly Classified Instances      43      86  %
Incorrectly Classified Instances     7      14  %
Kappa statistic                    0.8025
Mean absolute error                 0.0916
Root mean squared error             0.214
Relative absolute error             31.1201 %
Root relative squared error         56.0072 %
Total Number of Instances          50
    
```

Gambar 6. Hasil Output Weka Classifier Pengujian 1

Tabel II merupakan pengujian algoritma C4.5 pada 50 data peserta setelah ditambahkan metode RGFDT. Dari Tabel II

diperoleh hasil sesuai sebanyak 43 pengujian dan 7 pengujian tidak sesuai. Dari pengujian pengukuran validasi sistem dengan *tools* Weka diperoleh hasil validasi sebesar 86%.

Pada Gambar6 dapat disimpulkan bahwa data 50 peserta yang diuji menggunakan *tools*Weka berdasarkan algoritma C4.5, menghasilkan presentase sebesar 86% dimana 43 data dari 50 data terklasifikasi dengan benar.

#### V. KESIMPULAN

Sistem pendukung keputusan telah berhasil dirancang menggunakan algoritma C4.5 dengan pengembangan metode RGFDT(*Rule Generation From the Decision Tree*) untuk membentuk *rule* yang dihasilkan dari pohon keputusan. Hasil pengujian terhadap algoritma C4.5 dan metode RGFDT menunjukkan prosentase kebenaran sebesar 86%.

#### DAFTAR PUSTAKA

- [1] Karaolis, Minas A., IEEE., Moutiris, Joseph A., Hadjipanayi Demetra., And Pattichis, Constantinos S., Senior Member., IEEE, 2010, *Assessment Of The Risk Factors Of Coronary Heart Events Based On Data Mining With Decision Trees*, IEEE Transactions On Information Technology In Biomedicine, Vol. 14, No. 3, May.
- [2] Turban. dkk., 2005, *Decision Support System and Intelligent System (Sistem Pendukung Keputusan dan Sistem Cerdas)*, Andi, Yogyakarta.
- [3] Chih-Chiang Wei, Jiing-Yun You, 2011, *C4.5 Classifier For Solving The Problem Of Water Resources Engineering*, Proceeding Of The International Conference On Advanced Science, Engineering And Information Technology, Isbn 978-983-42366-4-9, January.
- [4] Al-Hegami, Ahmed Sultan , 2007, *Classical And Incremental Classification In Data Mining Process*, IJCSNS International Journal Of Computer Science And Network Security, Vol.7 No.12, December
- [5] Ding-An Chiang,Wei Chen, Yi-Fanwang And Lain-Jinn Hwang.2001, *Rules Generation From the Decision Tree*, Department of Information Engineering Tamkang University, JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 17, 325-339
- [6] Henry.2011. *Pengambilan Pola Kelulusan Tepat Waktu Pada Mahasiswa Stmik Amikom Yogyakarta Menggunakan Data Mining Algoritma C4.5*.STMIK AMIKOM Yogyakarta
- [7] Jie,Fan.2007. *Application Of C4.5 Algorithm In Web-Based Learning Assessment System*. IEEE, 2007, Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong.
- [8] Kusrini, Emha, 2009, *Algoritma Data Mining*, Penerbit Andi, Yogyakarta
- [9] LI Rui, WEI Xian-mei ,YU Xue-wei, IEEE 2009. *The Improvement of C4.5 Algorithm and Case Study*. IEEE Second International Symposium on Computational Intelligence and Design.
- [10] Moore, J. H. and M. G. Chang ,1980. *"Design of Decision Support Systems"* Data Base 12(1-2).