

3

Analisis Algoritma C4.5 Untuk Pengangkatan Karyawan Tetap Studi Kasus PT Citra Abadi Sejati Bogor

Oleh : Fitria Rachmawati

Abstrak

PT. Citra Abadi Sejati, merupakan perusahaan garment produk tekstil yang berlokasi di daerah Bogor, Jawa Barat dan merupakan anak perusahaan Texmaco Group. Peranan sumber daya manusia dalam perusahaan sangatlah penting karena sebagai penggerak utama seluruh kegiatan atau aktivitas perusahaan dalam mencapai tujuannya, baik untuk memperoleh keuntungan maupun untuk mempertahankan kelangsungan hidup perusahaan. Kualitas Sumber Daya Manusia (SDM) merupakan salah satu faktor yang untuk meningkatkan produktivitas kinerja suatu organisasi atau instansi. Salah satu upaya untuk mengoptimalkan produksi yang dihasilkan oleh PT. Citra Abadi Sejati perlu adanya pengangkatan karyawan kontrak dengan cara seleksi. Proses seleksi pengangkatan karyawan masih belum dilakukan secara profesional, tetapi dilakukan dengan cara-cara penyusunan, pertemanan, atau hubungan kerja. Menilai kinerja kerja berarti membandingkan kinerja actual bawahan dengan standar yang ditetapkan oleh perusahaan. Secara umum Algoritma C4.5 untuk membangun pohon keputusan, dan untuk analisa data dengan menggunakan Confusion Matrix. Tujuan penelitian ini adalah untuk mendapat karyawan seperti yang diharapkan, perusahaan sehingga kualitas dan kuantitas produksi semakin meningkat dengan baik. Harapan dari hasil seleksi dengan menggunakan metode Algoritma C4.5 adalah mendapatkan karyawan yang sesuai dengan kebutuhan dan standar nilai kinerja selaras dengan posisi.

Kata Kunci : *Pengangkatan Karyawan, Algoritma C4.5, Penilaian kinerja*

Pendahuluan

Latar Belakang

Perkembangan dunia usaha saat ini semakin bertambah pesat, sehingga perusahaan di dalam mengelola usaha diharapkan mampu menggunakan sumber daya manusia dengan baik dan benar. Sumber daya manusia merupakan salah satu faktor kunci dalam reformasi ekonomi, yakni bagaimana menciptakan sumber daya manusia yang berkualitas dan memiliki keterampilan serta berdaya saing tinggi dalam persaingan global yang selama ini kita abaikan. Karyawan merupakan salah satu sumber daya yang digunakan sebagai alat penggerak dalam memajukan perusahaan.

Peranan sumber daya manusia dalam perusahaan sangatlah penting karena sebagai penggerak utama seluruh kegiatan atau aktivitas perusahaan dalam mencapai tujuannya, baik untuk memperoleh keuntungan maupun untuk mempertahankan kelangsungan hidup perusahaan. Berhasil tidaknya suatu perusahaan dalam mempertahankan eksistensi perusahaan dimulai dari manusia itu sendiri dalam mempertahankan perusahaan dalam meningkatkan efektivitas dan efisiensi secara maksimal.

Kualitas Sumber Daya Manusia (SDM) merupakan salah satu faktor yang untuk meningkatkan produktivitas kinerja suatu organisasi atau instansi. Dalam bidang personalia permasalahan yang mungkin timbul adalah dari seleksi pengangkatan karyawan yang semula kontrak menjadi tetap. Kebutuhan karyawan baru didalam perusahaan tidak bisa dipastikan, walaupun sudah direncanakan dengan sebaik-baiknya. Hal ini dikarenakan berbagai faktor yang timbul baik dalam perusahaan (faktor intern) maupun faktor yang berasal dari luar perusahaan (faktor ekstern). Untuk itu sebelum melakukan seleksi pemilihan karyawan

perusahaan menentukan kebutuhan karyawan yang dibutuhkan baik kualitas maupun kuantitas (jumlah).

Pelaksanaan seleksi merupakan salah satu bagian penting dalam aktivitas atau kegiatan Manajemen Sumber Daya Manusia. Terutama untuk pengadaan tenaga kerja, karena dengan adanya seleksi maka akan menghasilkan karyawan-karyawan yang sesuai dengan kebutuhan perusahaan baik dan segi jumlah maupun mutu, yang akan menjamin kelancaran tugas-tugas dan jalannya aktivitas. Guna mendapat karyawan seperti yang diharapkan, perusahaan harus melakukan penyeleksian karyawan dengan selektif agar dapat meningkatkan produktivitas kinerja kerja. Seleksi adalah kegiatan suatu perusahaan untuk dapat memilih karyawan yang paling tepat dalam jumlah yang cepat pula dari calon-calon yang dapat ditariknya.

Setiap perusahaan juga harus dapat melakukan seleksi secara efektif dan efisien, dengan demikian metode seleksi yang dilaksanakan tersebut harus dapat memilih atau menetapkan karyawan yang paling tepat. Pengambilan keputusan dari suatu masalah, baik itu masalah yang sederhana maupun yang kompleks, diperlukan informasi-informasi yang menyeluruh dan akurat, sehingga dengan kemampuan analisa yang tajam, diharapkan dapat melahirkan keputusan-keputusan yang sesuai dengan permasalahan yang dihadapi. Proses seleksi pengangkatan karyawan masih belum dilakukan secara professional, tetapi dilakukan dengan cara-cara penyuaipan, pertemanan, atau hubungan kerja. Hal ini terjadi karena tidak ada metode standar yang sistematis untuk menilai kelayakan kinerja kerja karyawan. Menilai kinerja kerja berarti membandingkan kinerja actual bawahan dengan standar yang ditetapkan oleh perusahaan. Jika dikerjakan dengan benar, hal ini akan memberikan manfaat yang penting bagi karyawan, atasan serta departemen SDM dan perusahaan. Atasan /supervisor atau manajer menilai kinerja karyawan untuk mengetahui tindakan apa yang sudah dilakukan selanjutnya. Umpan balik spesifik dari atasan akan memudahkan karyawan untuk membuat perencanaan-perencanaan kerja serta keputusan-keputusan yang lebih efektif untuk kemajuan perusahaan.

Biasanya permasalahan yang seperti ini dapat menggunakan metode *algoritma C4.5* untuk cara menyelesaikannya. *Algoritma C4.5* merupakan metode perhitungan algoritma yang digunakan untuk membentuk pohon keputusan. Metode ini dapat membantu memecahkan suatu permasalahan dalam kasus pengangkatan karyawan berdasarkan penilaian kinerja. Penelitian dengan judul "Pengangkatan Karyawan berdasarkan penilaian kinerja Pada PT. CITRA ABADI SEJATI" membahas tentang memprediksi dalam menyeleksi karyawan yang semula kontrak menjadi karyawan tetap dengan cara menilai kinerja kerja guna mendapatkan karyawan yang tepat sesuai dengan harapan untuk menunjang produktivitas kinerja kerja pada perusahaan tersebut.

Rumusan Masalah

Pada PT. Citra Abadi Sejati seorang karyawan setelah di putuskan untuk di terima bekerja maka karyawan tersebut harus menjalani sama kontrak kerja terlebih dahulu selama 6

bulan. Selama 6 bulan menjalani masa kontrak kerjanya, kemudian di perpanjang kontrak ke dua yaitu 2 tahun, seharusnya karyawan tersebut dilakukan penilaian untuk menentukan apakah karyawan tersebut layak atau tidak untuk di putuskan untuk menjadi karyawan tetap.

Namun pada penentuan pengangkatan karyawan masih belum dilakukan secara professional, tetapi masih dilakukan dengan cara-cara penyuaipan, pertemanan, atau hubungan kerja. Hal ini dikarenakan belum adanya kriteria yang tepat untuk menetapkan pengangkatan karyawan kontrak menjadi tetap. Maka dalam menentukan seleksi harus dilakukan dengan cara yang tepat sehingga mendapatkan karyawan yang berkualitas sesuai dengan kebutuhan produksi.

Berdasarkan dari uraian di atas, maka dapat diidentifikasi masalah sebagai berikut:

1. Belum dapat ditentukan kriteria-kriteria yang lebih tepat untuk mengakurasi penetapan dalam pengangkatan karyawan tetap.
2. Belum adanya aturan-aturan yang akurat untuk menentukan pengangkatan karyawan tetap.

Problem Statement

Belum akuratnya penetapan dalam pengangkatan karyawan kontrak menjadi karyawan tetap.

Research Question

Bagaimana penerapan *Algoritma C4.5* untuk akurasi penetapan dalam pengangkatan karyawan tetap dari karyawan kontrak pada PT. Citra Abadi Sejati ?

Maksud dan Tujuan Penelitian

Maksud penelitian ini adalah untuk menerapkan metode *Algoritma C4.5* untuk akurasi penetapan dalam pengangkatan karyawan tetap.

Sedangkan tujuan dari penelitian ini adalah :

1. Dapat ditentukan kriteria-kriteria yang lebih tepat untuk mengakurasi penetapan dalam pengangkatan karyawan tetap.
2. Mendapatkan aturan-aturan yang akurat untuk menentukan pengangkatan karyawan tetap

Landasan Teori

Data Mining

Data mining (*Ari Fadli, 2011*), adalah kegiatan mengekstrasi atau menambang pengetahuan dari data yang berukuran/berjumlah besar, informasi inilah yang nantinya sangat berguna untuk pengembangan.

Algoritma C4.5

Algoritma C4.5 (Kusrini and Taufiq, Emha Luthfi. 2009) merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Secara umum *Algoritma C4.5*.

1. Entropy

Entropy merupakan pengukuran ketidakpastian rata-rata kumpulan data ketika kita tidak tahu hasil dari sumber informasi yang kita tidak punya. Ini juga menunjukkan jumlah rata-rata informasi yang akan diterima dari hasil sumber informasi.

Untuk menentukan nilai *entropy* gunakan rumus

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

Ket :

S = himpunan kasus

N = jumlah partisi S

Pi = proporsi Si terhadap S

2. Gain Ratio

Pada konstruksi pohon C4.5, di setiap simpul pohon, atribut dengan nilai gain ratio tertinggi sebagai atribut split untuk simpul. Rumus dari gain adalah sebagai berikut

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

Ket :

S = himpunan kasus

A = fitur

N = jumlah partisi atribut A

|Si| = proporsi Si terhadap S

|S| = jumlah kasus dalam S

Rapid Miner

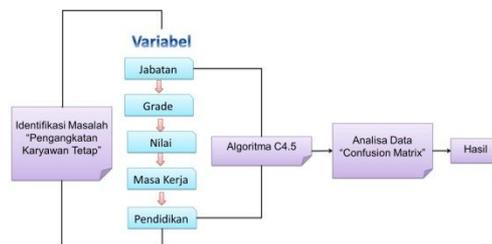
RapidMiner (Akhtar, Fareed dan Caroline Hahne. 2012), merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi.

Kerangka Pemikiran

Pada proses pengangkatan karyawan diperlukan cara yang baik untuk dapat dijadikan acuan dalam menentukan karyawan tetap dalam menilai kinerja kerja yang objektif seperti

menerapkan suatu metode atau upaya untuk memenuhi beberapa kinerja penilaian, sehingga umpan balik yang dihasilkan dapat digunakan sebagai dasar untuk pengambilan keputusan yang tepat berdasarkan kriteria yang ada di PT. Citra Abadi Sejati yang akan ditetapkan.

Metode yang digunakan untuk menganalisa dalam pengangkatan karyawan berdasarkan penilaian kinerja kerja adalah menggunakan metode *Algoritma C4.5*. Metode ini pertama dengan mendefinisikan permasalahan dalam perusahaan PT. Citra Abadi Sejati untuk di carikan solusi mengklasifikasikan class dan atribut yang telah di tentukan untuk mencari nilai gain yang tertinggi dari setiap kriteria/class tersebut untuk membuat pohon keputusan. Diharapkan dengan menggunakan metode perhitungan *Algoritma C4.5* ini dapat menentukan karyawan yang memiliki kinerja kerja yang baik dan mendapatkan prestasi pengangkatan karyawan tetap. Kerangka pemikiran dapat dilihat pada gambar 1.



Gambar 1. Kerangka Pemikiran

Desain Penelitian

Penelitian ini dilakukan pada PT. Citra Abadi Sejati untuk pengangkatan karyawan berdasarkan penilaian kinerja. Dimana penelitian ini menentukan karyawan yang tepat mendapatkan prestasi atau peningkatan grade dan merubah status karyawan tersebut yang semula karyawan kontrak menjadi karyawan tetap berdasarkan penilaian kinerja.

Adapun tahapan-tahapan yang dilakukan dalam proses pemecahan suatu masalah yang di mulai dari mengidentifikasi permasalahannya terlebih dahulu, selanjutnya menentukan metode yang akan digunakan untuk meneliti, pengumpulan data, menganalisis data dengan menggunakan metode yang telah dipilih, kemudian terakhir mendapatkan hasil akhir. Desain penelitian adalah sebagai berikut:

1. Identifikasi Masalah

Tahapan dimana melakukan identifikasi terhadap masalah yang ada, dari permasalahan yang telah ditentukan akan di carikan metode yang cocok untuk diterapkan, kemudian dilakukan penelitian atau uji coba untuk membuktikan metode tersebut

2. Pengumpulan Data

Pada tahap ini, ada beberapa tehnik yang digunakan pada pengumpulan data, salah satunya dengan dokumentasi. Guna memperoleh informasi data yang akurat dalam menunjang penelitian yang sedang dilakukan dengan tujuan mendapatkan hasil yang diharapkan.

3. Analisis Data

Data yang telah dikumpulkan akan dilakukan uji coba menggunakan teori yang telah ditentukan yaitu dengan metode *Algoritma C4.5*.

4. Hasil Akhir

Hasil dari proses analisis dengan menggunakan metode *Algoritma C4.5* adalah klasifikasi karyawan kontrak yang di angkat menjadi karyawan tetap

Hasil Penelitian

Proses pengklasifikasian Data Mining menggunakan Algoritma C4.5

Data dalam penelitian ini berdasarkan data karyawan pada bulan Oktober 2016 sebanyak 593 data pada bagian produksi. Contoh dari data karyawan dapat dilihat pada table 1

Tabel 1. Contoh Data Karyawan

| No | Nip | Grade | Nilai | Nama Pegawai | Jabatan | Bagian | Pendidikan | L/P | Tempat Lahir | Tgl Lahir | Tgl Masuk | Masa Kerja |
|----|--------|-------|--------|------------------|----------------------|-------------------|------------|-----|--------------|-----------|-----------|------------|
| 1 | 006376 | B | Rendah | Endang Saptaning | Operator | Sample | 3 | P | Wonorejo | 01-Apr-65 | 17-Oct-05 | 11 tahun |
| 2 | 006904 | B | Rendah | Siti Sopiiah | Operator | CAD | 3 | P | Bogor | 12-Jul-86 | 26-Sep-05 | 11 tahun |
| 3 | 000895 | B | Rendah | Sutarmi | Recorder | Marker | 3 | P | Klaten | 10-Jul-74 | 13-Oct-97 | 19 tahun |
| 4 | 002360 | B | Rendah | Nining S | Supervisor | Sewing Staff 3 | 1 | P | Bogor | 08-Mar-68 | 25-Mar-02 | 14 tahun |
| 5 | 001090 | B | Rendah | Evi Fristiwati | Q.C Audit | Quality Control 3 | 3 | P | Surabaya | 16-Nov-75 | 16-Mar-98 | 18 tahun |
| 6 | 000538 | B | Rendah | Luki Wahyudin | Assistant Supervisor | Cutting 2 | 3 | L | Bogor | 17-Feb-75 | 25-Sep-95 | 21 tahun |
| 7 | 001743 | B | Rendah | Yanah | Assistant Supervisor | Quality Control 2 | 3 | P | Bogor | 25-Jun-68 | 04-May-00 | 16 tahun |
| 8 | 001863 | B | Rendah | Nursetiasih | Staff | Industrial Eng. 2 | 3 | P | Bogor | 24-Feb-77 | 11-Oct-00 | 16 tahun |
| 9 | 001160 | B | Rendah | Zainul Hamidah | Assistant Supervisor | Cutting 2 | 3 | P | Sragen | 13-May-76 | 25-Apr-98 | 18 tahun |
| 10 | 001134 | B | Rendah | Hamsinar | Q.C Audit | Quality Control 3 | 3 | P | B.L.Lumpu | 11-Jan-76 | 20-Apr-98 | 18 tahun |

Dari data karyawan yang telah dikumpulkan, maka tahap berikutnya yaitu adalah dilakukan pengolahan data. Proses pengolahan data mining yang dilakukan pada penelitian ini terdiri dari beberapa tahap yaitu, tahap data selection dan data transformation.

1. Data selection

Data yang terkumpul pada table 1 terdapat beberapa atribut yang tidak terpakai dalam proses pengolahan data. Atribut yang tidak memiliki kontribusi terhadap tujuan penelitian, tidak digunakan dalam penelitian ini. Table 2 merupakan hasil dari proses data selection dapat dilihat pada table 2

Tabel 2 Data Karyawan setelah tahap *data selection*

| No | Grade | Nilai | Nama Pegawai | Jabatan | Pendidikan | Masa Kerja |
|----|-------|--------|------------------|----------------------|------------|------------|
| 1 | B | Rendah | Endang Saptaning | Operator | 3 | 11 tahun |
| 2 | B | Rendah | Siti Sopiiah | Operator | 3 | 11 tahun |
| 3 | B | Rendah | Sutarmi | Recorder | 3 | 19 tahun |
| 4 | B | Rendah | Nining S | Supervisor | 1 | 14 tahun |
| 5 | B | Rendah | Evi Fristiwati | Q.C Audit | 3 | 18 tahun |
| 6 | B | Rendah | Luki Wahyudin | Assistant Supervisor | 3 | 21 tahun |
| 7 | B | Rendah | Yanah | Assistant Supervisor | 3 | 16 tahun |
| 8 | B | Rendah | Nursetiasih | Staff | 3 | 16 tahun |
| 9 | B | Rendah | Zainul Hamidah | Assistant Supervisor | 3 | 18 tahun |
| 10 | B | Rendah | Hamsinar | Q.C Audit | 3 | 18 tahun |
| 11 | B | Rendah | Nanik R | Assistant Supervisor | 3 | 15 tahun |
| 12 | B | Rendah | Suprianto | Assistant Supervisor | 3 | 22 tahun |
| 13 | B | Rendah | Sartono | Assistant Supervisor | 3 | 17 tahun |
| 14 | B | Rendah | Suyadi | Supervisor | 3 | 11 tahun |
| 15 | B | Rendah | Suhendrik | Operator Jahit | 2 | 11 tahun |

2. Data Transformation

Tahap selanjutnya yaitu melakukan proses transformasi dengan cara mengubah pengetahuan dari data real yang terdapat pada table 2. transformasi data ini dilakukan pada atribut Lama Kerja atau masa kerja dan pada atribut Pendidikan.

Mengelompokan lama kerja berdasarkan pada data karyawan seperti terlihat pada table 3

Tabel 3 Klasifikasi Lama Kerja

| Lama Kerja (tahun) | Kriteria |
|--------------------|--------------|
| 3 Tahun | Masa Promosi |
| <3 Tahun | Baru |

Mengelompokan pendidikan berdasarkan pada data karyawan seperti terlihat pada table 4

Tabel 4.4 Klasifikasi Pendidikan

| Pendidikan | Kriteria |
|------------|----------|
| 1 | SD |
| 2 | SMP |
| 3 | SMA |

Hasil dari proses data transformation dapat dilihat pada table 5

Tabel 5 Data Karyawan hasil Tranformation

| No | Grade | Nilai | Nama Pegawai | Jabatan | Pendidikan | Masa Kerja |
|----|-------|--------|----------------|----------------|------------|--------------|
| 1 | B | Tinggi | Mulyaman Iskan | Operator | SMA | Masa Promosi |
| 2 | H | Rendah | Prisiyatun | Operator Jahit | SMP | Masa Promosi |
| 3 | H | Sedang | Elli Susanti | Operator Jahit | SMP | Masa Promosi |
| 4 | H | Tinggi | Erlinda | Operator | SMA | Masa Promosi |
| 5 | H | Rendah | Suryana | Operator Jahit | SMA | Masa Promosi |
| 6 | H | Rendah | Widha Hendraw | Operator Jahit | SMP | Masa Promosi |
| 7 | H | Rendah | Eka Jayanti | Operator Jahit | SMP | Masa Promosi |
| 8 | K | Rendah | Sally Siregar | Operator | SMA | Baru |
| 9 | K | Rendah | Anih Hanapiah | Helper | SMA | Baru |
| 10 | K | Rendah | Reni Suherni | Operator Jahit | SMA | Baru |
| 11 | K | Rendah | Enah | Operator Jahit | SMP | Baru |
| 12 | K | Rendah | Maryani | Operator Jahit | SD | Baru |
| 13 | K | Rendah | Nengsih | Operator Jahit | SD | Baru |
| 14 | K | Rendah | Nuraini | Operator Jahit | SMA | Baru |
| 15 | K | Rendah | Siti Soleha | Operator Jahit | SMP | Baru |
| 16 | K | Rendah | Erni Avrianti | Operator Jahit | SD | Baru |
| 17 | K | Rendah | Ria Ratnawati | Operator Jahit | SMA | Baru |
| 18 | B | Rendah | Tuti Setiawati | Operator Jahit | SMP | Baru |

Analisis Pengangkatan Karyawan Tetap Menggunakan Algoritma C4.5

Berikut ini data yang digunakan dalam penelitian dapat di lihat pada table 6

Tabel 6. Data Karyawan

| No | Grade | Nilai | Nama Pegawai | Jabatan | Masa Kerja | Hasil |
|----|-------|--------|-------------------|----------------|--------------|-----------|
| 1 | B | Tinggi | Mulyaman Iskandar | Operator | Masa Promosi | Di Terima |
| 2 | H | Rendah | Prisiyatun | Operator Jahit | Masa Promosi | Di Tolak |
| 3 | H | Sedang | Ellis Susanti | Operator Jahit | Masa Promosi | Di Terima |
| 4 | H | Tinggi | Erlinda | Operator | Masa Promosi | Di Terima |
| 5 | H | Rendah | Suryana | Operator Jahit | Masa Promosi | Di Tolak |
| 6 | H | Rendah | Widha Hendrawati | Operator Jahit | Masa Promosi | Di Tolak |
| 7 | H | Rendah | Eka Jayanti | Operator Jahit | Masa Promosi | Di Tolak |
| 8 | K | Rendah | Sally Siregar | Operator | Baru | Di Tolak |
| 9 | K | Rendah | Anih Hanapiah | Helper | Baru | Di Tolak |
| 10 | K | Rendah | Reni Suherni | Operator Jahit | Baru | Di Tolak |
| 11 | K | Rendah | Enah | Operator Jahit | Baru | Di Tolak |
| 12 | K | Rendah | Maryani | Operator Jahit | Baru | Di Tolak |
| 13 | K | Rendah | Nengsih | Operator Jahit | Baru | Di Tolak |
| 14 | K | Rendah | Nuraini | Operator Jahit | Baru | Di Tolak |
| 15 | K | Rendah | Siti Soleha | Operator Jahit | Baru | Di Tolak |
| 16 | K | Rendah | Erni Avrianti | Operator Jahit | Baru | Di Tolak |
| 17 | K | Rendah | Ria Ratnawati | Operator Jahit | Baru | Di Tolak |
| 18 | K | Rendah | Tuti Setiawati | Operator Jahit | Baru | Di Tolak |
| 19 | K | Rendah | Sutarsih | Operator Jahit | Baru | Di Tolak |
| 20 | K | Tinggi | Ismiyani | Operator | Baru | Di Terima |

Dari data yang tertera pada table 4.12 akan di buat pohon keputusan untuk menentukan karyawan Tetap dan Kontrak dengan melihat dari JABATAN, GRADE, MASA KERJA, PENILAIAN KINERJA, dan PENDIDIKAN.

Secara umum *Algoritma C4.5* untuk membangun pohon keputusan adalah sebagai berikut.

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap nilai.
- c. Bagi kasus dalam cabang.
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut atribut yang ada. Untuk menghitung gain digunakan rumus :

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

Keterangan :

S: himpunan Kasus

A: fitur

n : jumlah partisi S

P_i : proporsi dari S_i terhadap S

Entropy merupakan pengukuran ketidakpastian rata-rata kumpulan data ketika kita tidak tahu hasil dari sumber informasi yang kita tidak punya. Ini juga menunjukkan jumlah rata-rata informasi yang akan diterima dari hasil sumber informasi.

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

Keterangan :

S: himpunan kasus

A: atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke-i

$|S|$: jumlah kasus dalam S

Hasil Pengolahan Data Dengan Menggunakan Tools RapidMiner

Pengujian terhadap analisa, sangat penting dilakukan untuk menentukan dan memastikan apakah hasil analisa tersebut telah sesuai dengan keputusan yang diharapkan. Untuk menguji kebenaran dari hasil pengolahan data yang dilakukan secara manual, maka dapat menggunakan salah satu software aplikasi RapidMiner 5.3.0

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi.

RapidMiner Merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

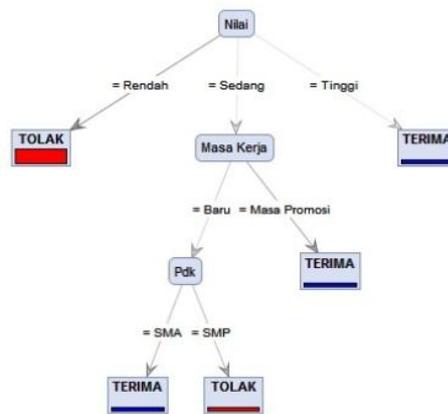
RapidMiner sebagai software pengolah datamining menyediakan tool untuk membuat decision tree. Hal ini tentu akan memudahkan kita membuat decision tree dengan menggunakan RapidMiner dibandingkan membuat decision tree secara manual yaitu dengan melakukan perhitungan menggunakan algoritma C4.5

Untuk memudahkan dalam menggunakan RapidMiner untuk membuat decision tree, kita gunakan data sederhana. Pertama-tama data pada table dibuat lagi dalam format excel seperti yang terlihat pada Tabel 7

Tabel 7. Tabel Keputusan dalam format xls

| No | Grade | Nilai | Nama Pegawai | Jabatan | Pendidikan | Masa Kerja | HASIL |
|----|-------|--------|-------------------|---------------|------------|--------------|--------|
| 1 | B | Tinggi | Mulyaman Iskandar | Operator | SMA | Masa Promosi | TERIMA |
| 2 | H | Rendah | Prisiyatun | Operator Jahl | SMP | Masa Promosi | TOLAK |
| 3 | H | Sedang | Elli Susanti | Operator Jahl | SMP | Masa Promosi | TERIMA |
| 4 | H | Tinggi | Erlinda | Operator | SMA | Masa Promosi | TERIMA |
| 5 | H | Rendah | Suryana | Operator Jahl | SMA | Masa Promosi | TOLAK |
| 6 | H | Rendah | Widha Hendrawati | Operator Jahl | SMP | Masa Promosi | TOLAK |
| 7 | H | Rendah | Eka Jayanti | Operator Jahl | SMP | Masa Promosi | TOLAK |
| 8 | K | Rendah | Sally Siregar | Operator | SMA | Baru | TOLAK |
| 9 | K | Rendah | Anih Hanapiah | Helper | SMA | Baru | TOLAK |
| 10 | K | Rendah | Reni Suherni | Operator Jahl | SMA | Baru | TOLAK |
| 11 | K | Rendah | Enah | Operator Jahl | SMP | Baru | TOLAK |
| 12 | K | Rendah | Maryani | Operator Jahl | SD | Baru | TOLAK |
| 13 | K | Rendah | Nengsih | Operator Jahl | SD | Baru | TOLAK |
| 14 | K | Rendah | Nuraini | Operator Jahl | SMA | Baru | TOLAK |
| 15 | K | Rendah | Siti Soleha | Operator Jahl | SMP | Baru | TOLAK |
| 16 | K | Rendah | Erni Avrianti | Operator Jahl | SD | Baru | TOLAK |
| 17 | K | Rendah | Ria Ratnawati | Operator Jahl | SMA | Baru | TOLAK |

Setelah data yang kita punya dibuat dalam bentuk tabel format xls, Selanjutnya lakukan Importing Data kemudian tarik atau drag pada main process kemudian klik Import Configuration Wizard untuk import data.



Gambar 2. Hasil Graphic View Decision Tree

Selain menampilkan hasil decision tree berupa graph atau tampilan pohon keputusan, RapidMiner juga menyediakan tool untuk menampilkan hasil berupa teks view dengan mengklik button Text View seperti yang tampak pada Gambar 3.

```

Tree
Nilai = Rendah: TOLAK {TERIMA=0, TOLAK=566}
Nilai = Sedang
| Masa Kerja = Baru
| | Pdk = SMA: TERIMA {TERIMA=2, TOLAK=0}
| | Pdk = SMP: TOLAK {TERIMA=0, TOLAK=2}
| Masa Kerja = Masa Promosi: TERIMA {TERIMA=11, TOLAK=0}
Nilai = Tinggi: TERIMA {TERIMA=12, TOLAK=0}
  
```

Gambar 3. Hasil Text View Decision Tree

Pembahasan

Perbandingan Antara Algoritma C4.5 dengan Tools RapidMiner

Dari hasil penelitian perhitungan Algoritma C4.5 dengan menggunakan Tools RapidMiner terdapat perbedaan sedikit hasil, yaitu:

- a. Pada perhitungan Algoritma C4.5 manual atau dengan cara menghitung dengan rumus dapat mengetahui secara detail atribut mana yang di jadikan node awal yang

paling penting dalam mengambil keputusan pengangkatan karyawan kontrak menjadi tetap dengan mengetahui gain yang tertinggi, dan dapat mengetahui ada beberapa karyawan yang semula kontrak berdasarkan penilaian menjadi karyawan tetap sesuai kebutuhan yang ada pada perusahaan tersebut.

- b. Pada perhitungan menggunakan tools RapidMiner dari hasil penelitian tidak berbeda jauh hanya memunculkan hasil *view Graphic* dari data keryawan yang dimasukan memakai format xls, juga tidak dapat mengetahui *information gain* secara detail.

Analisis Data dengan Confusion Matrix

Confusion matrix mempresentasikan hasil evaluasi model dengan menggunakan table matriks, jika data set terdiri dari dua kelas, kelas pertama dianggap *positive*, dan kelas kedua dianggap *negative*. Evaluasi menggunakan *confusion matrix* menghasilkan nilai *accuracy*, *precision*, *recall*, dan *error rate*. Akurasi dalam klasifikasi merupakan presentase ketepatan record data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. *Precision* atau *confidence* merupakan proporsi kasus yang diprediksi positif yang juga positif benar pada data sebenarnya. *Recall* merupakan proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar. *Error rate* merupakan kasus yang diidentifikasi salah dengan sejumlah semua kasus.

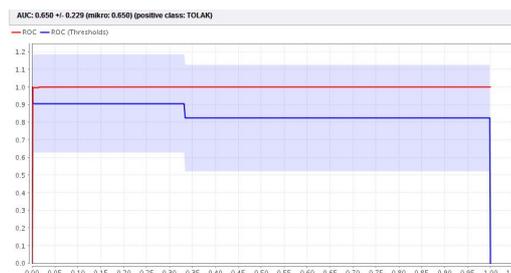
Tabel 8 Prediction Decision Tree

| Prediction | Decision Tree dengan Information Gain Ratio |
|------------|---|
| Akurasi | 99,49% |
| Precision | 99,65% |
| Recall | 99,82% |
| Error Rate | 0,51% |

Decision tree dengan information gain ratio mendapatkan akurasi yaitu sebesar 99,49%, *precision* 99,65%, *recall* 99,82%, dan *error rate* 0,51%.

Area Under Curve

Selain *confusion matrix*, kurva Receiver Operating Characteristic (ROC) dihasilkan oleh RapidMiner 5.3.0 kurva tersebut dapat dilihat pada gambar 4.



Gambar 4. Hasil AUC by RapidMiner

Kurva Receiver Operating Characteristic (ROC) digunakan untuk mengekspresikan data confusion matrix. Garis horizontal mewakili nilai *false positive* (FP) dan garis vertikal mewakili nilai *true positive* (TP). Dari gambar 4.20 dapat diketahui bahwa nilai Area Under Curve (AUC) model algoritma C4.5 adalah 0,650.

Kesimpulan

Dari hasil penelitian ini dapat disimpulkan sebagai berikut :

1. Variabel yang digunakan pada akurasi penetapan pengangkatan karyawan tetap yaitu Nilai, Pendidikan, dan Masa Kerja.
2. Dari hasil klasifikasi menggunakan metode Algoritma C4.5 dapat mempertimbangkan dari atribut Penilaian dengan nilai Tinggi mendapatkan karyawan dari atribut Masa kerja dengan nilai Masa Promosi 11 karyawan, dan dari atribut Pendidikan dengan nilai SMA mendapatkan 2 karyawan.
3. Penetapan karyawan tetap dari karyawan kontrak dari mempertimbangkan variabel yang telah ditentukan dengan akurasi 99,49 %

Saran

Adapun saran-saran yang dapat disampaikan yaitu:

1. Dapat dikembangkan dengan menggunakan aplikasi agar memaksimalkan waktu proses.
2. Menggunakan spesifikasi data yang representatif agar mendapatkan hasil yang lebih akurat.

Daftar Pustaka

- Akhtar, Fareed dan Caroline Hahne. 2012. *RapidMiner5 Operator Reference*, [online], (www.rapid-i.com, diakses tanggal 30 Desember 2016).
- Aprilla Dennis C, Aji Donny Baskoro, and Ambarwati Lia. *BELAJAR DATA MINING DENGAN RAPIDMINER*
- Ardiyanto, Dwi Puro. *SISTEM PENDUKUNG KEPUTUSAN PENETAPAN KENAIKAN GAJI KARYAWAN DENGAN MENGGUNAKAN METODE ALGORITMA C4.5 PADA PT. HARAPAN MAKMUR SURAKARTA*. Surakarta
- Ayu, Dwi Nursela (2010). *PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI TINGKAT KEGANASAN KANKER PAYUDARA*. Universitas Dian Nuswantoro: Semarang
- Kusrini and Taufiq, Emha Luthfi. 2009. *ALGORITMA DATA MINING*. ANDI. Yogyakarta