



PREDIKSI PENGANGKATAN KARYAWAN DENGAN METODE ALGORITMA C5.0 (STUDI KASUS PT. MATARAM CAKRA BUANA AGUNG

Ismasari Nawangsih¹, Indri Melani², Sifa Fauziah

Program Studi Teknik Informatika, Universitas Pelita Bangsa
Jl. Inspeksi Kalimalang Tegal Danas Arah DELTAMAS, Cikarang Pusat-Kab. Bekasi,
Indonesia

ismasari.n@pelitabangsa.ac.id

Abstrak

In the decision to appoint permanent employees that can be made at the company is determined in terms of attendance and good discipline. Determination of the appointment of permanent employees at PT. Mataram Cakra Buana Agung still looks imprecise and takes a long time. Then a study was conducted that aims to determine the prediction information on permanent employee appointments by looking at the criteria set by the company using the C5.0 classification algorithm with the decision tree method. The data used in this study are employee data owned by PT. Mataram Cakra Buana Agung as many as 403 data, the process of testing the method using Rapid Miner9.5. Based on the results of testing on research in predicting employee appointments at PT. Mataram Cakra Buana Agung, obtained the result from the C5.0 algorithm or decision tree, which is an accuracy of 90.00%. So it can be concluded that the C5.0 algorithm technique with the decision tree method is considered successful in predicting employee appointments at PT. Mataram

Informasi Artikel

Diterima: 20-08-2021

Direvisi: 19-09-2021

Dipublikasikan: 26-10-2021

Keywords:

Employee recruitment, Classification, C5.0 Algorithm.

I. Pendahuluan

Karyawan Kontrak adalah Perjanjian kerja antara pekerja dengan pengusaha untuk mengadakan hubungan kerja dalam waktu tertentu atau dapat juga disebut dengan Perjanjian Kerja Waktu Tertentu (PKWT) yang didasarkan suatu jangka waktu yang diadakan paling lama 2 tahun dan hanya dapat di perpanjang 1 kali untuk jangka waktu maksimal 1 tahun (Undang – Undang RI ketenagakerjaan 2003 dalam pasal 59 ayat 1).

Karyawan Tetap adalah aset utama perusahaan yang menjadi perencana dan pelaku aktif dari aktivitas organisasi. Karyawan adalah seorang pekerja yang bekerja dibawah perintah orang lain dan mendapat kompensasi serta jaminan (Hasibuan, 2006). Perjanjian Pekerja atau buruh dengan pengusaha untuk mengadakan hubungankerja yang bersifat tetap atau dapat juga disebut dengan Perjanjian Kerja Waktu Tidak Tertentu (PKWTT). Pekerja atau/buruh adalah setiap orang yang bekerja dengan menerima upah atau imbalan dalam bentuk lain, tertuang dalam Undang-Undang RI No.13 Tahun 2003 tentang Ketenagakerjaan dalam pasal 1 ayat 3.

Karyawan dan perusahaan merupakan dua hal yang sangat bergantung dan saling membutuhkan satu sama lain. Karyawan membutuhkan perusahaan sebagai tempat dimana ia dapat bekerja serta memperoleh imbalan yang dapat digunakan untuk memenuhi kebutuhan-kebutuhan hidupnya. Perusahaan membutuhkan karyawan sebagai sumber daya manusia yang menggerakkan roda perusahaan sehingga perusahaan dapat melakukan produksi dan menghasilkan produk dan kinerja yang memiliki kualitas yang baik. Suatu perusahaan akan berkembang atau tidak sangat bergantung pada kualitas sumber daya manusia yang menggerakannya karena setiap karyawan dalam perusahaan memiliki arti dan perannya masing-masing dimana aktivitas individual serta kegiatan sosial mereka mempengaruhi fungsi dari perusahaan tersebut.

Karyawan kontrak merupakan karyawan

yang diperbantukan untuk menyelesaikan pekerjaan-pekerjaan rutin perusahaan, dan tidak ada jaminan kelangsungan masa kerjanya. Dalam kelangsungan masa kerja karyawan kontrak ditentukan oleh prestasi kerja dan kinerjanya. Semakin bagus prestasi kerjanya, karyawan kontrak akan dipertahankan oleh perusahaan, namun jika prestasi kerjanya tidak ada peningkatan maka perusahaan akan memberhentikan karyawan tersebut. Kewajiban kerja karyawan kontrak terkadang hampir sama atau bahkan lebih berat dari pada karyawan tetap. Namun dari segi gaji atau fasilitas lainnya tentu saja sangat berbeda, termasuk tidak adanya ketentuan pesangon yang jelas apabila perusahaan tidak lagi menggunakan jasa si tenaga kerja kontrak. Beberapa kriteria atau syarat dari Pimpinan perusahaan untuk menjadi dasar pengambilan keputusan antara lain kinerja, kedisiplinan, loyalitas, dan pengalaman kerja Karyawan. Jika karyawan kontrak tersebut telah memenuhi nilai kriteria-kriteria yang sudah ditentukan maka karyawan kontrak tersebut dapat ditentukan sebagai karyawan tetap pada PT. Mataram Cakra Buana Agung.

Permasalahan yang ada dalam PT. Mataram Cakra Buana Agung dalam penentuan karyawan kontrak menjadi karyawan tetap selama ini dilakukan secara bertahap dengan penyeleksian berkas, seperti tes lisan, tes fisik, atau tertulis, wawancara dan lain sebagainya. Cara ini cukup membutuhkan waktu yang lama, sehingga kemungkinan akan sering terjadinya kesalahan dalam hasil akhir dari penentuan kariyawan kontrak menjadi karyawan tetap sering tidak memenuhi kriteria yang dibutuhkan perusahaan dan menghambat kinerja perusahaan.

Karena tugas ini melibatkan banyak managerial untuk memutuskan karyawan yang tepat bagi perusahaan, dengan keputusan tersebut sangatlah tidak mudah, karena perlu pertimbangan dari berbagai indokator-indikator seperti kinerja materi, kinerja lapangan, absensi,

kehadiran, loyalitas, dan kedisiplinan dengan demikian penulis ingin menerapkan sebuah teknik data mining untuk membantu dalam proses pengangkatan karyawan tetap. performansi algoritma decision tree C5.0, CHART, dan CHAID : kasus prediksi status resiko kredit di Bank X” menyatakan bahwa algoritma C5.0 menghasilkan tingkat keakuratan yang lebih tinggi dalam hal prediksi. Untuk itu dalam memprediksi permintaan pengambilan keputusan untuk II. pengangkatan karyawan tetap yaitu dengan metode klasifikasi algoritma C5.0 adalah didasarkan pada konsepnya dimana, dari beberapa faktor-faktor yang menjadi atribut penunjang dalam memprediksi pengangkatan karyawan tetap. Konsepnya sederhana dan mudah dipahami, komputasinya efisien, dan memiliki kemampuan untuk mengukur kinerja relatif dari alternatif-alternatif keputusan dalam bentuk matematis yang sederhana.

1.1 Identifikasi Masalah

Berdasarkan latar belakang masalah diatas, maka dapat diidentifikasi beberapa masalah sebagai berikut :

- 1.Data Penilaian Kerja untuk masing-masing karyawan masih disimpan dalam bentuk arsip hard copy sehingga terjadi kesulitan mencari data karna tidak akuratnya penyimpanan sehingga membutuhkan waktu yang lama dalam pencarian.
- 2.Perlu adanya informasi suatu prediksi untuk memudahkannya pengangkatan karyawan tetap.
- 3.Belum adanya suatu metode dalam memprediksi pengangkatan pada PT. Mataram Cakra Buana Agung.
- 4.Proses penilaian kinerja karyawan masih secara manual, sehingga membutuhkan waktu yang lama.

1.2 Batasan Masalah

Dari identifikasi masalah diatas maka penelitian ini bagaimana melihat dan mengukur metode klasifikasi algoritma C5.0 untuk mendapatkan keputusan dalam

pengangkatan karyawan pada PT. Mataram Cakra Buana Agung.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk memperoleh informasi mengenai prediksi pengangkatan karyawan dengan menerapkan metode klasifikasi algoritma C5.0 menggunakan data karyawan berdasarkan penilaian kriteria-kriteria pengangkatan dan mempermudah proses dalam pengangkatan karyawan tetap.

II. Metodologi

2.1 Definisi Data Mining

Data *mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. Data *mining* adalah proses yang menggunakan teknik *statistic*, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. [1]

2.2 Tahapan Proses Data Mining

Data mining merupakan salah satu dari rangkaian *Knowledge Discovery inDatabase* (KDD). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data.Serangkaian proses tahapan sebagai berikut (Vulandari, 2017) :

1. Data Selection

Langkah pertama adalah dengan melakukan seleksi data, data yang ada akan dilakukan pemilihan data dan atribut yang akan digunakan dalam proses selanjutnya. Pemilihan (seleksi) data dari sekumpulan data operasional dilakukan sebelum tahap penggalian informasi. Data hasil seleksi akan digunakan untuk proses data mining, dan disimpan dalam berkas yang terpisah dari data operasional.

2. Pre-Processing (Cleaning)

Pada tahap ini, data hasil seleksi akan dilakukan pemrosesan pendahuluan dan pembersihan data. Sebelum

proses data mining dilakukan, perlu dilakukan *cleaning* dengan tujuan untuk membuang duplikasi data, menghilangkan data yang tidak konsisten dan noise.

3. Transformation

Merupakan proses transformasi pada data yang telah dipilih untuk diubah menjadi bentuk yang sesuai untuk di-mining.

4. Data mining

Data mining adalah proses mencari suatu informasi atau pola yang menarik pada data yang telah dipilih menggunakan teknik atau metode tertentu. Pemilihan metode atau algoritma sangatlah penting karena sangat berpengaruh pada tujuan dan proses KDD secara keseluruhan.

5. Interpretation (Evaluation)

Pada tahap ini, dilakukan identifikasi pola-pola yang dihasilkan dari data mining. Pola informasi yang dihasilkan perlu ditampilkan kedalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahapan ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.3 Teknik Data Mining

Dengan definisi Data Mining yang luas, maka ada beberapa kelompok berdasarkan tugas yang dapat dilakukan.

- A. Asosiasi (*Association*)
- B. Klusteran (*Clustering*)
- C. Prediksi (*Prediction*)
- D. Estimasi (*Estimation*)
- E. Klasifikasi (*Classification*)

2.4 Pengertian Algoritma C5.0

Algoritma C5.0 tetap dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena

karakteristik data yang diklasifikasi dapat diperoleh dengan jelas baik dalam bentuk struktur pohon keputusan maupun aturan *if-then*, sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan. [2]

Algoritma ini memiliki fitur penting yang membuat algoritma C5.0 ini lebih unggul dari algoritma terdahulunya, dan fitur tersebut antara lain :

1. Algoritma C5.0 dirancang untuk menganalisis basis data substansial yang berisi puluhan sampai ratusan *record* dan satuan hingga ratusan field numeric dan nominal. Memaksimumkan tingkat penafsiran pengguna terhadap hasil yang disajikan dalam dua bentuk yaitu menggunakan pohon keputusan (*decision tree*) dan sekumpulan aturan *IF-Then* yang mudah dimengerti.
2. Algoritma C5.0 mudah digunakan dan tidak membutuhkan pengetahuan yang tinggi tentang statistik atau *machine learning*.
3. Algoritma ini dimulai dengan semua data yang dijadikan akar dari pohon keputusan sedangkan atribut yang dipilih akan menjadi pembagi bagi sampel tersebut.

2.5 Pengertian klasifikasi

Klasifikasi yaitu proses menemukan model atau fungsi yang menjelaskan pengelompokan sesuatu berdasarkan kelas-kelas tertentu, yang bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui.

2.6 Pengertian Confusion Matrix

Confusion matrix adalah tool yang digunakan untuk evaluasi model klasifikasi untuk memperkirakan objek yang benar atau salah. Sebuah matrix dari prediksi yang akan dibandingkan dengan kelas yang asli dari inputan atau dengan kata lain berisi informasi nilai actual dan prediksi pada klasifikasi.

$$\text{Info}(D) = - \sum_{i=1}^m p_i \log_2(p_i)$$

2.7 Pengertian prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi dimasa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti tentang kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Kemampuan membuat prediksi adalah kemampuan perkiraan kejadian atau manfaat yang akan datang berdasarkan hasil observasi, pengukuran, data yang telah dikumpulkan maupun hasil penelitian yang diperlihatkan kecenderungan gejala tertentu.[1]

2.8 Area Under Curve (AUC) dan Receiver Operating Characteristic (ROC)

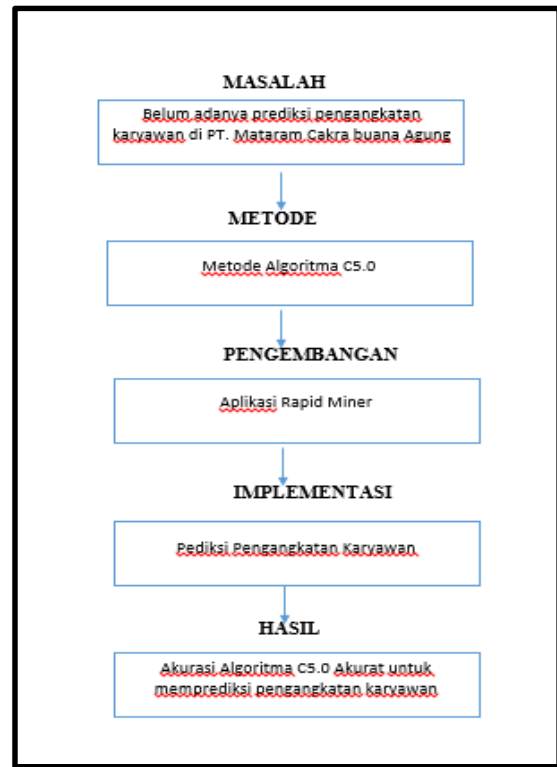
AUC (Area Under Curve) adalah luas daerah di bawah kurva ROC. Apabila nilainya mendekati satu, maka model yang didapat lebih akurat.

Grafik kurva ROC (*Receiver Operating Characteristic*) digunakan untuk mengevaluasi akurasi *classifier* dan untuk membandingkan klasifikasi yang berbeda model (Vercellis, 2009). [8]

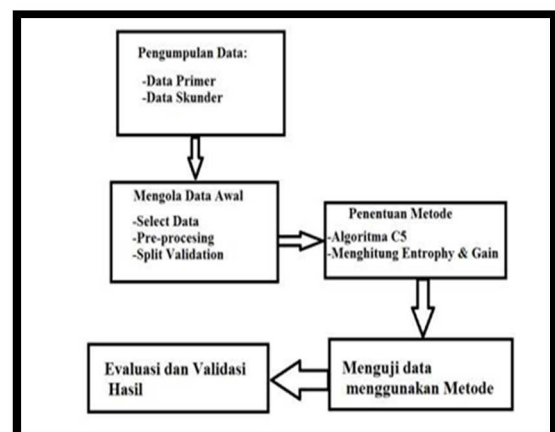
Rapid Miner merupakan salah satu software yang digunakan dalam pengolahan data mining.

2.9 Kerangka Berfikir

Dalam proses pengangkatan karyawan kontrak menjadi karyawan tetap butuh banyak step dan waktu untuk bisa menentukan layak dan tidak menjadi karyawan tetap, metode tersebut. Untuk pengembangan dan pengujian, metode yang dipakai.



Gambar 1. Kerangka Berfikir



Gambar 2. Tahapan Penelitian

2.10 Pengumpulan Data

1. Sumber Data

a) Data Primer

Data primer adalah data yang diperoleh atau dikumpulkan oleh peneliti secara

langsung dari sumber datanya, teknik yang bisa digunakan peneliti untuk mengumpulkan data primer antara lain observasi, diskusi terfokus.

b) Data Sekunder

Data sekunder adalah data yang diperoleh secara tidak langsung dari objek. Dengan memperoleh data dari studi pustaka (*library research method*). diantaranya buku ataupun penelitian yang membahas Data Mining, jurnal mengenai algoritma C5.0. Data sekunder diperoleh dari berbagai sumber seperti buku, laporan, jurnal, dan lain-lain.

2. Variabel Data

Variabel yang terdapat pada data penelitian ini meliputi variabel input dan variabel output atau target.

a) Variabel input

Variabel input dinyatakan sebagai data sampel dari label. Berikut ini adalah variabel input yang terdapat pada data label yang akan digunakan untuk mengklasifikasi yaitu Nama, Absen, Lama kerja, Kreatifitas, loyalitas, skill, Kinerja Lapangan, dan Disiplin.

b) Variabel Target

Variabel target dinyatakan sebagai hasil dari penelitian yang dilakukan terhadap variabel input. Sedangkan variabel target dari penelitian ini adalah atribut "Pengangkatan Karyawan" yang berisikan label "Ya" dan "Tidak".

2.11 Jenis Data

1. Data kualitatif

Data yang tidak berbentuk angka, atau dengan kata lain data yang berbentuk kalimat, kata, atau gambar. Dalam penelitian ini misalnya latar belakang dan data yang diambil dari dokumen.

2. Data kuantitatif

Data yang berbentuk angka dan dapat dihitung nilainya dengan statistic ataupun matematika secara pasti. Data kuantitatif dapat dibedakan menjadi 2 yaitu data diskrit dan data kontinuum. Yaitu data yang diperoleh dari hasil menghitung. Sedangkan data data kontinuum adalah data yang diperoleh dari hasil pengukuran.

2.12 Pengolahan Data Awal

Select Data

Pada tahap ini akan dilakukan pemilihan variabel data yang akan dianalisis, karena sering ditemukan bahwa tidak semua data yang dibutuhkan dengan mempertimbangkan tujuan penulisan, sehingga diperoleh beberapa variabel yang akan digunakan untuk menjadi masukan variabel input. 403 data record akan diambil 7 variabel, yaitu: absen, lama kerja, Kreatifitas, loyalitas, skill, Kinerja Lapangan, dan Disiplin hasil seleksi akan digunakan dalam proses data mining.

1. Pre-processing

Pada proses *pre-processing* akan dilakukan pembersihan data untuk membuang data yang missing value yaitu data yang tidak konsisten dan juga memperbaiki data yang rusak. Proses pembersihan data dilakukan secara manual untuk memastikan bahwa data yang telah dipilih layak untuk dilakukan proses pemodelan.

2. Split Validation

Split Validation merupakan teknik validasi yang membagi data menjadi dua bagian, sebagian data training dan sebagian data testing.

2.13 Pemodelan

Pada penelitian ini akan dilakukan pemodelan menggunakan Metode Algoritma C5.0 untuk mengolah data Karyawan. Metode Algoritma C5.0 dipilih karena satu kelebihan adalah dapat menangani data numerik dan diskret.

Dengan menghitung konsep *entropy* untuk membentuk pohon keputusan. Data kemudian dihitung menggunakan algoritma sesuai dengan metodenya kemudian dicari hasil akurasi.

• Pengujian Data Metode

Pengujian metode dilakukan untuk mengetahui hasil perhitungan yang dianalisa dan untuk melihat apakah fungsi bekerja dengan baik atau tidak. Setelah data dihitung secara manual, kemudian data diuji menggunakan *tools Rapid Miner* untuk memastikan

bahwa apakah hasil perhitungan manual dengan hasil yang diperoleh *RapidMiner* sesuai atau tidak.

• Evaluasi Validasi Hasil

Evaluasi dapat dilakukan dengan cara mengamati dan menganalisa hasil dari algoritma yang digunakan untuk memastikan bahwa hasil pengujian itu benar atau tidak sesuai dengan pembahasan. Sedangkan, validasi dilakukan dengan mengukur hasil prediksi untuk mengetahui tingkat akurasi, presisi dan recall.

III. Hasil dan Pembahasan

3.1 Perhitungan Algoritma C5.0

Tahapan pengolahan data awal antara lain : *select data*, *transformation* serta akan dilakukan *split validation*.

3.2 Proses *Select Data*

Data yang sudah tersedia selanjutnya pemilihan terhadap parameter yang akan di analisis.

Tabel 1. Select Data

No	Atribut	Indikator	Detail Pengguna
1	Nama	OK	ID
2	Absen	OK	NILAI MODEL
3	Lama Kerja	OK	NILAI MODEL
4	Kreatifitas	OK	NILAI MODEL
5	Loyalitas	OK	NILAI MODEL
6	Skill	OK	NILAI MODEL
7	Kinerja Lapangan	OK	NILAI MODEL

3.3 Proses *Tranformation Data*

Setelah data sudah dipilih maka akan dilakukan tahapan untuk melakukan transformasi terhadap atribut. Selanjutnya atribut data hasil proses *cleaning* data akandi transformasi ke dalam bentuk ordinal agar memudahkan proses pemodelan. Berikut adalah penjelasan proses transformasi atribut :

Tabel 2. Proses *Tranformasi Data*

Nilai	Keterangan
Baik	3
Cukup	2
Kurang	1

Tabel 3. Hasil *Transformasi*

ABSENSI	LAMA KERJA	KREATIFITAS	LOYALITAS	SKILL	KINERJA LAPANGAN	DISIPLIN	HASIL
3	1	3	3	3	3	3	ya
3	2	3	2	1	2	3	ya
3	2	3	2	1	2	3	tdk
3	2	1	2	2	1	2	tdk
2	1	2	2	2	2	2	tdk
...
1	2	2	2	3	3	1	tdk

3.4 *Split Validation*

Split Validation merupakan teknik validasi yang membagi data menjadi dua bagian, sebagian data *training* dan sebagian data *testing*. Data yang sudah disiapkan untuk klasifikasi dibagi menjadi dua menggunakan teknik *sampling random* untuk data *training* (80%) dan data *testing* (20%).

Tabel 4. Data Testing

NO	NAMA	ABSENSI	LAMA KERJA	KREATIFITAS	LOYALITAS	SKILL	KINERJA LAPANGAN	DISIPLIN	HASIL
1	AJAT S	3	1	3	3	3	3	3	YA
2	SUTRISNA	3	2	3	2	1	2	3	YA
3	SUTRISNO	3	2	1	2	2	1	2	TDK
4	IWAN M	3	1	2	2	2	2	2	TDK
5	ANDAKA	2	1	1	2	1	1	1	TDK
6	EDI S	2	2	2	2	2	2	2	YA
7	SUGENG R	2	1	3	1	2	2	1	TDK
8	SAPTO A	3	1	1	1	2	2	1	TDK
9	SARWONO	1	1	1	1	1	1	1	TDK
10	SAWAL S	1	2	2	2	3	3	1	TDK
...
403	ERNA N	1	2	2	2	3	3	1	YA

Tabel 5. Data Training

NO	NAMA	ABSENSI	LAMA KERJA	KREATIFITAS	LOYALITAS	SKILL	KINERJA LAPANGAN	DISIPLIN	HASIL
101	EKA O	3	1	2	2	2	2	2	YA
102	ENDAY D	2	1	1	2	1	1	1	TDK
103	FADLIZ	2	2	2	2	2	2	2	YA
104	JOKO S	2	2	3	1	2	2	1	YA
105	ZAINUDIN L	3	1	1	1	2	2	1	TDK
106	PANDU Y	1	1	1	1	1	1	1	TDK
107	SLAMET T	1	2	2	2	3	3	1	YA
108	INDAH F	2	2	2	2	2	2	2	YA
109	QORRI C	3	1	3	3	3	3	3	YA
...
403	ERNA N	1	2	2	2	3	3	1	YA

3.5 Hasil Pengujian

Pada perhitungan Performa *confusion matrix* harus ada 2 data yaitu data *training* dan data *testing*. Data *training* pada kali ini memakai 403 data dan data *testing* memakai sample 100 data. Sebelum proses *confusion matrix*, maka kita harus menentukan terlebih dahulu prediksi pengangkatan. Jika sudah diketahui,

maka langkah yang harus dilakukan adalah menentukan data aktual dengan data prediksi tersebut.

ROW NO	NAMA	PREDIKSI	CONFIDENCE	CONFIDENCE	ABSENSI	LAMA KERJA	KREATIFITAS	LOYALITAS	SKILL
31	M.SOLEHUDIN	TDK	0.445	0.555	5	2	2	5	3
32	HERIS	YA	1	0	5	2	5	4	1
33	ALI W	TDK	0.446	0.555	2	3	4	5	2
34	MANAP	TDK	0.445	0.555	1	2	4	4	2
35	ZAENAL A	TDK	0.445	0.555	2	1	3	5	1
36	NARUDIN	TDK	0.446	0.555	5	2	2	4	1
37	AGUS D	TDK	0.445	0.555	3	2	1	5	2
38	KARMIN S	YA	1	0	4	1	5	4	3

Gambar3. Perbandingan data actual dan data prediksi

3.6 Proses Data Mining Menggunakan Algoritma C5.0

Berikut adalah langkah-langkah perhitungan nilai Entropy dan Gain pada setiap Attribute yang memiliki hasil YA atau TIDAK

3.7 Perhitungan Nilai Entropy

1. Perhitungan Nilai Entropy Total

Langkah awal dalam Algoritma C5.0 adalah mencari nilai entropy.

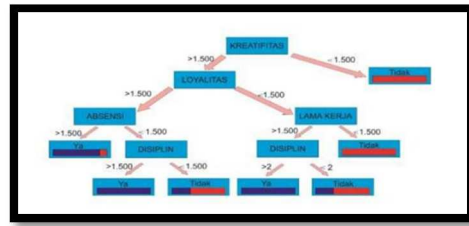
Tabel 6. Perhitungan Data Testing

		Jumlah(s)	Ya(s)	Tidak(s)	Entropy	Gain
Nilai Entropy total		100	44	56	0,989588	
Lama Kerja	Lama	53	32	21	0,9687	0,090977
	Sebentar	47	12	35	0,8196	
Absen	Baik	47	26	21	0,99182	
	Cukup	28	14	14	1	0,084879
	Kurang	25	4	21	0,63431	
Skill	Baik	27	21	6	0,764205	
	Cukup	51	20	31	0,966177	0,164143
	Kurang	22	3	19	0,574636	

2. Perhitungan Nilai Gain

Selanjutnya lakukan perhitungan gain pada setiap attribute berdasarkan pada jumlahkasus per subset attribute.

3. Decision Tree (Pohon Keputusan)



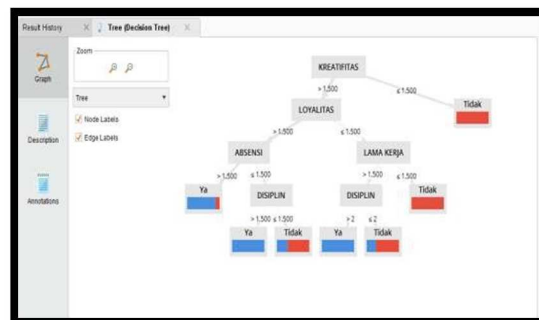
Gambar 4. Pohon Keputusan

Pada gambar diatas dapat disimpulkan bahwa pohon keputusan tersebut memiliki root node Lama Kerja dimana jika Lama kerja lebih dari 1.5 maka dikatakan diangkat karyawan tetap, tetapi jika kurang dari 1.5 maka tidak Diangkat. Pohon keputusan tersebut sesuai dengan pengujian oleh RapidMiner9.5.

4. Hasil Pengujian Rapid Miner

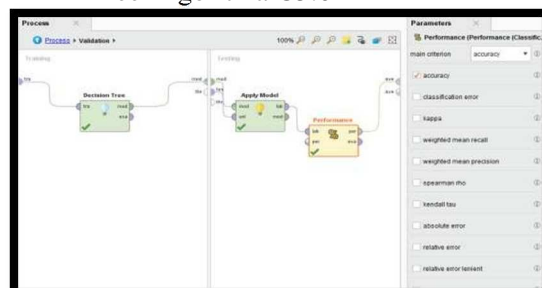
Pada tahap ini metode data mining diterapkan untuk menemukan pengetahuan tersembunyi dan berharga dari data. Metode yang digunakan adalah klasifikasi dengan algoritma C5.0.

Gambar 5. Pohon Keputusan Data Testing



6. Performance Vector

Performance vector adalah suatu cara untuk mengukur performa dari suatu datamining klasifikasi dengan Decision Tree Algoritma C5.0



Gambar 6. Tampilan Setting Validation

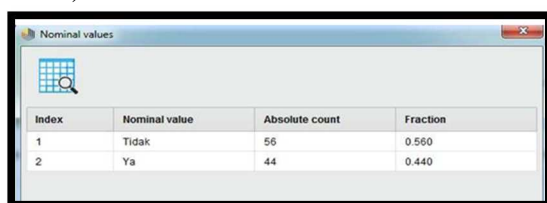
1. Confusion Matrix

Berdasarkan perhitungan akurasi data training dari 100 data, 40 diklasifikasikan prediksi diangkat dan ternyata diangkat, 6 data diprediksi diangkat tetapi ternyata tidak diangkat, 4 data diprediksi tidak diangkat tetapi ternyata diangkat dan sebanyak 50 data diprediksi tidak diangkat dan ternyata tidak diangkat, dengan nilai *accuracy* sebesar 90.00%.

Dari hasil *RapidMiner9.5* didapat nilai *Accuracy* 90.00% dengan 100 data training menghasilkan prediksi *recall* 91.07% dengan prediksi YA 86.96% dan prediksi TIDAK 92.59%.

3.8. Hasil Pengujian ROC dan AUC

AUC (*Area Under Curve*) *Optimistic* Hasil perhitungan divisualisasikan dengan kurva ROC (*Receiver Operating Characteristic*) atau AUC (*Area Under Curve*).



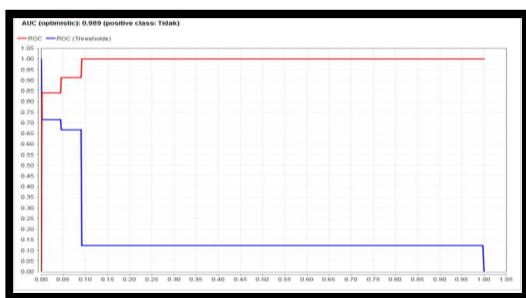
Index	Nominal value	Absolute count	Fraction
1	Tidak	56	0.560
2	Ya	44	0.440

Gambar 7. Grafik AUC (Area Under Curve)

Diketahui nilai dari kurve ROC (*Receiver Operating Characteristic*) diketahui dari SAUC (*Area Under Curve*) *optimistic* pada gambar diatas bernilai 0.989 masuk dalam range ROC nomor 1, itu berarti hasil klasifikasi penelitian ini masuk ke dalam tingkat diagnosa *excellent classification*.

3.9 Analisa Pengujian

Dari Perhitungan Rapid Miner karyawan yang diangkat yaitu 44 karyawan, dan tidak diangkat 56 karyawan. Dengan nilai *Fraction* 0.056 dan 0.044.



Gambar 8. Nominal Value Pengangkatan

Karyawan

IV . Kesimpulan

Dari hasil dan pembahasan dapat diambil beberapa kesimpulan sebagai berikut :

1. Dengan menggunakan metode algoritma C5.0 maka pengangkatan karyawan dapat di prediksi dengan hal tersebut dapat dijadikan kontribusi terhadap proses pengambilan keputusan bagi pihak perusahaan.
2. Proses untuk menghasilkan kombinasi item menggunakan algoritma C5.0 dilakukan dengan menerapkan nilai Entropy dan nilai Gain tertinggi. Dengan didapatnya nilai Entropy dan Gain tertinggi maka dapat membentuk pohon keputusan, dan menunjukkan penentuan atribut mana yang menjadi penentuan pengangkatan karyawan.
3. Evaluasi hasil klasifikasi dalam prediksi pengangkatan karyawan menggunakan data training dan data testing dengan Confusion Matrix dimana data testing dengan tingkat *accuracy* sebesar 91.00% dan data training dengan tingkat *accuracy* 96.75%. Jika melihat dengan kurva ROC menggunakan data training dengan akurasi Excellent Classification sebesar 0.989.

DAFTAR PUSTAKA

- [1]. Hutabarat, C. (2018). Penerapan Data Mining Untuk Memprediksi Permintaan Produk Kartu Perdana Internet Menggunakan Algoritma C5.0 (Studi Kasus: Vidha Ponsel). *Jurnal Pelita Informatika*, 6(4), 419–424.
- [2]. Kurniawan, I., & Saputra, R. A. (2017). Penerapan Algoritma C5.0 Pada Sistem Pendukung Keputusan Kelayakan penerimaan Beras Masyarakat Miskin. *Jurnal Informatika*, 4(@), 236–240.
- [3]. Saw, M. M. (2015). *Implementasi Sistem Keputusan Pengangkatan Karyawan Tetap Pt. Imanuel Surya Utama*. 2(1).
- [4]. Sunge, A. S., (2018). *Prediksi Kompetensi Karyawan Menggunakan Algoritma C4. 5 (Studi Kasus : Pt Hankook Tire Indonesia)*. 2018(Sentika), 23–24.

- [5].Sembiring, I. (2018). *Pengambilan Keputusan Pegawai Tidak Tetap Menjadi Pegawai Tetap Dengan Decission Tree*. 1, 26–37.
- [6].Sudarma, I. M. (2018). *Implementasi Algoritma C5.0 Pada Penilaian*. 17(3), 1–6.
- [7].Susanti, M., Kom, M., & Kom, M. (2018). *Prediksi Pengangkatan Karyawan Kontrak Menjadi Karyawan Tetap Menggunakan Decision Tree Pada Pt. Baskara Cipta Pratama*. 6(1), 1–7.
- [8]Vercellis, C. (2009). *Business Intelligence Data Mining And Optimization ForDecision Making*. United Kingdom : A John Wiley And Sons, Ltd., Publication.
- [9].Wijaya, A. C., Hasibuan, N. A., & Ramadhani, P. (2018). *Implementasi Algoritma C5.0 Dalam Klasifikasi Pendapatan Masyarakat (Studi Kasus : Kelurahan Mesjid Kecamatan Medan Kota)*. 13.
- [10].Tinggi, S., Informatika, M., & Komputer, D. A. N. (2018). *Sistem Pendukung Keputusan Pengangkatan Guru Dan Karyawan Tetap Pada Yayasan Dharma Ibu*.