

Prediksi Promosi Jabatan Karyawan JNE Pematangsiantar Dengan Algoritma C4.5

Rizki Ananda, Sumarno, Indra Gunawan*

Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ¹rzkyanand93@gmail.com, ²sumarno@amiktunasbangsa.ac.id, ^{3,*}indragunawan@amiktunasbangsa.ac.id,

Email Penulis Korespondensi: indragunawan@amiktunasbangsa.ac.id

Submitted: 28/01/2022; Accepted: 29/03/2022; Published: 31/03/2022

Abstrak—Penelitian ini melakukan prediksi promosi jabatan karyawan menggunakan data karyawan dari PT. JNE cabang Pematangsiantar. Dengan tujuan penelitian adalah untuk mendapatkan model aturan dalam mengklasifikasikan prediksi promosi jabatan di perusahaan tersebut. Dengan mengetahui prediksi ini, para pengambil keputusan dapat memilih karyawan yang akan dipromosikan jabatan sesuai dengan aturan yang berlaku di perusahaan. Metode yang digunakan dalam penelitian ini adalah Data Mining Algoritma C4.5, dengan data karyawan yang digunakan sebanyak 20 orang karyawan. Adapun variabel yang digunakan dalam penelitian ini diantaranya : Pengetahuan Pekerjaan, Kompetensi, Kerja Tim, Komunikasi dan Kualitas Pekerjaan. Hasil penelitian diperoleh 9 rules atau model aturan untuk klasifikasi prediksi ini dengan 5 rules dengan hasil akhir Dipromosikan dan 4 rules dengan hasil akhir Tidak Dipromosikan. Dengan tingkat akurasi sebesar 71,43% dan classification error yaitu sebesar 28,57%.

Kata Kunci: Data Mining; Algoritma C4.5; Prediksi; Promosi Jabatan; Karyawan

Abstract—This study predicts employee promotions using employee data from PT. JNE Pematangsiantar. The purpose of this research is to obtain a rule model in classifying predictions of promotions in the company. By knowing this prediction, decision makers can choose employees who will be promoted according to the rules that apply in the company. The method used in this research is Data Mining Algorithm C4.5, with employee data used as many as 20 employees. The variables used in this study include: Job Knowledge, Competence, Teamwork, Communication and Job Quality. The results of the study obtained 9 rules or rule models for this prediction classification with 5 rules with the final result being Promoted and 4 rules with the final result not being Promoted. With an accuracy rate of 71.43% and a classification error of 28.57%.

Keywords: Data Mining; C4.5 Algorithm; Prediction; Job Promotion; Employee

1. PENDAHULUAN

Jalur Nugraha Ekakurir atau biasa dikenal dengan JNE merupakan salah satu perusahaan yang bergerak dalam bidang jasa pengiriman logistik yang berpusat di Jakarta. Dibawah nama resmi yang berlabel Tiki Jalur Nugraha Ekakurir adalah perusahaan pengiriman terbesar di Indonesia. Beberapa unit kegiatan yang dilakukan oleh JNE antara lain urusan kepabeaan, impor barang, dokumen serta pengantaran. Tak hanya melayani pangsa pasar domestik saja, JNE juga melayani urusan pengantaran barang dan dokumen dari luar negeri untuk masuk ke Indonesia.

Seiring dengan perkembangan usaha yang dicapai JNE, pada tahun 1991 perusahaan ini mulai melakukan perluasan wilayah layanannya. JNE mulai memperluas jaringan hingga ke internasional dengan bergabung dengan asosiasi perusahaan-perusahaan yang bergerak dalam bidang yang sama dalam lingkup Asia yang dikenal dengan ACCA. Kelompok perusahaan pengiriman yang berpusat di Hong Kong tersebut selanjutnya memberikan kesempatan bagi JNE untuk "go international" atau berkembang hingga pelosok dunia.

Tuntutan penguasaan teknologi dalam produksi pengiriman dan penerimaan barang juga membutuhkan tenaga yang kompeten. Dengan begitu, PT. JNE Pematangsiantar juga harus mampu melaksanakan pemilihan karyawan yang efektif guna mendapatkan SDM yang sesuai dengan visi dan misi perusahaan.

Sumber daya manusia (SDM) adalah individu produktif yang bekerja sebagai penggerak suatu organisasi, baik itu di dalam institusi maupun perusahaan yang memiliki fungsi sebagai aset sehingga harus dilatih dan dikembangkan kemampuannya. Sumber daya manusia (SDM) adalah salah satu faktor yang sangat penting bahkan tidak dapat dilepaskan dari sebuah organisasi, baik institusi maupun perusahaan. SDM juga merupakan kunci yang menentukan perkembangan perusahaan. Pada hakikatnya, SDM berupa manusia yang dipekerjakan di sebuah organisasi sebagai penggerak, pemikir dan perencana untuk mencapai tujuan organisasi itu[1].

Perusahaan yang memiliki karyawan berkualitas dan berinovasi dapat dengan mudah mengelola perusahaannya sehingga tujuan yang telah ditetapkan dapat tercapai. Untuk melakukan penilaian kinerja karyawan untuk promosi jabatan sangat sulit dilakukan karena keterbatasan pemantauan karyawan dengan atasannya. Karena itulah dalam perusahaan ini dibutuhkan sistem yang dapat membantu dalam penilaian kinerja karyawan untuk promosi jabatan untuk proses seleksi dilakukan secara obyektif dan efisien[2].

Kinerja adalah keluaran (*result*) yang merupakan real value bagi perusahaan, misalnya berupa jumlah penjualan, jumlah produksi, tingkat kualitas, efisiensi biaya, profit, dan sebagainya. Untuk menghasilkan keluaran (*result*) tersebut dibutuhkan kompetensi tertentu, seperti kemampuan mengoperasikan mesin, kemampuan mengatur sumber daya, dan lain-lain[3].

Proses penilaian kinerja karyawan dalam prediksi promosi jabatan melalui tahap seleksi penilaian kriteria-kriteria seperti pengetahuan pekerjaan, kompetensi, tanggung jawab, kerajinan dan disiplin. Pada proses penilaiannya dilakukan secara manual oleh manager berdasarkan kriteria yang ada. Kriteria yang sudah ditentukan menjadi tolak ukur dalam menentukan karyawan yang akan dipromosikan dan tidak dipromosikan. Permasalahan dalam penelitian ini adalah adanya penilaian kualitas kinerja untuk promosi jabatan karyawan pada PT. JNE Cabang Pematangsiantar yaitu adanya posisi yang harus ditempati dan layak diisi, evaluasi kinerja setiap tahun dan untuk memperkuat kerja tim. Perusahaan melakukan evaluasi atau penilaian terhadap karyawan, merupakan bagian yang sangat penting karena dengan mengetahui kualitas kinerja karyawan yang baik maka akan diperoleh hasil kinerja yang baik. Target yang diinginkan perusahaan akan tercapai, kelangsungan perusahaan dan regenerasi dari karyawan dapat terlaksana dengan baik.

Pada suatu studi kasus di tahun 2019 yang menggunakan Algoritma C4.5 menunjukkan bahwa algoritma C4.5 memiliki kemampuan dalam pengambilan keputusan untuk menentukan prediksi promosi jabatan. Hasil perolehan perhitungan klasifikasi data dengan model decision tree menggunakan algoritma C4.5 yang dievaluasi dengan confusion matrix menghasilkan angka akurasi $78\% \pm 14,00\%$, *precision* $83,17\% \pm 14,67$ dan *recall* $89,17\% \pm 17,50\%$ serta dengan angka curva ROC 0.867. Dengan demikian hasil penelitian ini dapat memberikan rekomendasi kepada Apartemen Senayan Jakarta dalam melakukan kegiatan promosi jabatan sesuai dengan kriteria yang ada[2].

Dalam penulisan lain di tahun yang sama, berdasarkan hasil pengujian yang telah dilakukan oleh penulis menggunakan metode algoritma C4.5 terhadap data yang dihipung dari BAZNAS Kabupaten Agam dalam mengklasifikasikan jenis zakat yang akan disalurkan kepada penerima zakat di ruang lingkup BAZNAS Kabupaten Agam. Algoritma C4.5 berhasil menghasilkan pohon keputusan yang dapat dijadikan sebagai acuan bagi BAZNAS Kabupaten Agam dalam menentukan jenis zakat yang akan disalurkan kepada 8 asnaf. Dengan memanfaatkan rule yang dihasilkan oleh algoritma C4.5, penyaluran zakat akan menjadi tepat sasaran dan memakan waktu lebih sedikit untuk merealisasikannya[4].

Tulisan di tahun sebelumnya juga mengungkapkan bahwa Algoritma C4.5 dapat digunakan sebagai metode klasifikasi dalam memprediksi penjualan motor oleh PT. Capella Dinamik Nusantara Cabang Muka Kuning dengan memperhatikan nilai *gain* (penguatan) tertinggi dari empat atribut seperti Semester, Warna, Tahun Produksi, Segmen dan Harga. Dari pohon keputusan yang telah terbentuk, variabel tertinggi dalam memprediksi penjualan adalah variabel penjualan via, artinya pendistribusian motor dari berbagai tempat sangat mempengaruhi terhadap penjualan motor. Harga dalam range Sangat Terjangkau Sekali (STS) yaitu harga dibawah Rp 17.699.000,00 masih mendominasi motor matic. Artinya, matic masih menguasai pasar penjualan[5].

Penulisan terkait tentang Algoritma C4.5 pada studi memprediksi ketersediaan uang pada mesin ATM di tahun 2021 menjelaskan bahwa pengujian terhadap algoritma C4.5 menggunakan metode validasi *K-Fold Cross validation* dengan nilai *fold*=10 dapat diketahui tingkat akurasi sebesar 85%, nilai *Precision* sebesar 80% dan nilai *Recall* sebesar 66,67%. Sedangkan nilai AUC (*Area Under Curve*) adalah 0.833, hal ini menunjukkan bahwa jika nilai AUC mendekati nilai 1 maka tingkat akurasi semakin bagus[6].

Pada tahun yang sama, yang mengambil judul tentang diagnosa klasifikasi gangguan kepribadian didapatkan hasil bahwa algoritma C4.5 memiliki nilai akurasi sebesar 57.50%, mean *precision* sebesar 57.50%, *classification error* sebesar 0%, dan *recall* sebesar 57.50% untuk data training sedangkan data testing memiliki tingkat akurasi sebesar 72.67 %, *precision* sebesar 72.67%, dan *recall* sebesar 100%. Dan adapun hasil prediksinya despresi mayor adalah 73% masyarakat mengalami GAD 25% masyarakat dan masyarakat yang mengalami gangguan konversi 2%, hal ini berdasarkan nilai *confidance*. Nilai klasifikasi akurasi yang didapat pada penerapan algoritma C4.5 menghasilkan nilai lebih dari 60% yaitu akurasi data 72.67%[7].

Sesuai dengan penelitian terkait sebelumnya, maka penulisan ini akan menggunakan Data Mining dengan algoritma pemecahan masalahnya adalah algoritma C4.5. Dengan harapan prediksi tersebut dapat membantu perusahaan dalam mengambil keputusan mengenai karyawan yang akan di promosikan jabatan. Hasil dari penelitian ini sekiranya dapat membantu mengoptimalkan SDM sehingga perusahaan dapat memberikan pelayanan yang diharapkan.

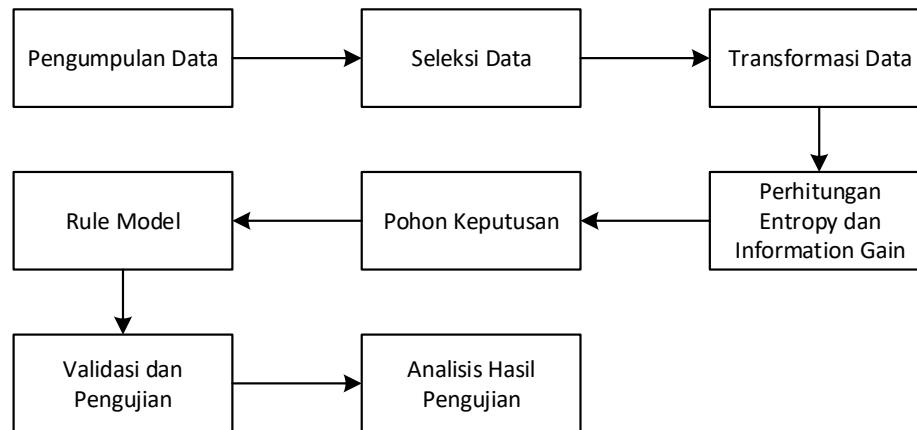
2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Berikut adalah tahapan penerapan metode dalam penelitian serta pengujian metode dalam mendapatkan hasil penelitian sesuai dengan harapan dan gambaran penelitian[8].

- Pengumpulan Data : Mengumpulkan data-data yang akan digunakan dalam proses algoritma klasifikasi C4.5.
- Seleksi Data : Memilih data yang akan digunakan dalam proses algoritma klasifikasi C4.5. Tujuan dari seleksi data adalah menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (*discovery*) akan dilakukan.
- Transformasi Data : Proses mentransformasi atau mengubah data ke dalam bentuk yang sesuai, agar dapat di proses dengan perhitungan algoritma C4.5.

- d. Perhitungan *Entropy* dan *Information Gain* Perhitungan semua atribut/variabel, *entropy* menggunakan rumus (1) dan *information gain* menggunakan rumus (2) untuk mengetahui *information gain* tertinggi yang akan dijadikan simpul akar pada pembuatan pohon keputusan.
- e. Pohon Keputusan (*Decision Tree*) : Hasil dari proses perhitungan *entropy* dan *information gain*, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan.
- f. Aturan-aturan/ *Rule Model* : Uraian penjelasan yang merepresentasikan sebuah pohon keputusan.
- g. Validasi dan Pengujian : Pengujian yang dilakukan untuk mengetahui semua fungsi bekerja dengan baik atau tidak. Validasi dilakukan dengan metode *Cross Validation*. *Cross validation* adalah salah satu metode dalam statistika untuk mengevaluasi dan membandingkan algoritma, dengan metode ini maka tingkat akurasi sebuah algoritma dapat ditampilkan dalam melakukan proses klasifikasi terutama kasus promosi jabatan[9]. Pada metode *cross validation* terdapat 2 segmen yaitu:
 1. Data latih, berfungsi membentuk model regresi;
 2. Data uji, berfungsi memvalidasi terhadap model dari data latih.
 Validasi dan pengujian dilakukan untuk mengetahui tingkat akurasi, presisi, dan *recall* dari hasil prediksi klasifikasi. Akurasi adalah persentase dari catatan yang diklasifikasikan dengan benar dalam pengujian dataset. Presisi adalah persentase data yang diklasifikasikan sebagai model baik yang sebenarnya juga baik. *Recall* adalah pengukuran tingkat pengenalan positif sebenarnya.
- h. Analisis Hasil : Dilakukan untuk memastikan bahwa hasil pengujian benar-benar sesuai dengan pembahasan. Analisa dilakukan dengan melakukan perhitungan kembali hasil validasi dan pengujian(akurasi, presisi, dan *recall*) secara manual.



Gambar 1. Tahapan Penelitian

2.2 Decision Tree

Decision Tree adalah sebuah struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok kelas tertentu. Level node teratas dari sebuah *Decision Tree* adalah node akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya *Decision Tree* melakukan strategi pencarian secara *top-down* untuk solusinya. Pada proses mengklasifikasi data yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari node akar (*root*) sampai node akhir (daun) dan kemudian akan diprediksi kelas yang dimiliki oleh suatu data baru tertentu[10].

Decision tree atau pohon keputusan adalah metode atau algoritma klasifikasi data mining dengan membentuk pola pohon keputusan yang digunakan untuk mendapatkan jawaban dari masalah yang dimasukkan. Dengan pohon keputusan, dapat dengan mudah mengidentifikasi hubungan antar faktor-faktor yang mempengaruhi masalah dan mencari solusi yang baik dengan memperhitungkan faktor-faktor tersebut. Sebuah pohon keputusan yang dibangun akan dapat lebih mudah dimengerti, fleksibel, menarik karena dapat divisualisasikan dalam bentuk gambar[11].

2.3 Data Mining

Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, *statistic*, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar. Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai database besar.

Berdasarkan pengertian data mining yang telah dijelaskan di atas, maka data mining merupakan pengetahuan yang tersembunyi di dalam database yang di proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan dari database tersebut[12]. Beberapa teknik dan sifat data mining adalah sebagai berikut[4]:

- a. Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dengan mudah dapat dimengerti.
- b. Prediksi merupakan sesuatu yang menghasilkan nilai yang ada di masa yang akan datang.
- c. Estimasi memiliki kemiripan dengan prediksi, yang membedakannya terletak pada target estimasi yang lebih kearah numerik dari pada kearah kategori.
- d. Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas.
- e. Pengklusteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan dan tidak memiliki kemiripan dengan kelompok yang lain.
- f. Asosiasi memiliki peran dalam menemukan atribut yang muncul dalam satu waktu. Asosiasi bertugas untuk mengungkap aturan untuk mengukur hubungan antar dua atau lebih atribut.

2.4 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipaham dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari record pada kategori tertentu[13].

Cara algoritma C4.5 untuk membangun pohon keputusan yaitu[13]:

- a. Pilih atribut yang akan digunakan sebagai akar
- b. Buatlah sebuah cabang untuk setiap nilai
- c. Bagilah kasus dalam sebuah cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Algoritma C4.5 merupakan algoritma yang digunakan untuk membangun sebuah pohon keputusan (decision tree) dari data. Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang juga merupakan algoritma untuk membangun sebuah pohon keputusan. Algoritma C4.5 secara rekursif mengunjungi tiap simpul keputusan, memilih percabangan optimal, sampai tidak ada cabang lagi yang mungkin dihasilkan. Adapun tahapan-tahapan untuk melakukan perhitungan algoritma C4.5 diantaranya, menyiapkan data training, menentukan akar dari pohon dengan menghitung *entropy*, kemudian menghitung nilai *gain*, setelah itu menentukan tupel yang ingin dipartisi[7].

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

Keterangan :

- S = Himpunan Kasus
- A = Fitur
- n = Jumlah Partisi S
- pi = Proporsi dari Si terhadap S

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \tag{2}$$

Keterangan :

- S = Himpunan Kasus
- A = Atribut
- n = Jumlah Partisi Atribut A
- | Si | = Jumlah Kasus pada partisi ke-i
- | S | = Jumlah Kasus dalam S

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data

Data yang digunakan adalah contoh data karyawan yang diambil dari JNE cabang Pematangsiantar.

Tabel 1. Data Karyawan

No	Karyawan	Pengetahuan Pekerjaan	Kompetensi	Kerja Tim	Komunikasi	Kualitas Pekerjaan	Promosi Jabatan
1	K1	Baik	Baik	Baik	Baik	Sangat Baik	Dipromosikan

No	Karyawan	Pengetahuan Pekerjaan	Kompetensi	Kerja Tim	Komunikasi	Kualitas Pekerjaan	Promosi Jabatan
2	K2	Baik	Baik	Kurang Baik	Baik	Baik	Dipromosikan
3	K3	Sangat Baik	Baik	Baik	Kurang Baik	Baik	Dipromosikan
4	K4	Baik	Tidak Baik	Baik	Tidak Baik	Tidak Baik	Tidak Dipromosikan
5	K5	Baik	Baik	Sangat Baik	Kurang Baik	Baik	Dipromosikan
6	K6	Baik	Baik	Baik	Baik	Sangat Baik	Dipromosikan
7	K7	Sangat Baik	Baik	Kurang Baik	Baik	Baik	Dipromosikan
8	K8	Baik	Sangat Baik	Baik	Baik	Kurang Baik	Dipromosikan
9	K9	Tidak Baik	Kurang Baik	Baik	Baik	Tidak Baik	Tidak Dipromosikan
10	K10	Baik	Baik	Baik	Baik	Kurang Baik	Dipromosikan
11	K11	Baik	Baik	Kurang Baik	Kurang Baik	Baik	Dipromosikan
12	K12	Tidak Baik	Baik	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Dipromosikan
13	K13	Baik	Baik	Baik	Kurang Baik	Baik	Dipromosikan
14	K14	Sangat Baik	Baik	Baik	Baik	Baik	Dipromosikan
15	K15	Baik	Baik	Baik	Baik	Sangat Baik	Dipromosikan
16	K16	Sangat Baik	Baik	Kurang Baik	Baik	Sangat Baik	Dipromosikan
17	K17	Kurang Baik	Tidak Baik	Baik	Tidak Baik	Baik	Tidak Dipromosikan
18	K18	Baik	Baik	Kurang Baik	Baik	Sangat Baik	Dipromosikan
19	K19	Kurang Baik	Baik	Sangat Baik	Baik	Baik	Dipromosikan
20	K20	Sangat Baik	Baik	Sangat Baik	Baik	Baik	Dipromosikan
21	K21	Sangat Baik	Baik	Sangat Baik	Sangat Baik	Sangat Baik	Dipromosikan
22	K22	Kurang Baik	Tidak Baik	Baik	Tidak Baik	Tidak Baik	Tidak Dipromosikan
23	K23	Baik	Kurang Baik	Sangat Baik	Baik	Baik	Dipromosikan
24	K24	Kurang Baik	Kurang Baik	Tidak Baik	Baik	Kurang Baik	Tidak Dipromosikan
25	K25	Sangat Baik	Baik	Kurang Baik	Kurang Baik	Baik	Tidak Dipromosikan

Data karyawan tersebut lalu diklasifikasikan menjadi 5 bagian yaitu, Pengetahuan Pekerjaan, Kompetensi, Kerja Tim, Komunikasi dan Kualitas Pekerjaan.

3.2 Pengolahan Data

Berdasarkan Tabel 1, maka selanjutnya dilakukan perhitungan *entropy* menggunakan rumus (1) dan *gain* dengan rumus (2). Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel 2 berikut.

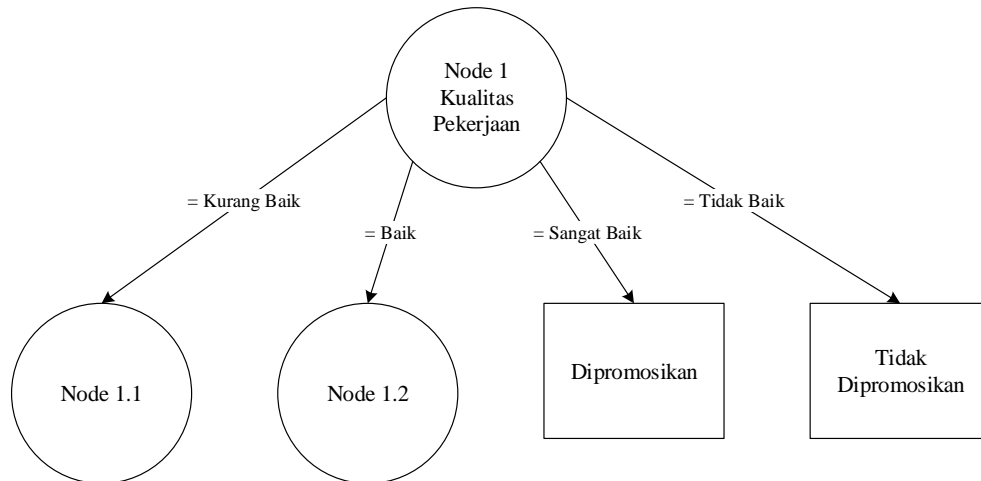
Tabel 2. Hasil Perhitungan Node 1

Node 1	Jumlah Kasus	Dipromosikan	Tidak Dipromosikan	Entropy	Information Gain
Total	25	18	7	0,855450811	
Pengetahuan Pekerjaan					0,361345844
Sangat Baik	7	6	1	0,591672779	
Baik	12	11	1	0,41381685	
Kurang Baik	4	1	3	0,811278124	
Tidak Baik	2	0	2	0	
Kompetensi					0,382909309
Sangat Baik	1	1	0	0	
Baik	18	16	2	0,503258335	
Kurang Baik	3	1	2	0,918295834	
Tidak Baik	3	0	3	0	
Kerja Tim					0,258663429
Sangat Baik	5	5	0	0	
Baik	12	8	4	0,918295834	
Kurang Baik	6	5	1	0,650022422	
Tidak Baik	2	0	2	0	
Komunikasi					0,371159488
Sangat Baik	1	1	0	0	
Baik	15	13	2	0,566509507	
Kurang Baik	5	4	1	0,721928095	
Tidak Baik	4	0	4	0	
Kualitas Pekerjaan					0,433244548
Sangat Baik	6	6	0	0	
Baik	12	10	2	0,650022422	

Node 1	Jumlah Kasus	Dipromosikan	Tidak Dipromosikan	Entropy	Information Gain
Kurang Baik	3	2	1	0,918295834	0
Tidak Baik	4	0	4		

Berdasarkan perhitungan *entropy* dan *gain* di Tabel 2, maka diperoleh nilai atribut tertinggi adalah Kualitas Pekerjaan dengan nilai *gain* sebesar 0,433244548. Maka atribut Kualitas Pekerjaan dipilih sebagai node akar. Nilai kelas atribut Sangat Baik mengklasifikasikan kasus menjadi keputusan yaitu Dipromosikan. Nilai kelas atribut Tidak Baik mengklasifikasikan kasus menjadi keputusan Tidak Dipromosikan. Sedangkan nilai kelas atribut Baik dan Kurang Baik belum diperoleh hasil antara Dipromosikan atau Tidak Dipromosikan, sehingga perlu dilakukan perhitungan lebih lanjut.

Berikut ini adalah proses pemodelan pohon keputusan berdasarkan hasil perhitungan algoritma C4.5 dari perhitungan node 1.

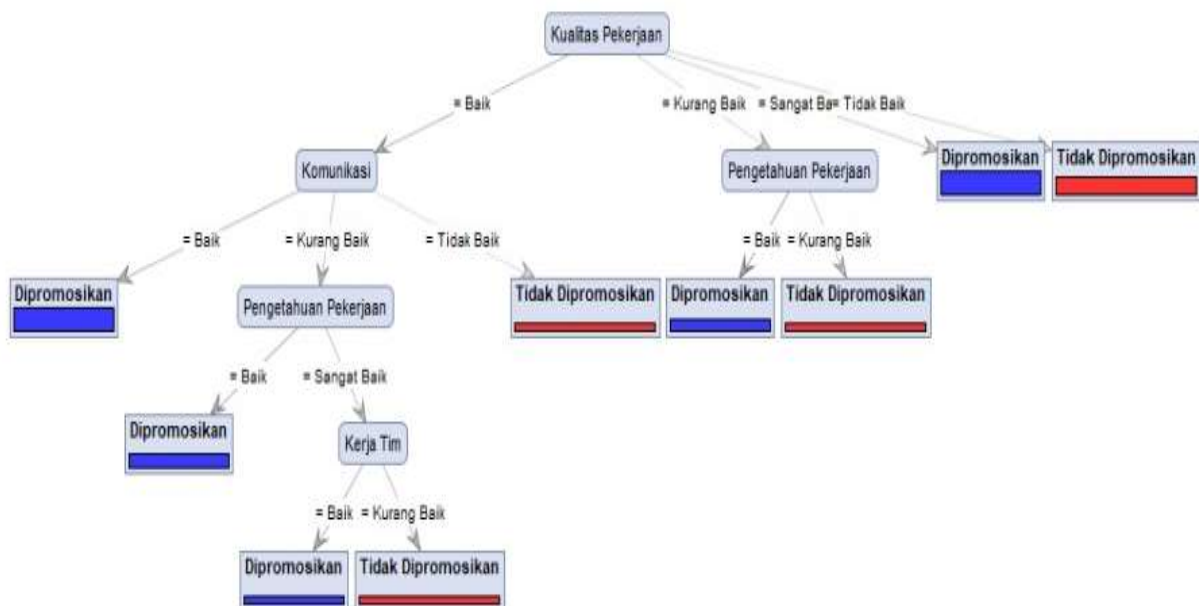


Gambar 2. Pohon Keputusan Hasil Perhitungan *Entropy* dan *Gain* untuk Mencari Akar

Proses perhitungan *entropy* dan *gain* akan berulang sampai semua cabang memiliki sebuah keputusan.

3.2 Implementasi

Hasil pengolahan Data Promosi Jabatan selanjutnya diimplementasikan ke dalam aplikasi *RapidMiner 7.3*. Sehingga pohon keputusan akhir dapat dilihat pada gambar 3 berikut.



Gambar 3. Implementasi data dengan Aplikasi *RapidMiner 7.3*

3.2.1 Rule dari Pohon Keputusan

Pada gambar 3 di atas merupakan pohon keputusan yang dihasilkan pada *RapidMiner* dengan aturan atau *rule* yang dapat dilihat pada *text view* pada gambar 4 berikut.

Tree

```

Kualitas Pekerjaan = Baik
| Komunikasi = Baik: Dipromosikan {Dipromosikan=6, Tidak Dipromosikan=0}
| Komunikasi = Kurang Baik
| | Pengetahuan Pekerjaan = Baik: Dipromosikan {Dipromosikan=3, Tidak Dipromosikan=0}
| | Pengetahuan Pekerjaan = Sangat Baik
| | | Kerja Tim = Baik: Dipromosikan {Dipromosikan=1, Tidak Dipromosikan=0}
| | | Kerja Tim = Kurang Baik: Tidak Dipromosikan {Dipromosikan=0, Tidak Dipromosikan=1}
| Komunikasi = Tidak Baik: Tidak Dipromosikan {Dipromosikan=0, Tidak Dipromosikan=1}
Kualitas Pekerjaan = Kurang Baik
| Pengetahuan Pekerjaan = Baik: Dipromosikan {Dipromosikan=2, Tidak Dipromosikan=0}
| Pengetahuan Pekerjaan = Kurang Baik: Tidak Dipromosikan {Dipromosikan=0, Tidak Dipromosikan=1}
Kualitas Pekerjaan = Sangat Baik: Dipromosikan {Dipromosikan=6, Tidak Dipromosikan=0}
Kualitas Pekerjaan = Tidak Baik: Tidak Dipromosikan {Dipromosikan=0, Tidak Dipromosikan=4}
    
```

Gambar 4. Rule Decision Tree pada *RapidMiner*

3.2.2 Validasi Decision Tree

Hasil penerapan Algoritma C4.5 menggunakan aplikasi *RapidMiner* dengan operator *Cross Validation* diperoleh nilai akurasi yaitu sebesar 71,43% dan *classification error* yaitu sebesar 28,57%. Hasil akurasi tersebut diperoleh dengan pengaturan pada *Cross Validation* dengan nilai *number of validation* = 10 dan *sampling type* = *shuffled sampling*. Berikut ini adalah hasil akurasi yang diperoleh.

Table View Plot View

accuracy: 71.43%

	true Dipromosikan	true Tidak Dipromosikan	class precision
pred. Dipromosikan	4	2	66.67%
pred. Tidak Dipromosikan	0	1	100.00%
class recall	100.00%	33.33%	

Gambar 5. Nilai Akurasi Algoritma C4.5

PerformanceVector

```

PerformanceVector:
accuracy: 71.43%
ConfusionMatrix:
True:  Dipromosikan  Tidak Dipromosikan
Dipromosikan:  4      2
Tidak Dipromosikan:  0      1
classification_error: 28.57%
ConfusionMatrix:
True:  Dipromosikan  Tidak Dipromosikan
Dipromosikan:  4      2
Tidak Dipromosikan:  0      1
    
```

Gambar 6. Performance Vector Algoritma C4.5

3.3 Pembahasan

Hasil yang dilakukan oleh peneliti dalam perhitungan Algoritma C4.5 diperoleh 9 model aturan atau rule prediksi promosi jabatan di PT. JNE cabang Pematangsiantar. Implementasi dalam bentuk pohon keputusan dalam dilihat pada gambar 3.

Berdasarkan pengolahan data menggunakan aplikasi *RapidMiner* di dapat nilai akurasi sebesar 71,43%, artinya rule yang dihasilkan tingkat kebenarannya mendekati 100%. Dimana hasil dari *class precision* label Dipromosikan sebesar 66,67% dan prediksi label Tidak Dipromosikan sebesar 100%.

Parameter yang digunakan pada decision tree telah dilakukan penyesuaian terhadap kriteria *decision tree* yang digunakan pada *RapidMiner* yaitu *information gain*,

Maksimal depth = 20
 Confidence = 0,25
 Minimal gain = 0,1
 Minimal leaf size = 1,

Minimal size for split = 2

Number or pruning alt = 3

Model aturan dalam bentuk pohon keputusan yang dihasilkan dari *RapidMiner* dapat dilihat pada gambar

4. KESIMPULAN

Berdasarkan pembahasan sebelumnya, maka dapat disimpulkan penerapan klasifikasi data mining dengan Algoritma C4.5 dapat diterapkan dalam memprediksi promosi jabatan terhadap karyawan di PT. JNE cabang Pematangsiantar. Data Karyawan yang diolah dalam penelitian ini berjumlah 25 orang karyawan. Dengan hasil yang diperoleh bahwa atribut kualitas pekerjaan adalah atribut yang paling berpengaruh. Atribut yang digunakan dalam penelitian ini ada 5, diantaranya Pengetahuan Pekerjaan, Kompetensi, Kerja Tim, Komunikasi dan Kualitas Pekerjaan. Dan menghasilkan 9 model aturan keputusan yaitu 5 keputusan Dipromosikan dan 4 keputusan Tidak Dipromosikan. Hasil pengujian Algoritma C4.5 dengan menggunakan aplikasi *RapidMiner 7.3* memperoleh tingkat akurasi sebesar 71,43%. Dengan prediksi Dipromosikan untuk true Dipromosikan mencapai 4 Karyawan dan untuk true Tidak Dipromosikan sebanyak 0 Karyawan. Sedangkan prediksi Tidak Dipromosikan untuk true Dipromosikan sebanyak 0 Karyawan dan untuk true Tidak Dipromosikan sebanyak 1 Karyawan. Untuk class recall true Dipromosikan mencapai 100% dan class recall true Tidak Dipromosikan mencapai 33,33%. Dan untuk *class precision* untuk prediksi Dipromosikan mencapai 66,67% dan *class precision* prediksi Tidak Dipromosikan mencapai 100%. Untuk lebih jelasnya dapat dilihat pada gambar 5.

REFERENCES

- [1] E. Susan, “Manajemen Sumber Daya Manusia,” *J. Manaj. Pendidik. Islam*, vol. 9, no. 2, pp. 952–962, 2019.
- [2] Sunarti, “Prediksi Promosi Jabatan Karyawan Dengan Algoritma C4.5 (Studi Kasus: Apartemen Senayan Jakarta),” *Techno.Com*, vol. 18, no. 4, pp. 288–298, 2019, doi: 10.33633/tc.v18i4.2471.
- [3] M. S. Ma’arif and L. Kartika, *Kinerja*. Bogor: PT Penerbit IPB Press, 2021.
- [4] R. S. Asa, S. Defit, and J. Na’am, “Identifikasi Penyaluran Zakat Menggunakan Algoritma C4.5 (Studi Kasus di BAZNAS Kabupaten Agam),” *J. Sains dan Inform.*, vol. 5, no. 1, p. 50, 2019, doi: 10.22216/jsi.v5i1.4048.
- [5] N. Azwanti, “Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning,” *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 13, no. 1, p. 33, 2018, doi: 10.30872/jim.v13i1.629.
- [6] F. Syahputra, H. Hartono, and R. Rosnelly, “Penerapan Algoritma C4.5 Dalam Memprediksi Ketersediaan Uang Pada Mesin ATM,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 556, 2021, doi: 10.30865/mib.v5i2.2933.
- [7] S. Febriani and H. Sulistiani, “Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 4, pp. 89–95, 2021.
- [8] E. P. Cynthia and E. Ismanto, “Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan,” *J. Ris. Sist. Inf. Dan Tek. Inform.*, vol. (3) Juli, no. July, pp. 1–13, 2018, [Online]. Available: <http://tunasbangsa.ac.id/ejurnal/index.php/jurasik/article/download/60/pdf>.
- [9] C. Anam and H. B. Santoso, “Perbandingan Kinerja Algoritma C4. 5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa,” *Ejournal.Upm.Ac.Id*, vol. 8, no. 1, pp. 2088–4591, 2018, [Online]. Available: <https://ejournal.upm.ac.id/index.php/energy/article/view/111>.
- [10] W. Julianto, R. Yunitarini, and M. K. Sophan, “Algoritma C4.5 Untuk Penilaian Kinerja Karyawan,” *Scan*, vol. Vo. IX, no. No. 2, pp. 33–39, 2014.
- [11] A. Rufiyanto, M. Rochcham, and A. Rohman, *Isi Decision Tree*. Yogyakarta: Deepublish Publisher, 2020.
- [12] D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [13] J. Eska, “Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5,” *JURTEKSI (Jurnal Teknol. dan Sist. Informatika)*, vol. 2, pp. 9–13, 2018, doi: 10.31227/osf.io/x6svc.