

IMPLEMENTASI ALGORITMA C 4.5 DALAM PEMBUATAN APLIKASI PENUNJANG KEPUTUSAN PENERIMAAN PEGAWAI CV. DINAMIKA ILMU

Joko Purnomo (ke_piss@yahoo.com)
Wawan Laksito YS (wlaksito@yahoo.com)
Yustina Retno Wahyu U (yustina_retno@yahoo.com)

ABSTRAK

Sebuah perusahaan penerbitan dan percetakan perlu mengadakan penerimaan karyawan secara mandiri dilingkungan perusahaan. Kendala yang sering ditemukan dalam proses penerimaan karyawan di sebuah perusahaan yaitu sulitnya menentukan pelamar mana yang memenuhi kriteria untuk menjadi karyawan sesuai keinginan perusahaan dari sekian banyak pelamar sedangkan pelamar yang diterima menjadi karyawan terbatas. Banyaknya variabel yang digunakan dalam penunjang keputusan penerimaan pegawai di sebuah perusahaan seperti pendidikan terakhir, pengalaman kerja yang sudah pernah dijalani, usia calon karyawan, nilai transkrip ijazah, nilai test tertulis, dan nilai test wawancara. Sehingga sulit bagi sebuah perusahaan untuk mengelompokkan nilai dari tiap-tiap variabel sesuai kriteria sebuah perusahaan yang nantinya digunakan sebagai hasil akhir penerimaan pegawai di sebuah perusahaan. Untuk menentukan calon pegawai yang memenuhi kriteria maka dibutuhkan rekomendasi yang tepat. Perekomendasi yang tepat membutuhkan jumlah data yang banyak, Algoritma C4.5 merupakan salah satu metode yang dapat menganalisis data dalam jumlah yang banyak atau biasa disebut data mining. Pada penelitian ini dibuat suatu aplikasi yang mengimplementasikan pohon keputusan dengan algoritma C4.5, data yang digunakan dalam penelitian ini adalah data calon pegawai. Dalam perancangan aplikasi menggunakan diagram konteks, DFD, dan penyajian aplikasi menggunakan Code Gear RAD Studio Delphi 2009 serta database MySQL. Hasil yang dicapai adalah sebuah Aplikasi Penunjang Keputusan Penerimaan Pegawai yang Implementasinya dapat membantu kepala bagian Human Resource Development (HRD) dalam penyeleksian dan penerimaan calon pegawai.

Kata Kunci : Algoritma C 4.5, Sistem Penunjang Keputusan, Data Mining.

I. PENDAHULUAN

Dengan semakin berkembangnya teknologi dan ilmu pengetahuan, semakin berkembang pula proses pendidikan di Indonesia ini. Menyebabkan banyaknya perusahaan atau instansi-instansi sulit memilih pegawai / karyawan yang sesuai dengan kebutuhan dan ahli pada bidangnya masing-masing. Oleh karena itu menyadari betapa pentingnya memilih calon pegawai yang tepat, maka dirancang program aplikasi sistem pendukung keputusan untuk pemilihan penerimaan pegawai untuk suatu perusahaan yang berbasis aplikasi dekstop. Aplikasi ini dapat mempermudah suatu perusahaan dalam pemilihan atau seleksi calon pegawai yang akan diterima.

Untuk menentukan calon pegawai yang memenuhi kriteria maka dibutuhkan rekomendasi yang tepat. Perekomendasi yang tepat membutuhkan jumlah data yang banyak, Algoritma C4.5 merupakan salah satu

metode yang dapat menganalisis data dalam jumlah yang banyak atau biasa disebut data mining. Pada penelitian ini dibuat suatu aplikasi yang mengimplementasikan pohon keputusan dengan algoritma C4.5, data yang digunakan dalam penelitian ini adalah data calon pegawai.

Algoritma C 4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Misalnya dapat mengolah data numerik kontinyu dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain.

Implementasi Aplikasi yang dibangun diharapkan dapat membantu bagian Human Resource Development (HRD) sehingga membantu dalam penyeleksian dan penerimaan calon pegawai.

II. METODE PENELITIAN

2.1 Metode Pengembangan Sistem Pendukung Keputusan

Pengembangan suatu sistem pendukung keputusan, terutama SPK yang besar, merupakan suatu proses yang rumit. Pengembangan tersebut melibatkan hal-hal teknis (misal pemilihan perangkat keras dan *networking*) sampai perilaku (misal antarmuka pengguna dan dampak potensial SPK terhadap individu, kelompok, dan organisasi keseluruhan).

Prototipe adalah suatu pendekatan yang sangat rapi dan berurutan untuk membuat sebuah sistem menjadi suatu nyata.

Prototipe mempunyai proses design berulang dan mengkombinasikan empat fase utama SDLC (*System Development Life Cycle*) tradisional (analisis, desain, konstruksi dan implementasi) ke dalam satu langkah yang berulang-ulang.

a) Perencanaan

Fase perencanaan dimulai dari kebutuhan dalam penerimaan pegawai di perusahaan yang terdapat suatu masalah yang perlu dipecahkan. Sehingga direncanakan pembuatan sistem pendukung keputusan penerimaan pegawai. Inisiasi proyek melibatkan sebuah sistem yang telah diputuskan, jika kelihatan bermanfaat, maka dilakukan (analisis) studi kelayakan. Studi kelayakan mempertimbangkan apakah gagasan tersebut masuk akal.

b) Analisis

Fase ini melakukan wawancara dan observasi di perusahaan, menanyakan dan menjawab pertanyaan penting seperti siapa pengguna sistem, data-data yang dibutuhkan SPK penerimaan pegawai, jika sebelumnya sudah ada sistem yang berjalan maka sistem dianalisa tersebut bersama untuk mengarah sistem yang baru.

c) Desain

Fase desain menandai bagai mana sistem yang baru akan bekerja, mempertimbangkan semua detail perangkat keras, perangkat lunak, infrastruktur jaringan, antar muka pengguna.

Dalam fase ini, antar muka pengguna, form, display, program dan laporan, database dan file ditetapkan.

d) Implementasi

Fase implementasi adalah pengujian sistem yang telah selesai dibuat dan menverifikasi bahwa sistem bekerja dalam pendukung keputusan penerimaan pegawai.

Untuk pengujian aplikasi ini menggunakan metode *black box*.

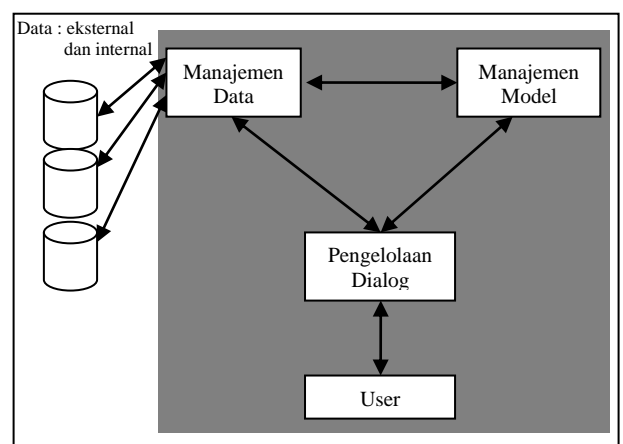
III. TINJAUAN PUSTAKA

3.1 Komponen Sistem Pendukung Keputusan

Sistem pendukung keputusan terdiri atas tiga komponen utama yaitu :

1. Subsistem pengolahan data (database).
2. Subsistem pengolahan model (modelbase).
3. Subsistem pengolahan dialog (user interface).

Hubungan antara ketiga komponen ini dapat dilihat pada gambar 1 [1].



Gambar 1. Hubungan Antara Tiga Komponen Sistem Penunjang Keputusan

3.2 Data Mining

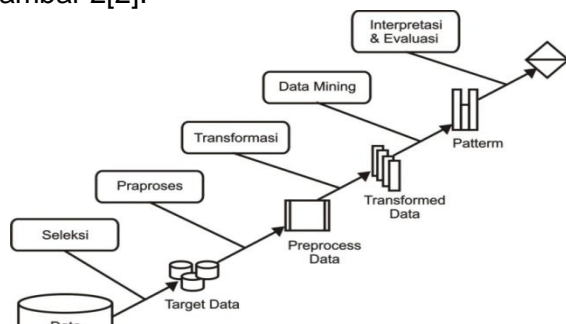
Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* (*mesin pembelajaran*) untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa data mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari disiplin ilmu, data mining bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani :

- Jumlah data yang sangat besar
- Dimensi yang tinggi
- Data yang heterogen dan berbeda sifat.

Data-data yang ada, tidak dapat langsung diolah dengan menggunakan sistem data mining. Data-data tersebut harus dipersiapkan lebih dulu agar hasil yang

diperoleh lebih maksimal, dan waktu komputasinya lebih minimal. Proses persiapan data ini sendiri dapat mencapai 60% dari keseluruhan proses dalam data mining. Adapun tahapan-tahapan yang harus dilalui dalam proses data mining disajikan pada gambar 2[2].



Gambar 2. Data Mining sebagai Tahapan dalam Penemuan Pengetahuan

Terdapat beberapa teknik data mining yang sering disebut-sebut dalam literatur. Namun ada 3 teknik data mining yang populer, yaitu [2][3]:

a. *Association Rule Mining*

Association Rule Mining (Asosiasi Peraturan Pertimbangan) adalah teknik mining untuk menemukan asosiatif antara kombinasi atribut. Contoh dari aturan asosiatif dari analisa pembelian di suatu pasar swalayan dapat mengatur penempatan barangnya atau merancang strategi pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

b. *Pengklusteran (Pengelompokkan)*

Berbeda dengan *association rule mining* dan klasifikasi dimana kelas data telah ditentukan sebelumnya, *pengklusteran* dapat dipakai untuk memberikan label pada kelas data yang belum diketahui. Karena itu *pengklusteran* sering digolongkan sebagai metode unsupervised learning. Prinsip *pengklusteran* adalah memaksimalkan kesamaan antar kluster. *Pengklusteran* dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensi.

c. *Klasifikasi*

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, pendapatan rendah.

3.3 Algoritma C.45

Pencarian Entropy dan Gain menggunakan rumus, untuk menghitung *gain* digunakan rumus sebagai berikut :

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

Keterangan :

S : himpunan kasus

A : Atribut

n : jumlah partisi atribut

|S_i| : jumlah kasus pada partisi ke –i

|S| : jumlah kasus dalam S

Rumus dasar dari *entropy* tersebut adalah sebagai berikut:

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

Keterangan :

S : Himpunan kasus

A : Fitur

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

IV. HASIL DAN PEMBAHASAN

4.1. Pemberian Bobot Setiap Variabel

Pemberian bobot pada nilai-nilai variabel yang akan digunakan sebagai formulasi penyeleksian penerimaan calon pegawai CV. Dinamika Ilmu. Prosentase 100% dibagi jumlah variabel yang digunakan, kemudian hasil pembagian tersebut kemudian dibagi untuk menentukan setiap nilai variabel. Prosentase nilai variabel dalam penerimaan calon pegawai sudah ditentukan sesuai kebijakan direksi, dari bobot variabel yang paling tinggi hingga bobot variabel yang paling rendah. Adapun pemberian bobot dari setiap variabel sebagai berikut :

a. **Pemberian bobot nilai variabel pendidikan**

Tabel 1. Pemberian bobot nilai variabel pendidikan

Pendidikan	Bobot
S1	8%
D3	7%
SMA	5%

b. **Pemberian bobot nilai variabel usia**

Tabel 2. Pemberian bobot nilai variabel usia

Usia	Kasifikasi	Bobot
17-35	Produktif	10%
>35 - 40	Tidak Produktif	6%

c. Pemberian bobot nilai variabel pengalaman

Tabel 3. Pemberian bobot nilai variabel pengalaman

Pengalaman	Bobot
Ya	10%
Tidak	6%

d. Pemberian bobot nilai variabel nilai transkrip

Tabel 4. Pemberian bobot nilai variabel nilai transkrip(SMA).

Nilai (Untuk SMA)	Klasifikasi	Bobot
55 – 65	D	2%
> 65 – 75	C	3%
> 75 – 85	B	5%
> 85 – 100	A	6%

Tabel 5. Pemberian bobot nilai variabel nilai transkrip (D3 dan S1)

Nilai (Untuk D3 dan S1)	Klasifikasi	Bobot
2 – 2,5	D	2%
> 2,5 - 3	C	3%
>3 – 3,5	B	5%
> 3,5 - 4	A	6%

e. Pemberian bobot nilai variabel nilai test tertulis

Tabel 6. Pemberian bobot nilai variabel nilai test tertulis

Nilai Test Tertulis	Klasifikasi	Bobot
55 – 65	D	2%
> 65 – 75	C	3%
> 75 – 85	B	5%
> 85 – 100	A	6%

f. Pemberian bobot nilai variabel nilai wawancara

Tabel 7. Pemberian bobot nilai variabel nilai test tertulis

Nilai Wawancara	Bobot
Baik	10%
Buruk	6%

Formulasi penyeleksian pegawai:

$$\text{Pegawai baru} = \frac{(v1 + v2 + v3 + v4 + v5 + v6)}{\text{jumlah variabel} \%}$$

Keterangan :

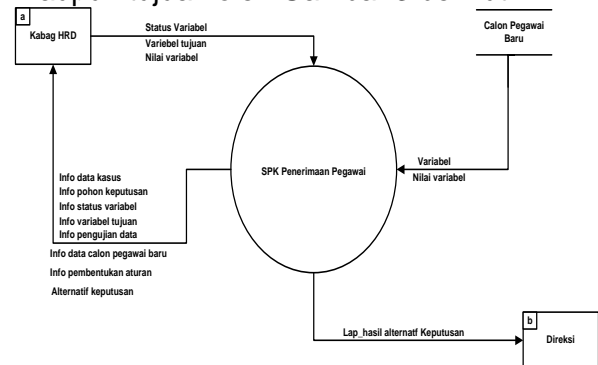
V-[n] : nilai bobot variabel

Calon pegawai dinyatakan diterima apabila hasil dari formula ≥ 6.5

4.2. Perancangan Sistem

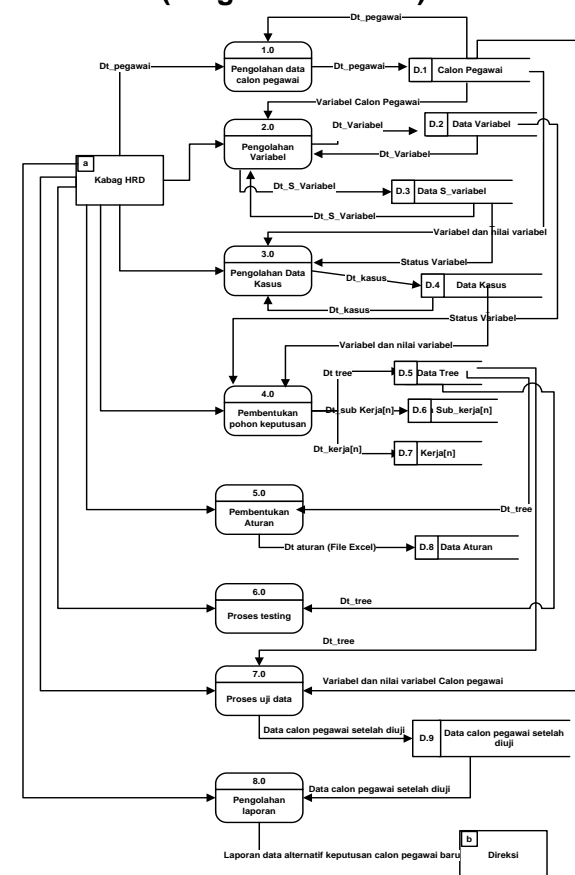
a. Diagram Konteks

Diagram konteks merupakan salah satu alat bantu dalam melakukan analisis terstruktur. Diagram konteks ini menggambarkan suatu sistem secara garis besarnya atau keseluruhannya saja. Dalam diagram konteks juga digambarkan entitas eksternal yang merupakan *brainware* yang menghasilkan data yang akan diolah sistem maupun tujuan oleh Gambar 3 berikut ini.



Gambar 3. Diagram Konteks

b. DAD (Diagram Alir Data)

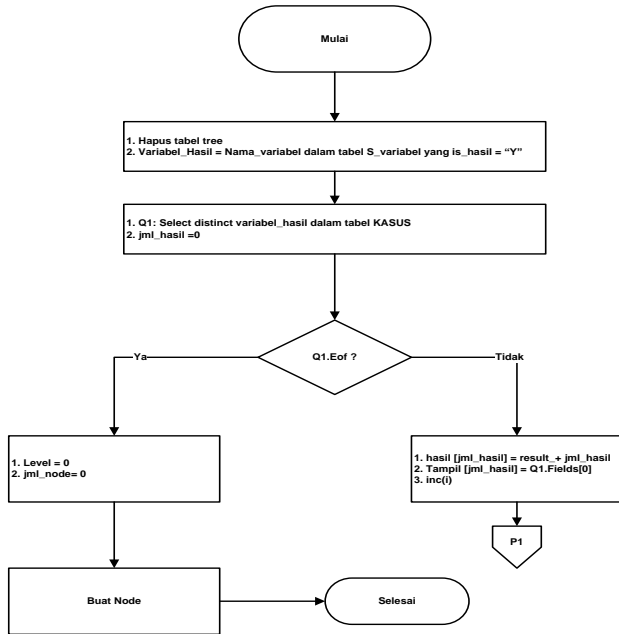


Gambar 4. DAD Level 0

4.3. Implementasi

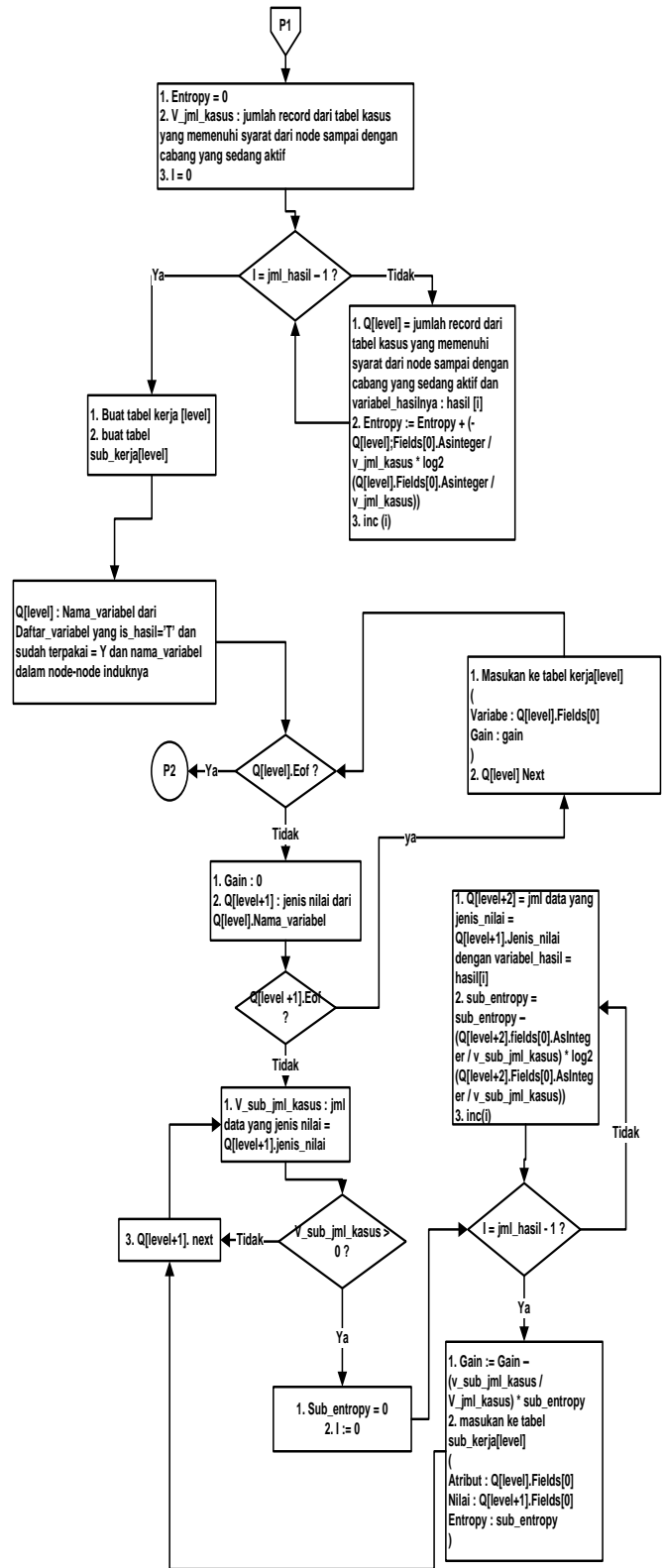
4.3.1 Implementasi Algoritma C4.5

Dalam Algoritma C4.5 pembentukan pohon keputusan dilakukan secara rekursif. Inisialisasi awal pembentukan pohon keputusan ditunjukkan gambar 5.

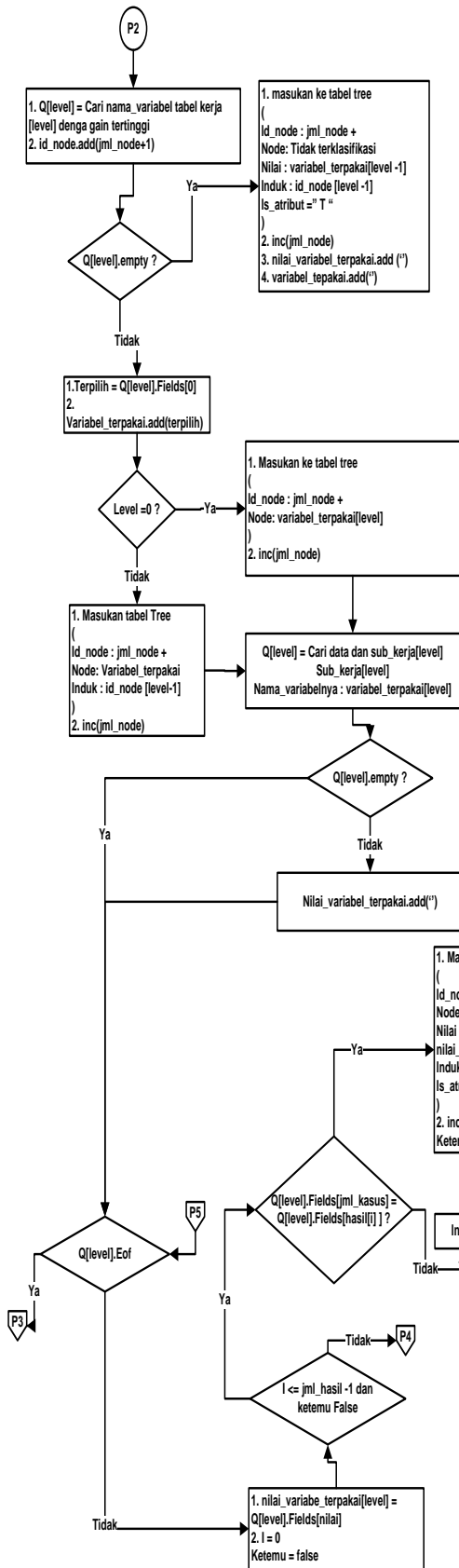


Gambar 5. Algoritma C4.5 inisialisasi pembentukan node

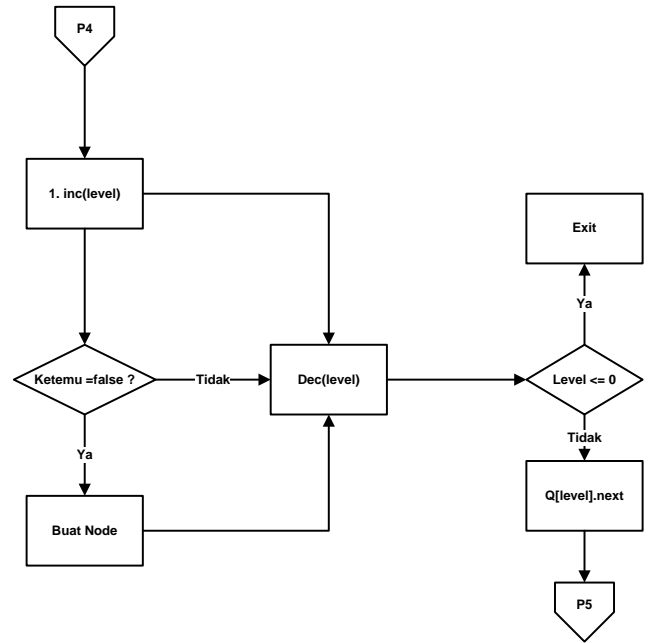
Salah satu langkah pada algoritma inisialisasi pembentukan node adalah buat node. Algoritma prosedur buat node 1 dan node 2 ditunjukkan pada Gambar 6 dan Gambar 7. Sedangkan algoritma prosedur untuk node 3, node 4, dan node 5 ditunjukkan pada Gambar 8, Gambar 9, dan Gambar 10.



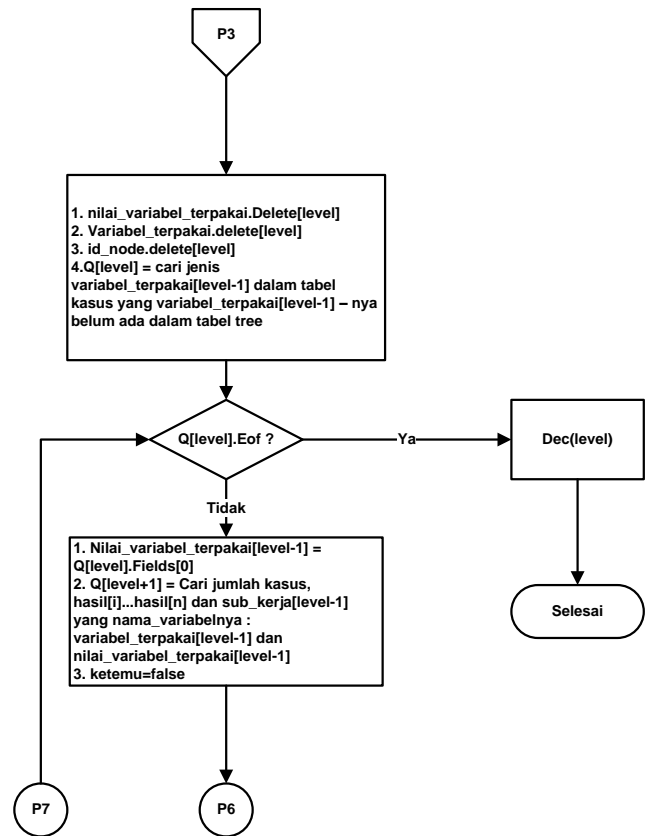
Gambar 6. Algoritma pembentukan node 1



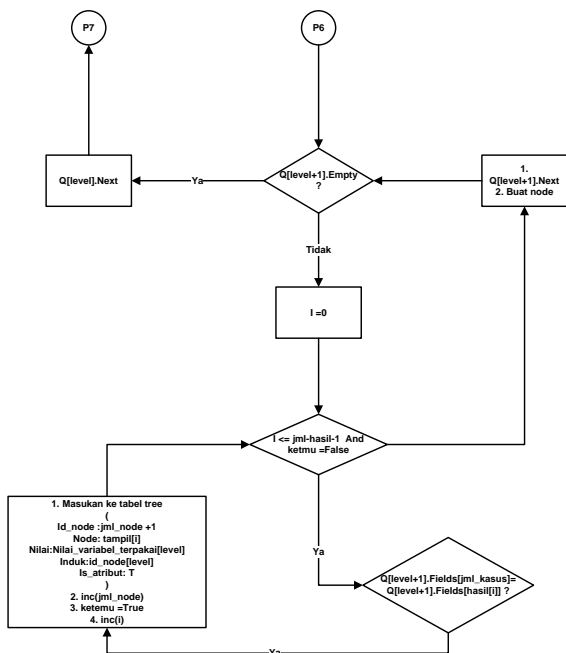
Gambar 7. Algoritma pembentukan node 2



Gambar 9. Algoritma pembentukan node 3



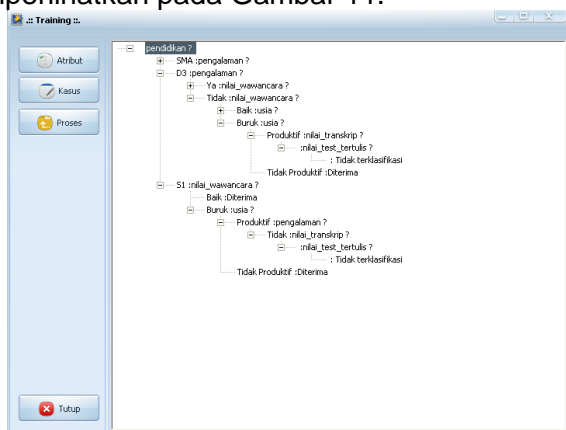
Gambar 11. Algoritma pembentukan node 4



Gambar 10. Algoritma pembentukan node 5

4.3.2 Implementasi Antar Muka

Pohon Keputusan yang dihasilkan sistem dari penilaian beberapa atribut diperlihatkan pada Gambar 11.



Gambar 11. Pohon Keputusan

Tampilan halaman Data Calon Pegawai ditunjukkan pada Gambar 12.

Gambar 12. Tampilan Halaman Data Calon Pegawai

Tampilan halaman Data Calon Pegawai ditunjukkan pada Gambar 13.

No	Aturan
1	Jika pendidikan = SMA Dan pengalaman = Ya Dan usia = Produktif Dan nilai_wawancara = Baik Dan nilai_transkrip = Dan nilai_test_tertulis = Maka Tidak Terklasifikasi
2	Jika pendidikan = SMA Dan pengalaman = Ya Dan usia = Produktif Dan nilai_wawancara = Buruk Maka Tidak Diterima
3	Jika pendidikan = SMA Dan pengalaman = Ya Dan usia = Tidak Produktif Dan nilai_wawancara = Baik Dan nilai_transkrip = Dan nilai_test_tertulis = Maka Tidak Terklasifikasi
4	Jika pendidikan = SMA

Gambar 13. Tampilan Halaman Aturan

4.4. Pengujian Sistem

4.4.1 Pengujian algoritma C 4.5

a. Hasil perhitungan entropy node 1

Hasil perhitungan entropy dan gain node 1 menggunakan aplikasi dan hasil perhitungan secara manual, dengan menghitung jumlah kasus **Diterima** dan **Tidak Diterima** dengan berdasarkan variabel **Pendidikan, Pengalaman, Usia, Nilai Transkrip, Nilai Test Tertulis** dan **Nilai Wawancara** dapat dilihat pada Tabel 2.

b. Hasil Perhitungan Entropy dan Gain node 1.2

Hasil perhitungan entropy dan gain node 1.2 menggunakan aplikasi dan hasil perhitungan secara manual, dengan menghitung jumlah kasus **Diterima** dan **Tidak Diterima** dengan berdasarkan variabel **Pendidikan, Usia, Nilai Transkrip, Nilai Test Tertulis** dan **Nilai Wawancara** dapat dilihat pada Tabel 3.

4.4.2 Pengujian Aplikasi

Pengujian adalah pengujian terhadap fungsi-fungsi yang ada dalam sistem, apakah fungsi tersebut berfungsi sesuai yang diharapkan atau tidak. Pengujian aplikasi sistem pendukung keputusan penerimaan pegawai menggunakan metode *black box* dengan memberikan kuisioner kepada responden yaitu kabag HRD, pegawai perusahaan dan mahasiswa dapat dilihat Tabel 4.

Tabel 2. Hasil perhitungan *entropy* dan *Gain* node 1

Node	Nama Variabel	Nilai	Kasus	Res ult1	Res ult2	Entropy	Gain
1	Total		93	61	32	0.9286	
	Pendidikan						0.0943
		SMA	33	19	14	0.9834	
		S1	21	17	4	0.7025	
		D3	39	30	9	0.7793	
	Usia						0.1229
		Tidak Produktif	38	16	22	0.9819	
		Produktif	55	45	10	0.6840	
	Pengalaman						0.1530
		Ya	55	46	9	0.6429	
		Tidak	38	15	23	0.9876	
	Nilai Transkrip						0.0537
		A	27	21	6	0.7642	
		B	27	20	7	0.8256	
		C	15	9	6	0.9710	
		D	24	11	13	0.9950	
	Nilai Test Tertulis						0.1206
		A	31	21	10	0.9072	
		B	20	17	3	0.6098	
		C	26	19	7	0.8404	
		D	16	4	12	0.8113	
	Nilai Wawancara						0.1031
		Baik	61	48	13	0.7474	
		Buruk	32	13	19	0.9950	

Tabel 3. Hasil Perhitungan *entropy* dan *gain* node 1.2

Node	Nama Variabel	Nilai	Kasus	Res ult1	Res ult2	Entropy	Gain
1.2	Pengalaman : Tidak		38	15	23	0.9876	
	Pendidikan						0.0634
		SMA	14	3	11	0.7496	
		S1	9	5	4	9911	
		D3	15	7	8	9968	
	Usia						0.4454
		Tidak Produktif	16	0	16	0	
		Produktif	22	15	7	0.9024	
	Nilai Transkrip						0.2439
		A	15	9	6	0.9710	
		B	13	6	7	0.9957	
		C	4	0	4	0	
		D	6	0	6	0	
	Nilai Test Tertulis						0.1749
		A	13	5	8	0.9612	
		B	8	5	3	0.9544	
		C	10	5	5	1	
		D	7	0	7	0	
	Nilai Wawancara						0.0243
		Baik	16	8	8	1	
		Buruk	22	7	15	0.9024	

Tabel 4. Item Uji

Item Uji	Detail Pengujian
Pengolahan data calon pegawai	1. Input data calon pegawai. 2. Ubah data calon pegawai 3. Hapus data calon pegawai
Pengaturan Variabel	1. Pilih Variabel dari data calon pegawai. 2. pilih Variabel tujuan
Data kasus	1. Variabel aktif yang muncul
Pembentukan pohon keputusan	1. Proses pembentukan pohon keputusan
Testing	1. Pengujian hasil pohon keputusan
Pembentukan aturan	1. Pembentukan daftar aturan hasil pohon keputusan. 2. Penyimpanan daftar aturan dalam format Excel
Uji data	1. Pengujian data calon pegawai baru dengan mencocokkan nilai variabel dengan aturan yang sudah dibuat 2. Simpan tabel calon pegawai
Data hasil pengujian	1. Cetak laporan alternatif keputusan calon pegawai

Hasil pengujian sistem penunjang keputusan penerimaan pegawai yang di uji langsung oleh kabag HRD, kesimpulan yang didapat adalah sebagai berikut :

1. Algoritma C.45 dapat digunakan dalam system pendukung keputusan penerimaan pegawai.
2. Data calon pegawai dapat diinputkan, diedit, dan dihapus sesuai dengan kebijakan kabag HRD, sehingga dapat mempermudah perusahaan dalam merekap data calon pegawainya.
3. Variabel untuk menunjang keputusan penerimaan pegawai dapat diubah nilainya dan dapat menambahkan variabel, sehingga mempermudah perusahaan dalam menilai calon pegawai sesuai variabel-variabel yang digunakan.

V. PENUTUP

5.1. Kesimpulan

Setelah melakukan perancangan Aplikasi Penunjang Keputusan Penerimaan Pegawai pada sebuah perusahaan, maka dihasilkan sebuah program aplikasi berbasis komputer. Program aplikasi penunjang keputusan penerimaan pegawai ini dibuat sesuai dengan kebijaksanaan dan permintaan dari pihak perusahaan untuk mempermudah kegiatan penerimaan pegawai di perusahaan

tersebut. Setelah perancangan sistem tersebut, maka dapat disimpulkan :

1. Algoritma C4.5 dapat diimplementasikan dalam pembuatan aplikasi sistem penunjang keputusan penerimaan pegawai.
2. Pohon keputusan untuk mengklasifikasikan calon pegawai baru dengan variabel tujuan diterima atau tidak diterima dibentuk dengan algoritma C4.5.
3. Informasi yang dihasilkan berupa klasifikasi calon pegawai baru, laporan data aturan dan laporan data calon pegawai baru.
4. Aplikasi mudah dipahami dan digunakan oleh kepala bagian HRD.
5. Aplikasi dapat digunakan untuk alternatif keputusan penerimaan calon pegawai baru.
6. Membantu pihak Kabag HRD dalam menentukan pegawai yang tepat dari sejumlah calon pegawai yang mengirimkan lamaran ke Perusahaan.

- [4] Ramdhani, M.A dan Suryadi, K., "Sistem Pendukung Keputusan", PT. Remaja Rosdakarya, 2003.

5.2. SARAN

Berdasarkan hasil dari kesimpulan diatas, penulis menyarankan beberapa hal sebagai berikut:

1. Pembuatan program ini masih dapat dikembangkan seiring dengan berkembangnya spesifikasi kebutuhan pengguna, terutama dalam hal tampilan ada baiknya dibuat tampilan yang lebih menarik dan dikembangkan lebih lanjut.
2. Masih banyak fasilitas lain yang dapat dikembangkan dalam perangkat lunak ini, pengembangan tersebut tentunya dapat meningkatkan mutu perangkat lunak yang lebih baik serta sesuai dengan tuntunan dan kebutuhan yang bisa dipenuhi sebuah perusahaan.

DAFTAR PUSTAKA

- [1] Turban, Efraim & Aronson, Jay E., "Decision Support Systems and Intelligent Systems", 6th edition, Prentice Hall, Upper Saddle River, NJ, 2006
- [2] Kusriani dan Emha Taufiq Luthfi, "Algoritma Data Mining", Andi Offset, Yogyakarta, 2009.
- [3] Fairuz El Said, 2009. Data Mining – Konsep Pohon Keputusan, diakses dari "<http://fairuzelsaid.wordpress.com/2009/11/24/data-mining-konsep-pohon-keputusan/>", pada tanggal 12 Maret 2010.