TOFEDU: The Future of Education Journal
Volume 5 Number 2 (2026) Page: 3856-3873
E-ISSN 2961-7553 P-ISSN 2963-8135
https://journal.tofedu.or.id/index.php/journal/index

The C4.5 Algorithm in Data Mining for Classifying Prospective Student Admissions (A Case Study of AMIK-DP)

Rezki Fauzi*[1], Gustri Efendi[2], Fitriany[3]
* gustriefendi@gmail.com
[1] Fakultas Teknik, Universitas Nahdlatul Ulama Sumatera Barat, Padang, Indonesia
[2] Fakultas Teknik, Universitas Nahdlatul Ulama Sumatera Barat, Padang, Indonesia
[3] Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK, Padang , Indonesia

ABSTRACT
Algorithm c4.5 data classification algorithm bertipe decision tree. algorithm decision tree c4.5 built with a few stage that cover attribute election as root, make branch to every value and divide case in branch. tahap-tahan this be repeatinged to every branch until all cases in has class same. From decision tree completion so be got several rule a case. in this case author clasify pass from a student university at one particular college (AMIK Depati Parbo Kerinci) not only based on accomplishment index criteria komulatif but also extracurrucular skill likes practice labor, campus organization, assistant labor and age also be somebody yardstick to acceptable at a private resort / government. with algorithm applications c4.5 this can help side baak in recommend student university genuinely skilled.

Keyword: data mining, data classification, algorithm c4.5, decision tree, student university pass

PENDAHULUAN
Perkembangan teknologi informasi telah mendorong peningkatan volume data yang dihasilkan oleh berbagai organisasi, termasuk institusi pendidikan tinggi. Data yang tersimpan dalam sistem informasi akademik tidak lagi hanya berfungsi sebagai arsip administratif, tetapi juga menjadi sumber informasi yang dapat dimanfaatkan untuk mendukung pengambilan keputusan secara objektif dan berbasis data. Salah satu pendekatan yang banyak digunakan untuk mengekstraksi pengetahuan dari data adalah data mining, yaitu proses menemukan pola atau informasi tersembunyi dari sekumpulan data sehingga menghasilkan pengetahuan yang bernilai bagi pengambilan keputusan (Pramudiono, 2006). Dalam bidang pendidikan, pemanfaatan Educational Data Mining (EDM) semakin berkembang untuk mendukung berbagai proses akademik, termasuk evaluasi pembelajaran, prediksi prestasi mahasiswa, identifikasi risiko putus studi, dan seleksi penerimaan mahasiswa baru.

Salah satu teknik utama dalam data mining adalah klasifikasi, yaitu proses membangun model berdasarkan data historis untuk mengelompokkan data baru ke dalam kelas tertentu. Berbagai algoritma klasifikasi telah dikembangkan, salah satunya adalah algoritma C4.5 yang membentuk model dalam bentuk pohon keputusan (decision tree) berdasarkan nilai information gain dan gain ratio (Kusrini & Emha Taufiq Luthfi, 2009). Keunggulan algoritma C4.5 terletak pada kemampuannya menghasilkan aturan keputusan (decision rules) yang mudah dipahami sehingga tidak hanya memberikan hasil klasifikasi, tetapi juga menjelaskan faktor-faktor yang memengaruhi suatu keputusan.

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Perkembangan penelitian mengenai penerapan algoritma klasifikasi dalam bidang pendidikan menunjukkan peningkatan yang signifikan dalam beberapa tahun terakhir. Berbagai penelitian telah memanfaatkan algoritma Decision Tree C4.5, Random Forest, Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN) untuk menyelesaikan berbagai permasalahan akademik, seperti prediksi prestasi belajar, klasifikasi tingkat kelulusan, identifikasi mahasiswa berisiko, serta seleksi penerimaan mahasiswa baru. Hasil penelitian tersebut menunjukkan bahwa algoritma klasifikasi mampu menghasilkan model prediksi dengan tingkat akurasi yang baik sehingga dapat dimanfaatkan sebagai pendukung pengambilan keputusan di lingkungan perguruan tinggi.

Di antara berbagai algoritma tersebut, C4.5 masih menjadi salah satu metode yang banyak digunakan karena memiliki keseimbangan antara tingkat akurasi dan kemudahan interpretasi model. Berbagai penelitian melaporkan bahwa algoritma ini mampu mengidentifikasi atribut-atribut yang paling berpengaruh terhadap hasil klasifikasi melalui pohon keputusan dan aturan keputusan (decision rules). Dengan demikian, hasil klasifikasi tidak hanya menghasilkan prediksi, tetapi juga memberikan informasi yang mudah dipahami oleh pengambil kebijakan sebagai dasar dalam penyusunan strategi maupun evaluasi proses seleksi mahasiswa.

Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada peningkatan performa model klasifikasi yang diukur menggunakan akurasi, precision, recall, atau F1-score. Pemanfaatan aturan keputusan yang dihasilkan algoritma C4.5 sebagai dasar penyusunan kebijakan dalam proses seleksi mahasiswa baru masih relatif terbatas. Selain itu, implementasi algoritma C4.5 pada proses klasifikasi kelulusan calon mahasiswa baru di perguruan tinggi vokasi, khususnya AMIK Depati Parbo Kerinci, masih belum banyak dilaporkan sehingga efektivitas penerapannya pada karakteristik data lokal belum diketahui secara memadai.

Secara empiris, proses Seleksi Penerimaan Mahasiswa Baru (SPMB) di AMIK Depati Parbo Kerinci masih didasarkan pada nilai tes seleksi sebagai indikator utama dalam menentukan kelulusan calon mahasiswa. Pendekatan tersebut belum memanfaatkan pola hubungan antar atribut yang tersimpan pada data historis penerimaan mahasiswa sehingga informasi yang diperoleh masih terbatas pada hasil akhir seleksi. Kondisi ini menunjukkan adanya kebutuhan akan suatu model klasifikasi yang tidak hanya mampu memprediksi kelulusan calon mahasiswa, tetapi juga menjelaskan faktor-faktor yang paling berpengaruh terhadap keputusan tersebut secara objektif dan transparan.

Berdasarkan kesenjangan tersebut, di rumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana penerapan algoritma C4.5 dalam mengklasifikasikan kelulusan calon mahasiswa baru di AMIK Depati Parbo Kerinci;
2. Atribut apa saja yang paling berpengaruh terhadap keputusan kelulusan berdasarkan model pohon keputusan yang dihasilkan; dan
3. Bagaimana mengevaluasi kinerja model klasifikasi yang dihasilkan oleh algoritma C45.

Penelitian ini bertujuan membangun model klasifikasi kelulusan calon mahasiswa baru menggunakan algoritma C4.5, mengidentifikasi atribut yang paling berpengaruh terhadap keputusan kelulusan, serta mengevaluasi performa model yang dihasilkan. Kebaruan penelitian ini terletak pada penerapan algoritma C4.5 pada data historis penerimaan mahasiswa baru di AMIK Depati Parbo Kerinci dengan menghasilkan model pohon keputusan beserta aturan klasifikasi yang tidak hanya digunakan untuk prediksi, tetapi juga sebagai dasar pendukung pengambilan keputusan yang objektif, transparan, dan mudah diinterpretasikan dalam proses seleksi mahasiswa baru.

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain penelitian eksperimen pada bidang data mining. Tahapan penelitian diawali dengan identifikasi permasalahan, pengumpulan data historis penerimaan mahasiswa baru, praproses data (data preprocessing), pembentukan model klasifikasi menggunakan algoritma C4.5, pengujian performa model, serta interpretasi aturan keputusan (decision rules) yang dihasilkan. Proses penelitian meliputi seleksi data, pembersihan data (data cleaning), transformasi data, proses data mining, dan evaluasi hasil. Pendekatan ini dipilih karena mampu menghasilkan model klasifikasi sekaligus mengidentifikasi atribut-atribut yang berpengaruh terhadap keputusan kelulusan calon mahasiswa baru.

Populasi dalam penelitian ini adalah seluruh data historis Seleksi Penerimaan Mahasiswa Baru (SPMB) di AMIK Depati Parbo Kerinci yang tersimpan dalam sistem informasi akademik. Sampel penelitian menggunakan teknik sampling jenuh (total sampling), yaitu seluruh data yang memenuhi kriteria kelengkapan atribut dan memiliki label status kelulusan digunakan sebagai data penelitian. Data yang digunakan merupakan data sekunder yang diperoleh dari dokumen administrasi penerimaan mahasiswa baru dengan tetap menjaga kerahasiaan identitas calon mahasiswa melalui proses anonimisasi. Atribut yang digunakan dalam proses klasifikasi disesuaikan dengan data yang tersedia, seperti nilai tes seleksi dan atribut lain yang berkaitan dengan proses penerimaan mahasiswa, sedangkan atribut target (class label) adalah status kelulusan calon mahasiswa, yaitu lulus atau tidak lulus.

Teknik pengumpulan data dilakukan melalui studi dokumentasi terhadap data historis penerimaan mahasiswa baru yang diperoleh dari bagian akademik AMIK Depati Parbo Kerinci. Sebelum dilakukan proses pemodelan, data terlebih dahulu melalui tahap praproses yang meliputi pemeriksaan kelengkapan data, penghapusan data duplikat, penanganan nilai yang hilang (missing value), transformasi tipe data apabila diperlukan, serta seleksi atribut yang relevan dengan tujuan penelitian. Karena penelitian menggunakan data sekunder, instrumen penelitian berupa lembar observasi data (data extraction sheet) yang digunakan untuk mengidentifikasi, mencatat, dan memverifikasi atribut-atribut yang akan digunakan dalam proses klasifikasi. Seluruh data yang telah melalui tahap praproses kemudian disusun dalam format yang sesuai untuk proses pengolahan menggunakan perangkat lunak data mining.

Analisis data dilakukan menggunakan teknik klasifikasi dengan algoritma C4.5. Tahap analisis diawali dengan pembentukan model pohon keputusan berdasarkan data latih (training data) untuk memperoleh aturan-aturan klasifikasi (decision rules) yang menggambarkan hubungan antar atribut terhadap status kelulusan calon mahasiswa. Selanjutnya, model diuji menggunakan data pengujian (testing data) untuk mengetahui kemampuan generalisasi model dalam mengklasifikasikan data baru. Selain mengevaluasi performa model, penelitian ini juga menganalisis struktur pohon keputusan dan aturan klasifikasi yang dihasilkan untuk mengidentifikasi atribut-atribut yang paling berpengaruh terhadap keputusan kelulusan. Hasil analisis tersebut kemudian diinterpretasikan sebagai dasar dalam memberikan rekomendasi bagi pihak AMIK Depati Parbo Kerinci dalam mendukung proses pengambilan keputusan pada seleksi penerimaan mahasiswa baru secara lebih objektif, transparan, dan berbasis data.

HASIL DAN PEMBAHASAN

Sistem yang akan dirancang dan digunakan dalam menentukan kelulusan calon mahasiswa baru dirancang dengan menggunakan metode analisis terstruktur, pada metode trans*form*asi input, proses dan *output* dinyatakan dengan Diagram Arus Data (*Data Flow Diagram)* sebagai pernyataan algoritma pelatihan dan pengujian. Pada tahapan ini penggunaan notasi-notasi digunakan untuk menggambarkan arus data sistem, di mana akan sangat membantu dalam proses komunikasi dengan pemakai. Proses pengambilan keputusan dalam klasifikasi calon kelulusan mahasiswa baru dikelompokkan menjadi beberapa kriteria yaitu :

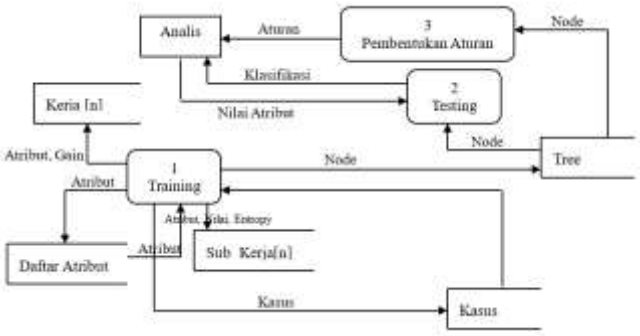
This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

1. Kriteria Nilai Ujian Tertulis
2. Kriteria Nilai Wawancara
3. Kriteria Nilai Ujian Praktek
4. Kriteria Nilai Rata-rata NEM

Perancangan Sistem

Aliran data yang terjadi pada proses pembentukan sistem kelulusan calon mahasiswa baru ditunjukkan dalam Diagram Alir Data level 1 pada gambar 1 di bawah ini :



Gambar 1: Diagram Alir Data Level 1

Untuk mendapatkan model, kita harus melakukan analisis terhadap data latih *(training set)*. Sedangkan data uji *(test set)* digunakan untuk mengetahui tingkat akurasi dari model yang telah dihasilkan.

Pemilihan Variabel

Data kelulusan calon mahasiswa baru pada Perguruan Tinggi Swasta AMIK Depati Parbo Kerinci tahun ajaran 2009/2010 memiliki *form*at seperti berikut :

1. No. Ujian
2. Nama
3. Tempat/ Tanggal Lahir
4. Jenis Kelamin
5. Alamat
6. Nilai Ujian Tertulis
7. Keterangan Kelulusan

Melakukan Pra-Proses

Berdasarkan variabel-variabel yang sudah terpilih, *form*at data menjadi seperti tampak pada tabel 1 di bawah ini :

Tabel 1 *Form*at Data Calon Mahasiswa Baru Setelah Pemilihan Variabel

No.	No. Ujian	Nilai Tertulis	Wawancara	Nilai Praktek	Rata-rata NEM	Kelulusan
1	70001	81	HADIR	80	6.32	YA
2	70002	80	HADIR	76	6.8	YA
3	70003	83	HADIR	82	6.25	YA
4	70004	65	TDKHADIR	75	6.72	TIDAK
5	70005	80	TDKHADIR	74	6.12	TIDAK
6	70006	82	HADIR	68	6.24	YA
7	70007	64	HADIR	69	5.39	TIDAK
8	70008	69	HADIR	80	5.52	YA
9	70019	83	HADIR	82	5.64	YA
10	70010	82	HADIR	73	5.78	YA
11	70011	70	HADIR	69	6.8	YA
12	70012	85	HADIR	70	6.94	YA
13	70013	72	TDKHADIR	71	5.67	TIDAK
14	70014	80	HADIR	73	6.23	YA
15	70015	77	HADIR	84	5.86	YA

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

16	70016	75	HADIR	68	6.56	TIDAK
17	70017	81	HADIR	83	7.20	YA
18	70018	76	HADIR	79	6.54	TIDAK
18	70019	82	HADIR	77	6.32	YA
20	70020	85	HADIR	80	6.81	YA
21	70021	67	TDKHADIR	65	6.72	TIDAK
22	70022	83	HADIR	80	6.56	YA
23	70023	81	HADIR	81	6.86	YA
24	70024	66	TDKHADIR	81	6.67	YA
25	70025	80	HADIR	65	7.24	YA
26	70026	64	HADIR	66	5.68	YA
27	70027	81	HADIR	80	6.35	YA
28	70028	68	HADIR	69	5.79	TIDAK
29	70029	69	TDKHADIR	70	6.21	TIDAK
30	70030	73	HADIR	80	5.87	YA
31	70031	80	HADIR	82	7.16	YA
32	70032	82	HADIR	79	5.86	TIDAK
33	70033	81	HADIR	78	7.43	YA
34	70034	84	HADIR	77	6.25	YA
35	70035	82	HADIR	78	5.95	TIDAK
36	70036	81	HADIR	80	5.67	YA
37	70037	76	HADIR	76	7.14	TIDAK
38	70038	72	HADIR	75	6.89	TIDAK
39	70039	80	HADIR	74	5.45	TIDAK
40	70040	81	HADIR	80	5.63	YA

Selanjutnya dilakukan pra-proses dengan mengelompokkan data-data tersebut ke dalam kelas (K) dengan Sturges (Drs. Riduwan, M.B.A, 2003) dengan menggunakan Rumus :

Jumlah Kelas (K) = 1 + 3.3 log n

Dimana : n = jumlah data

K = 1 + 3.3 log(40)

K = 1 + 3.3 * 1.6020

K = 1 + 5.286798

K = 6.286798 $\Longrightarrow$ 6

Berarti dengan jumlah data 40, maka akan mendapatkan 6 kelas, data tersebut akan dikelompokkan berdasarkan atribut sebagai berikut.

1. Mengelompokkan Nilai Ujian Tertulis, pengelompokkan nilai ujian tertulis ini berdasarkan dari hasil ujian yang didapat oleh calon mahasiswa baru tersebut sehingga nilai tersebut dikelompokkan seperti terlihat pada tabel 4.2 dibawah ini.

Tabel 2 Klasifikasi Nilai Ujian Tertulis

Nilai Ujian Tertulis	Klasifikasi
0 – 15	1
16 – 32	2
33 – 47	3
48 – 63	4
64 – 79	5
>80	6

2. Mengelompokkan Nilai Wawancara berdasarkan keikutsertaan ataupun kehadiran pada pelaksanaan tes wawancara, seperti terlihat pada tabel 3 dibawah ini :

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Tabel 3 Klasifikasi Wawancara

Wawancara	Klasifikasi
Hadir	Y
Tidak Hadir	T

3. Mengelompokkan Nilai Prakek, berdasarkan hasil yang didapat pada pelaksanaan ujian praktek, seperti terlihat pada tabel 4 dibawah ini :

Tabel 4 Nilai Praktek

Nilai Ujian Praktek	Klasifikasi
0 – 15	1
16 – 32	2
33 – 47	3
48 – 63	4
64 – 79	5
>80	6

4. Mengelompokkan Nilai Rata-rata NEM, pengelompokkan ini diambil dari nilai NEM yang diraih oleh calon mahasiswa baru tersebut pada waktu tamat SMA atau sederajat, pengelompokkan nilai Rata-rata NEM tersebut dapat dilihat pada tabel 5 dibawah ini :

Tabel 5 Nilai Rata-rata NEM

Nilai Rata-rata NEM	Klasifikasi
0,00 – 1,59	1
1,60 – 3,29	2
3,30 – 4,79	3
4,80 – 6,39	4
6,40 – 7,90	5
>8,00	6

*Form*at data akhir setelah dilakukan pra-proses tampak seperti tabel 6 berikut ini :

Tabel 6 *Form*at Data Akhir

No.	No. Ujian	Nilai Tertulis	Wawancara	Nilai Praktek	Rata-rata NEM	Kelulusan
1	70001	6	Y	6	5	YA
2	70002	6	Y	5	6	YA
3	70003	6	Y	6	5	YA
4	70004	5	T	5	6	TIDAK
5	70005	6	T	5	5	TIDAK
6	70006	6	Y	5	5	YA
7	70007	5	Y	5	4	TIDAK
8	70008	5	Y	6	4	YA
9	70019	6	Y	6	5	YA
10	70010	6	Y	5	5	YA
11	70011	5	Y	5	6	YA
12	70012	6	Y	5	5	YA
13	70013	5	T	5	4	TIDAK
14	70014	6	Y	5	5	YA
15	70015	5	Y	6	4	YA
16	70016	5	Y	5	5	TIDAK
17	70017	6	Y	6	6	YA
18	70018	5	Y	5	5	TIDAK
19	70019	6	Y	5	5	YA
20	70020	6	Y	6	6	YA
21	70021	5	T	5	5	TIDAK
22	70022	6	Y	6	5	YA

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

23	70023	6	Y	6	6	YA
24	70024	5	T	6	5	YA
25	70025	6	Y	5	6	YA
26	70026	5	Y	5	4	YA
27	70027	6	Y	6	5	YA
28	70028	5	Y	5	4	TIDAK
29	70029	5	T	5	5	TIDAK
30	70030	5	Y	6	4	YA
31	70031	6	Y	6	6	YA
32	70032	6	Y	5	4	TIDAK
33	70033	6	Y	5	6	YA
34	70034	6	Y	5	5	YA
35	70035	6	Y	5	4	TIDAK
36	70036	6	Y	6	4	YA
37	70037	5	Y	5	6	TIDAK
38	70038	5	Y	5	6	TIDAK
39	70039	6	Y	5	4	TIDAK
40	70040	6	Y	6	4	YA

Format data akhir pada tabel 6 diatas didapat berdasarkan dari attribute yang sudah dikelompokkan atau diklasifikasi, misalkan data pada tabel 1 Nilai Ujian Tertulis adalah "81", setelah diklasifikasi menjadi "6", Wawancara adalah "hadir", maka setelah klasifikasi menjadi "Y", Nilai Praktek adalah "80" berubah menjadi "6", dan Rata-rata Nem adalah "6.32" berubah menjadi "5" dan seterusnya.

Pohon Keputusan

Dari *form*at data akhir kelulusan calon mahasiswa baru maka akan dilakukan klasifikasi data algoritma C4.5 dengan membuat pohon keputusan sebagai berikut :

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Dalam kasus yang tertera pada tabel 6 diatas, akan dibuat pohon keputusan untuk menentukan klasifikasi kelulusan calon mahasiswa baru (ya dan tidak) dengan melihat Nilai Ujian Tertulis, Nilai Wawancara, Nilai Ujian Praktek dan Nilai Rata-rata NEM. Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Menghitung jumlah kasus, jumlah kasus untuk keputusan "Ya", jumlah kasus untuk keputusan "Tidak", dan kasus yang dibagi berdasarkan atribut Nilai Ujian Tertulis, Nilai Wawancara, Nilai Ujian Praktek, Nilai Rata-rata NEM. Setelah itu, lakukan perhitungan *gain* untuk setiap atribut.

Menghitung Nilai Entopy tiap-tiap atribut :

***Entropy* (Total)**

$$Entropy(total) = \left(-\frac{26}{40} * log_2\left(\frac{26}{40}\right)\right) + \left(-\frac{14}{40} * log_2\left(\frac{14}{40}\right)\right)$$
$$= 0{,}934068$$

Entropy(total) adalah menghitung nilai total keputusan ya (26) dan tidak (14), sedangkan 40 adalah jumlah keseluruhan kasus.

Atribut Nilai Rata-rata NEM

$$Entropy(1) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$
$$Entropy(2) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

$$Entropy(3) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(4) = \left(-\frac{6}{12} * log_2\left(\frac{6}{12}\right)\right) + \left(-\frac{6}{12} * log_2\left(\frac{6}{12}\right)\right) = 1$$

$$Entropy(5) = \left(-\frac{12}{17} * log_2\left(\frac{12}{17}\right)\right) + \left(-\frac{5}{17} * log_2\left(\frac{5}{17}\right)\right) = 0.873981$$

$$Entropy(6) = \left(-\frac{8}{11} * log_2\left(\frac{8}{11}\right)\right) + \left(-\frac{3}{11} * log_2\left(\frac{3}{11}\right)\right) = 0.845351$$

Menghitung nilai *entrophy* atribut Nilai Rata-rata NEM berdasarkan dari tiap-tiap kelas (1, 2, 3, 4, 5, 6) pada Atribut Nilai Rata-rata NEM.

Atribut Nilai Ujian Tertulis

$$Entropy(1) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(2) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(3) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(4) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(5) = \left(-\frac{6}{16} * log_2\left(\frac{6}{16}\right)\right) + \left(-\frac{10}{16} * log_2\left(\frac{10}{16}\right)\right) = 0{,}954434$$

$$Entropy(6) = \left(-\frac{20}{24} * log_2\left(\frac{20}{24}\right)\right) + \left(-\frac{4}{24} * log_2\left(\frac{4}{24}\right)\right) = 0{,}650022$$

Menghitung nilai *entrophy* Atribut Nilai Ujian Tertulis berdasarkan dari tiap-tiap kelas (1, 2, 3, 4, 5, 6) pada Atribut Nilai Ujian Tertulis.

a. Atribut Nilai Ujian Praktek

$$Entropy(1) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(2) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(3) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(4) = \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * log_2\left(\frac{0}{0}\right)\right) = 0$$

$$Entropy(5) = \left(-\frac{11}{25} * log_2\left(\frac{11}{25}\right)\right) + \left(-\frac{14}{25} * log_2\left(\frac{14}{25}\right)\right) = 0{,}989588$$

$$Entropy(6) = \left(-\frac{15}{15} * log_2\left(\frac{15}{15}\right)\right) + \left(-\frac{0}{15} * log_2\left(\frac{0}{15}\right)\right) = 0$$

Menghitung nilai *entrophy* Nilai Ujian Praktek berdasarkan dari tiap-tiap kelas (1,

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

2, 3, 4, 5, 6) pada Atribut Nilai Ujian Praktek.

b. Atribut Wawancara

$$Entropy(Y) = \left(-\frac{25}{34} * \log_2\left(\frac{25}{34}\right)\right) + \left(-\frac{9}{34} * \log_2\left(\frac{9}{34}\right)\right)$$
$$= 0{,}833765$$

$$Entropy(T) = \left(-\frac{1}{6} * \log_2\left(\frac{1}{6}\right)\right) + \left(-\frac{5}{6} * \log_2\left(\frac{5}{6}\right)\right)$$
$$= 0{,}650022$$

Menghitung Nilai Gain tiap-tiap atribut :
Gain (Total, Nilai Rata-rata NEM)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Rata\ NEM_i|}{|Total|} * Entropy(Rata\ NEM_i)$$
$$= 0.934068 - \left(\left(\frac{12}{40} * 1\right) + \left(\frac{17}{40} * 0.873981\right) + \left(\frac{11}{40} * 0.845351\right)\right) = 0.030155$$

Gain (Total, Nilai Tertulis)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Tertulis_i|}{|Total|} * Entropy(Nilai\ Tertulis_i)$$
$$= 0.934068 - \left(\left(\frac{16}{40} * 0.954434\right) + \left(\frac{24}{40} * 0.650022\right)\right)$$
$$= 0.162281$$

Gain(Total, Nilai Ujian Praktek)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Praktek_i|}{|Total|} * Entropy(Nilai\ Praktek_i)$$
$$= 0.934068 - \left(\left(\frac{25}{40} * 0.989588\right) + \left(\frac{55}{40} * 0\right)\right) = 0.315576$$

Gain(Total, Nilai Wawancara)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Wawancara_i|}{|Total|} * Entropy(Nilai\ Wawancara_i)$$
$$= 0.934068 - \left(\left(\frac{34}{40} * 0.833765\right) + \left(\frac{6}{40} * 0.650022\right)\right)$$
$$= 0.127864$$

Setelah nilai *entropy* dan *gain* dihitung, kemudian hasil dari perhitungan tersebut dimasukkan ke dalam tabel 7 berikut ini :

Tabel 7 Perhitungan Node 1

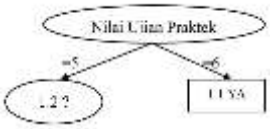
Node			Jumlah Kasus (S)	Ya (S1)	Tidak (S2)	Entropy	Gain
1	TOTAL		40	26	14	0.934068	
	Nilai Rata-rata NEM						0.030155
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	12	6	6	1	

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

		5	17	12	5	0.873981	
		6	11	8	3	0.845351	
	Nilai Ujian Tertulis						0.162281
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	0	0	0		
		5	16	6	10	0.954434	
		6	24	20	4	0.650022	
	Nilai Ujian Praktek						0.315576
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	0	0	0		
		5	25	11	14	0.989588	
		6	15	15	0	0	
	Nilai Wawancara						0.127864
		Y	34	25	9	0.833765	
		T	6	1	5	0.650022	

Dari perhitungan pada tabel 7 dapat diketahui bahwa atribut dengan *gain* tertinggi adalah Nilai Ujian Praktek sebesar 0.315576. Berarti Nilai Ujian Praktek dapat menjadi node akar. Ada dua nilai atribut dari Nilai Ujian Praktek yaitu "5" dan "6". Dari nilai atribut tersebut, nilai "6" sudah Lulus, sehingga tidak perlu dilakukan perhitungan, tetapi nilai atribut "5" masih perlu dilakukan perhitungan lagi, seperti pada gambar 2 berikut ini :



Gambar 3: Pohon Keputusan Hasil Perhitungan Node 1

Selanjutnya adalah menyelesaikan untuk menghitung Node 1.2 sebagai akar, sama dengan cara yang diatas dengan menghitung nilai *entropy* dari atribut yang tersisa yaitu Nilai Rata-rata NEM, Nilai Ujian Tertulis dan Nilai Wawancara, setelah di hitung *entropy*, kemudian menghitung *gain* untuk tiap-tiap atribut.

***Entropy* (Nilai Ujian Praktek, 5)**

$$= \left(-\frac{11}{25} * \ log_2\left(\frac{11}{25}\right)\right) + \left(-\frac{14}{25} * \ log_2\left(\frac{14}{25}\right)\right)$$

$= 0.989588$

Entropy (Nilai Ujian Praktek, 5) adalah nilai entropy setelah dilakukan perhitungan sebelumnya yang bernilai Tidak, dan masih dilakukan lagi perhitungan berikutnya untuk menetukan kelulusan dari calon mahasiswa baru tersebut.

a. Atribut Nilai Rata-rata NEM

$$Entropy(4) = \left(-\frac{1}{7} * \ log_2\left(\frac{1}{7}\right)\right) + \left(-\frac{6}{7} * \ log_2\left(\frac{6}{7}\right)\right)$$

$= 0.591673$

$$Entropy(5) = \left(-\frac{6}{11} * \ log_2\left(\frac{6}{11}\right)\right) + \left(-\frac{5}{11} * \ log_2\left(\frac{5}{11}\right)\right)$$

$$= 0.99403$$

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

$$Entropy(6) = \left(-\frac{4}{7} * log_2\left(\frac{4}{7}\right)\right) + \left(-\frac{3}{7} * log_2\left(\frac{3}{7}\right)\right)$$

$= 0.9852228$

Menghitung nilai *entrophy* Atribut Nilai Rata-rata NEM berdasarkan dari tiap-tiap kelas (1, 2, 3, 4, 5, 6) pada Atribut Nilai Rata-rata NEM.

b. Atribut Nilai Ujian Tertulis

$$Entropy(5) = \left(-\frac{3}{13} * log_2\left(\frac{3}{13}\right)\right) + \left(-\frac{10}{13} * log_2\left(\frac{10}{13}\right)\right)$$
$$= 0.77935$$
$$Entropy(6) = \left(-\frac{8}{12} * log_2\left(\frac{8}{12}\right)\right) + \left(-\frac{4}{12} * log_2\left(\frac{4}{12}\right)\right)$$

$= 0.918296$

Menghitung nilai *entrophy* atribut Nilai Ujian Tertulis berdasarkan dari tiap-tiap kelas (1, 2, 3, 4, 5, 6) pada atribut Nilai Ujian Tertulis.

c. Atribut Nilai Wawancara

$$Entropy(Y) = \left(-\frac{11}{20} * log_2\left(\frac{11}{20}\right)\right) + \left(-\frac{9}{20} * log_2\left(\frac{9}{20}\right)\right)$$

$= 0.992774$

$$Entropy(T) = \left(-\frac{0}{5} * log_2\left(\frac{0}{5}\right)\right) + \left(-\frac{5}{5} * log_2\left(\frac{5}{5}\right)\right) = 0$$

Menghitung nilai *entrophy* atribut Wawancara berdasarkan dari tiap-tiap kelas (Y,T) pada atribut Wawancara.

Gain (Total, Nilai Rata-rata NEM)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Rata\ NEM_i|}{|Total|} * Entropy(Rata\ NEM_i)$$

$$= 0.989588 - \left(\left(\frac{7}{25} * 0.591673\right) + \left(\frac{11}{25} * 0.99403\right) + \left(\frac{7}{25} * 0.985228\right)\right)$$

$= 0.110683$

Gain (Total, Nilai Tertulis)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Tertulis_i|}{|Total|} * Entropy(Nilai\ Tertulis_i)$$

$$= 0.989588 - \left(\left(\frac{13}{25} * 0.77935\right) + \left(\frac{12}{25} * 0.918296\right)\right)$$

$= 0.143544$

Gain(Total, Nilai Wawancara)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Wawancara_i|}{|Total|} * Entropy(Nilai\ Wawancara_i)$$

$$= 0.989588 - \left(\left(\frac{20}{25} * 0.992774\right) + \left(\frac{5}{25} * 0\right)\right)$$
$$= 0.195369$$

Setelah nilai *entropy* dan *gain* dihitung, kemudian hasil dari perhitungan tersebut dimasukkan ke dalam tabel 8 berikut ini :

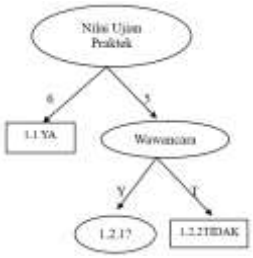
This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Tabel 8 Perhitungan Node 1.2

Node			Jumlah Kasus (S)	Ya (S1)	Tidak (S2)	Entropy	Gain
1.2	Nilai Ujian Praktek-5		25	11	14	0.989588	
	Nilai Rata-rata NEM						0.110683
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	7	1	6	0.591673	
		5	11	6	5	0.99403	
		6	7	4	3	0.985228	
	Nilai Ujian Tertulis						0.143544
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	0	0	0		
		5	13	3	10	0.77935	
		6	12	8	4	0.918296	
	Wawancara						0.195369
		Y	20	11	9	0.992774	
		T	5	0	5	0	

Dari hasil tabel 8 dapat diketahui bahwa atribut *gain* tertinggi adalah Nilai Wawancara, yaitu sebesar 0.195369, berarti Wawancara dapat menjadi node akar. Ada dua nilai atribut dari Wawancara yaitu "Y" dan "T". Dari nilai atribut tersebut, nilai "T" sudah Tidak Lulus, sehingga tidak perlu dilakukan perhitungan, tetapi nilai atribut "Y" masih perlu dilakukan perhitungan lagi, seperti pada gambar 4 berikut ini :



Gambar 4: Pohon Keputusan Hasil Perhitungan Node 1.2

Selanjutnya adalah menyelesaikan untuk menghitung Node 1.2.1 sebagai akar, sama dengan cara yang diatas dengan menghitung nilai *entropy* dari atribut yang tersisa yaitu Nilai Rata-rata NEM dan Nilai Ujian Tertulis, setelah di hitung *entropy*, kemudian menghitung *gain* untuk tiap-tiap atribut.

***Entropy* (Wawancara, Y)**

$$= \left(-\frac{11}{20} * log_2\left(\frac{11}{20}\right)\right) + \left(-\frac{9}{20} * log_2\left(\frac{9}{20}\right)\right)$$

$$= 0.992774$$

a. Atribut Nilai Rata-rata NEM

$$Entropy(4) = \left(-\frac{1}{6} * log_2\left(\frac{1}{6}\right)\right) + \left(-\frac{5}{6} * log_2\left(\frac{5}{6}\right)\right)$$

$$= 0.650022$$

$$Entropy(5) = \left(-\frac{6}{8} * log_2\left(\frac{6}{8}\right)\right) + \left(-\frac{2}{8} * log_2\left(\frac{2}{8}\right)\right)$$

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

$$= 0.811278$$

$$Entropy(6) = \left(-\frac{4}{6} * log_2\left(\frac{4}{6}\right)\right) + \left(-\frac{2}{6} * log_2\left(\frac{2}{6}\right)\right)$$

$$= 0.918296$$

b. Atribut Nilai Ujian Tertulis

$$Entropy(5) = \left(-\frac{6}{8} * log_2\left(\frac{6}{8}\right)\right) + \left(-\frac{2}{8} * log_2\left(\frac{2}{8}\right)\right)$$

$$= 0.811278$$

$$Entropy(6) = \left(-\frac{8}{12} * log_2\left(\frac{8}{12}\right)\right) + \left(-\frac{4}{12} * log_2\left(\frac{4}{12}\right)\right)$$

$$= 0.918296$$

Gain (Total, Nilai Rata-rata NEM)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Rata\ NEM_i|}{|Total|} * Entropy(Rata\ NEM_i)$$

$$= 0.992744 - \left(\left(\frac{6}{20} * 0.650022\right) + \left(\frac{8}{20} * 0.811278\right) + \left(\frac{6}{20} * 0.918296\right)\right)$$

$$= 0.197737$$

Gain (Total, Nilai Tertulis)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Nilai\ Tertulis_i|}{|Total|} * Entropy(Nilai\ Tertulis_i)$$

$$= 0.992774 - \left(\left(\frac{8}{20} * 0.811278\right) + \left(\frac{12}{20} * 0.918296\right)\right)$$

$$= 0.117255$$

Setelah nilai *entropy* dan *gain* dihitung, kemudian hasil dari perhitungan tersebut dimasukkan ke dalam tabel 9 berikut ini :

Tabel 9 Perhitungan Node 1.2.1

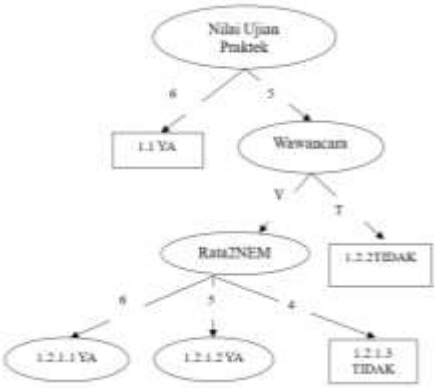
Node			Jumlah Kasus (S)	Ya (S1)	Tidak (S2)	Entropy	Gain
1.2.1	Wawancara –Y		20	11	9	0.992774	
	Nilai Rata-rata NEM						0.197737
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	6	1	5		
		5	8	6	2		
		6	6	4	3		
	Nilai Ujian Tertulis						0.117255
		1	0	0	0		
		2	0	0	0		
		3	0	0	0		
		4	0	0	0		
		5	8	6	2		
		6	12	8	4		

Dari hasil tabel 9 dapat diketahui bahwa atribut *gain* tertinggi adalah Rata-rata NEM, yaitu sebesar 0.197737, berarti Rata-rata NEM dapat menjadi node akar selanjutnya. Ada tiga nilai atribut dari Nilai Rata-rata NEM yaitu “4”, “5” dan “6”. Dari nilai atribut tersebut, nilai

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

"4" sudah Tidak Lulus, sehingga tidak perlu dilakukan perhitungan, tetapi nilai atribut "5" dan "6" masih perlu dilakukan perhitungan lagi. Pohon keputusan yang terbentuk pada saat ini adalah seperti terlihat pada gambar 5 berikut ini :



Gambar 5: Pohon Keputusan Hasil Perhitungan Node 1.2.1

Berdasarkan pohon keputusan terakhir yang terbentuk pada gambar 5 diatas, maka aturan atau *rule* yang terbentuk adalah sebagai berikut :

1. Jika Nilai Ujian Praktek = 6, maka Kelulusan = YA
2. Jika Nilai Ujian Praktek = 5 dan Wawancara = T, Maka Kelulusan = TIDAK
3. Jika Nilai Ujian Praktek = 5 dan Wawancara = Y dan Rata-rata NEM = 6 dan Nilai Tertulis = 6, maka kelulusan = YA
4. Jika Nilai Ujian Praktek = 5 dan Wawancara = Y dan Rata-rata NEM = 6, dan Nilai Tertulis = 5, maka kelulusan = TIDAK
5. Jika Nilai Ujian Praktek = 5 dan Wawancara = Y dan Rata-rata NEM = 5, dan Nilai Tertulis = 6, maka kelulusan = YA
6. Jika Nilai Ujian Praktek = 5 dan Wawancara = Y dan Rata-rata NEM = 5 dan Nilai Tertulis = 5, maka kelulusan = TIDAK
7. Jika Nilai Ujian Praktek = 5 dan Wawancara = Y dan Rata-rata NEM = 4, maka kelulusan = TIDAK

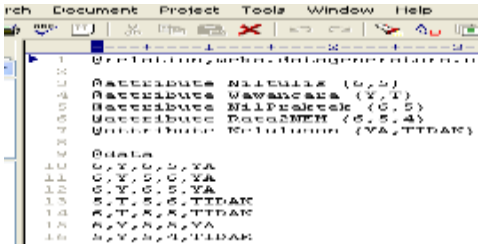
Hasil Pengujian

Untuk menguji kebenaran dari hasil pengolahan data digunakan salah satu *software* aplikasi *WEKA Knowlege Explorer* dengan langkah-langkah sebagai berikut :

1. Seluruh variabel-variabel (terdiri dari attribut kondisi dan atribut keputusan) yang digunakan untuk menentukan kelulusan calon mahasiswa baru disimpan pada *microsoft excel* dengan nama file datatesis.xls.
2. Agar data tersebut dapat digunakan pada aplikasi *WEKA decision tree*, maka data tersebut harus disimpan dalam format 'ARFF' (Gambar 7). File ARFF (*Attribute-Relation File Format*) adalah sebuah file teks ASCII yang berisi daftar *instances* dalam sekumpulan atribut. File ARFF dikembangkan oleh *Machine Learning Project* di *Department of Computer Science of The University of Waikato* untuk digunakan dalam perangkat lunak *WEKA*. Data dalam format .arff tersebut dapat dipemenuhi dengan cara:
    - Data dipisahkan dengan koma, dengan kelas sebagai atribut terakhir.
    - Bagian *header* diawali dengan @RELATION.
    - Tiap atribut ditandai dengan @ATTRIBUTE. Tipe-tipe data dalam *WEKA*: numerik(REAL atau INTEGER), nominal, String, dan *Date*.
    - Bagian data diawali dengan @DATA

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135



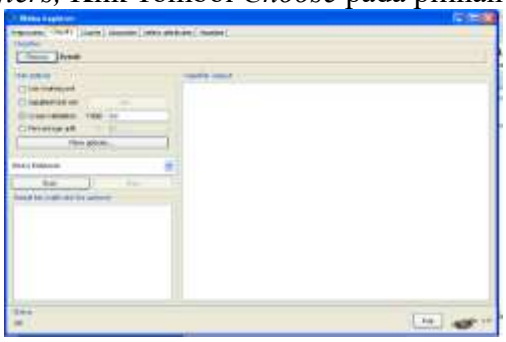
Gambar 6 datatesis.arff

3. Visualisasi tiap atribut dapat dilihat dengan meng-klik tombol *Visualize.* Visualisasi ini menggunakan diagram batang, yang mengilustrasikan jumlah dari masing-masing tipe pada atribut yang ada.



Gambar 7 *WEKA Explorer* dengan *Visualisasi Data*

4. Kemudian Pilih Tab *Classifiers,* Klik Tombol *Choose* pada pilihan *Classifier*



Gambar 8: *WEKA Explorer* dengan *Tab Classifiers*

5. Pada *menu Classifier* pilih *root Tree* dan Klasifikasi dengan pohon keputusan, yaitu metode klasifikasi dengan struktur pohon yang merepresentasikan kriteria pembagian kelas.



Gambar 9: Pilihan model *Classifiers*

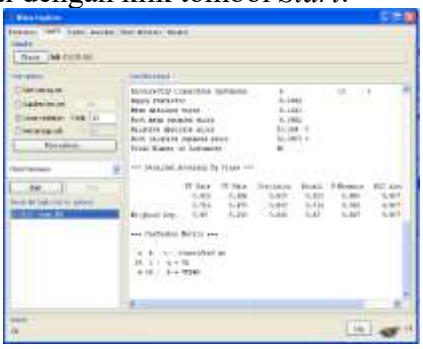
6. Tersedia pula pilihan pengujian bagi *user* di bawah bagian pemilihan *classifiers,* yaitu:

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

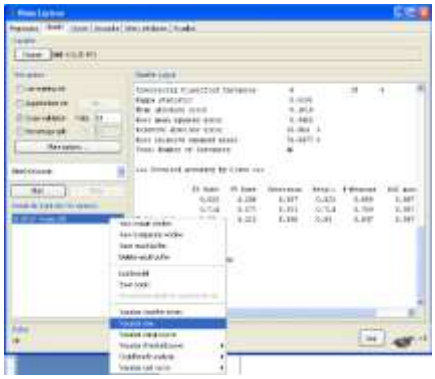
a. *Use training set. Classifier* dievaluasi pada kemampuannya memprediksi *class* dari *instances* yang diujikan.
b. *Supplied test set.* Pengujian kemampuan *classifier* dilakukan terhadap himpunan *instances* terpisah yang di-*load* dari sebuah file.
c. *Cross-validation*. *Classifier* dievaluasi dengan *cross-validation*, menurut jumlah *folds* yang dimasukkan pada kolom *Folds*.
d. *Percentage split.* Evaluasi *clasifier* dilakukan pada sejumlah persentase tertentu dari data yang digunakan untuk pengujian.

7. Setelah *classifier,* pilihan-pilihan pengujian, dan *class* telah ditentukan, proses pembelajaran dapat dimulai dengan klik tombol *Start*.



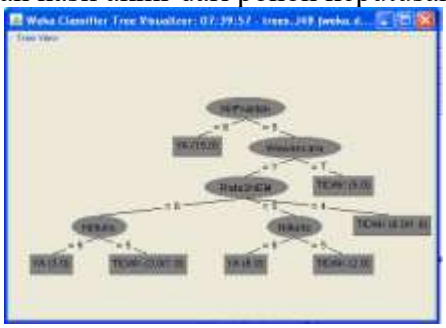
Gambar 10: Hasil *Classifier*

Errors klasifikasi dapat divisualisasikan dalam sebuah *tool* visualisasi data *pop-up*. Jika *classifier* menghasilkan sebuah *decision tree*, dapat ditampilkan secara grafis dalam sebuah *pop-up tree visualizer*.



Gambar 11 Pilihan *Visualisasi Tree*

8. Hasil Visualize Tree, adalah hasil akhir dari pohon keputusan.



Gambar 12 Hasil *Visualisasi Tree*

Dimana *rule* yang dihasilkan pada gambar 13 adalah :

- Bila Nilai praktek = 6 maka kelulusan = Y1

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

- Bila Nilai Praktek = 5 dan Nilai Wawancara = T maka kelulusan = T
- Bila Nilai Praktek = 5 dan Nilai Wawancara = Y dan Rata-rata NEM = 6 dan Nilai Tulis = 6 maka Kelulusan = Y
- Bila Nilai Praktek = 5 dan Nilai Wawancara = Y dan Rata-rata NEM = 6 dan Nilai Tulis = 5 maka Kelulusan = T
- Bila Nilai Praktek = 5 dan Nilai Wawancara = Y dan Rata-rata NEM = 5 dan Nilai Tulis = 6 maka Kelulusan = Y
- Bila Nilai Praktek = 5 dan Nilai Wawancara = Y dan Rata-rata NEM = 6 dan Nilai Tulis = 5 maka Kelulusan = T
- Bila Nilai Praktek = 5 dan Nilai Wawancara = Y dan Rata-rata NEM = 4 maka Kelulusan = Y

Setelah dilakukan pengujian pada software *WEKA* 3.6 untuk memprediksi tingkat kelulusan calon mahasiswa baru seperti yang sudah dijelaskan di atas ternyata hasil *rule* yang dihasilkan pada *WEKA* pada gambar 13 hampir sama dengan *rule* yang dihasilkan pada proses manual penemuan *knowledge* menggunakan pohon keputusan.

KESIMPULAN

Dari uraian pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

1. Metode pohon keputusan yang diproses dengan Weka lebih efektif dan fleksibel jika digunakan pada proses pengklasifikasian kelulusan calon mahasiswa baru.
2. Pemilihan variabel (atribut kondisi dan attribut keputusan) yang akan digunakan dalam memprediksi juga sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan.
3. Sistem yang dibangun dapat membantu dalam pengklasifikasian kelulusan calon mahasiswa baru yang umumnya masih dilakukan secara manual.
4. Algoritma C4.5 dengan metode pohon keputusan dapat memberikan informasi eksekutif dan sistem digunakan untuk menggambarkan proses yang terkait dengan pengklasifikasian kelulusan calon mahasiswa baru.
5. Sistem pengklasifikasian kelulusan calon mahasiswa baru menggunakan algoritma C4.5 dapat digunakan dalam pengambilan keputusan untuk mencari alternatif yang baik.

DAFTAR PUSTAKA

Iko Pramudiono. (2003). "*Jurnal Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data*"

Kusrini dan Luthfi Taufiq Emha. (2009). " *Algoritma Data Mining*". Yogyakarta : Andi

Kusrini1, Sri Hartati2, Retantyo Wardoyo3, Agus Harjoko4. (2009). "*Jurnal Perbandingan Metode Nearest Neighbor dan Algoritma C4.5 untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa di STMIK Amikom Yogyakarta*"

Sani Susanto1, Dedy Suryadi2. (2010). *"Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data"*. Yogyakarta. Andi

Veronica Sri Moertini. (2007). "*Disertasi Pengembangan Skalabilitas Algoritma Klasifikasi C4.5 dengan Pendekatan Konsep Operator Relasi (Studi Kasus: Pra-Pengolahan dan Klasifikasi Citra Batik)*"

Veronica S. Moertini. "*Jurnal Penanganan Atribut Citra dengan Wavelet untuk Pengembangan Aloritma C4.5*"

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Yudho Giri Sucahyo. (2003). "*Jurnal Data Mining Menggali Informasi yang Terpendam"*

Yudho Giri Sucahyo. (2003) " *Penerapan Data Mining Permasalahan apa saja yang bisa diselesaikannya?"*

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135