

## ANALISA PENENTUAN KARYAWAN TERBAIK MENGGUNAKAN DATA MINING DENGAN METODE ALGORITMA C4.5 DI PT.SHEI TAI INDUSTRIAL

U Darmanto Suer, Mohammad Faqih Mustofa

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa  
darmanto.soer@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

### Abstrak

PT.Shei Tai Industrial adalah sebuah perusahaan yang bergerak dalam bidang otomotif. Perusahaan ini memiliki karyawan yang mencapai kurang lebih 1.000 orang. PT.Shei Tai Industrial melakukan pemilihan karyawan terbaik untuk memacu semangat karyawan dalam meningkatkan dedikasi dan kinerjanya. Pemilihan karyawan terbaik dilakukan secara periodik akan tetapi belum optimal dalam pelaksanaannya. PT.Shei Tai Industrial mendapat kendala dalam memutuskan karyawan yang akan diprioritaskan. Kendala yang dihadapi adalah manager tidak menggunakan metode yang dapat menangani permasalahan prioritas dengan banyak kriteria. Selain itu, sering kali SDM kesulitan memilih karyawan terbaik dikarenakan banyaknya karyawan yang dinilai. Hal ini menjadi sebuah kekurangan untuk menentukan tepat atau tidaknya seseorang terpilih sebagai karyawan terbaik. Penerapan data mining dengan metode algoritma C4.5 yang dilakukan dapat mempercepat dan tingkat keakuratan dalam pengambilan keputusan terhadap penilaian karyawan. Hasil uji coba dengan algoritma C4.5 memiliki nilai accuracy, precision dan recall yang bagus yaitu dengan accuracy 91.21%, precision 85.71% dan recall 84.00% serta kurva ROC optimistic dengan akurasi excellent classification sebesar 0.982 %.

Kata kunci: algoritma C4.5, data mining, decision tree

### Abstract

*PT.Shei Tai Industrial is a company engaged in the automotive sector. This company has employees that reach approximately 1,000 people. PT.Shei Tai Industrial selects the best employees to spur employee morale in improving their dedication and performance. Selection of the best employees is carried out periodically but has not been optimal in the implementation. PT.Shei Tai Industrial faces problems in deciding which employees to prioritize. The obstacle faced is that managers do not use methods that can handle priority problems with many criteria. In addition, HR often finds it difficult to choose the best employees because of the many employees being assessed. This becomes a drawback in determining whether or not someone is selected as the best employee. The application of data mining with the C4.5 algorithm method that is carried out can accelerate and the level of accuracy in making decisions on employee assessments. The test results with the C4.5 algorithm have good accuracy, precision and recall values, namely 91.21% accuracy, 85.71% precision and 84.00% recall*

*and an optimistic ROC curve with an excellent classification accuracy of 0.982%.*

*Keywords: C4.5 algorithm, data mining, decision tree*

### 1. Pendahuluan

Didalam dunia pekerjaan adanya karyawan terbaik menjadi tolak ukur kemajuan dari perusahaan itu sendiri. Dalam penentuan biasanya dengan melihat kinerja karyawan tersebut missal dari kerajinan, kedisiplinan dan juga prestasi lainnya. Dengan cara seperti ini agak kurang efektif dan akurat, oleh karena itu dalam penelitian menganalisa penentuan karyawan terbaik dengan metode algoritma C4.5 agar hasil yang didapat lebih akurat dan memudahkan pihak perusahaan yang sangat penting adalah Sumber Daya Manusia (SDM). Pengelolaan SDM dari suatu perusahaan sangat mempengaruhi banyak aspek penentu keberhasilan kerja dari perusahaan tersebut. Jika SDM dapat diorganisir dengan baik maka diharapkan perusahaan dapat menjalankan semua proses usahanya dengan baik (Larose, Daniel.T.2005).

PT.Shei Tai Industrial adalah sebuah perusahaan yang bergerak dalam bidang otomotif. Perusahaan ini memiliki karyawan yang mencapai kurang lebih 1.000 orang. PT.Shei Tai Industrial melakukan pemilihan karyawan terbaik untuk memacu semangat karyawan dalam meningkatkan dedikasi dan kinerjanya. Pemilihan karyawan terbaik dilakukan secara periodik akan tetapi belum optimal dalam pelaksanaannya. PT.Shei Tai Industrial mendapat kendala dalam memutuskan karyawan yang akan diprioritaskan. Kendala yang dihadapi adalah manager tidak menggunakan metode yang dapat menangani permasalahan prioritas dengan banyak kriteria. Selain itu, sering kali SDM kesulitan memilih karyawan terbaik dikarenakan banyaknya karyawan yang dinilai. Hal ini menjadi sebuah kekurangan untuk menentukan tepat atau tidaknya seseorang terpilih sebagai karyawan terbaik.

Sumber daya manusia merupakan aset terpenting perusahaan karena perannya sebagai subjek pelaksana kebijakan dan kegiatan operasional perusahaan, agar perusahaan tetap eksis maka harus berani menghadapi tantangan dan implikasinya yaitu menghadapi perubahan dan memenangkan persaingan. Sumber daya yang dimiliki oleh perusahaan seperti modal, metode dan mesin tidak bisa memberikan hasil yang maksimal apabila tidak didukung oleh sumber daya manusia yang mempunyai kinerja yang optimum. Douglas,2000:34 menjelaskan bahwa perusahaan membutuhkan karyawan yang mempunyai kinerja yang tinggi.

Menurut Setiyawan, Budi dan Waridin, 2006 kinerja karyawan merupakan hasil atau prestasi kerja karyawan yang di nilai dari segi kualitas maupun kuantitas berdasarkan standar kerja yang ditentukan oleh pihak organisasi. Kinerja yang baik adalah kinerja yang optimal, yaitu kinerja yang sesuai standar organisasi dan mendukung tercapainya tujuan organisasi. Penentuan karyawan terbaik dilakukan tidak hanya dengan cara penunjukan langsung oleh pimpinan atau manager, akan tetapi sebuah perusahaan harus melakukan penilaian kinerja yang telah dilakukan oleh karyawannya dalam jangka waktu tertentu, dan tentunya akan ada sebuah rewards atas keberhasilan yang telah dicapai oleh karyawannya. Hal ini dimaksudkan untuk mendorong setiap karyawannya untuk selalu memberikan kinerja yang terbaik bagi perusahaan dalam melaksanakan tugas dan kewajibannya di perusahaan.

Melihat pentingnya pemilihan karyawan terbaik bagi perusahaan maupun karyawan, maka proses penilaian karyawan harus dilakukan secara objektif dan menyeluruh. Penilaian karyawan yang objektif akan memberikan umpan balik antara

atasan dan bawahan yang sinergi. Bawahan lambat laun akan memahami objektifitas kerja dan mampu mendorong produktivitas perusahaan.

## 2. Tinjauan Studi

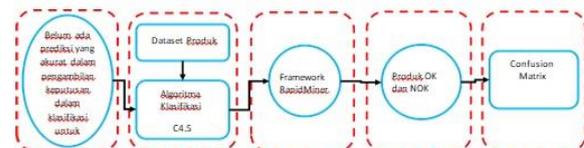
Secara sederhana data mining adalah pengembangan atau penemuan informasi baru dengan mencari pola atau ukuran tertentu dari sejumlah data yang sangat besar (Davies,2004). Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono,2007). Data mining, sering juga disebut Knowledge discovery in database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santoso,2007).

Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti database system, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi.

Data mining didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar (Witten,2005).

## 3. Kerangka Studi

Adapun kerangka pemikiran dari penelitian yang dilakukan adalah sebagai berikut:



## 4. Desain Penelitian/Metodologi

Dalam penyusunan penelitian ini penulis melakukan penelitian tentang data mining mengenai penentuan karyawan terbaik pada sebuah perusahaan yang bergerak di bidang otomotif yaitu PT.Shei Tai Industrial, dimana penelitian ini untuk mengetahui ke akuratan dalam pemilihan karyawan terbaik dalam perusahaan tersebut.

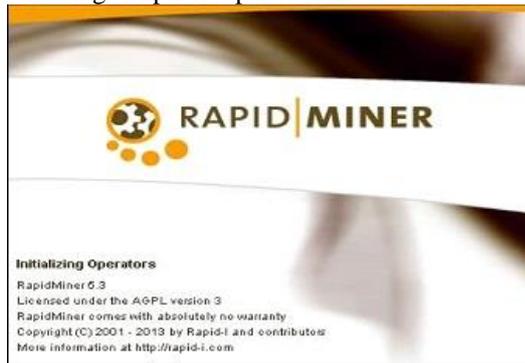
Adapun faktor yang diambil dalam acuan penilaian ini adalah seperti Prestasi Kerja, Kemampuan, Kepribadian, Kepemimpinan dan Absensi dalam bekerja.

## 5. Hasil Penelitian dan Pengujian

### 5.1 Proses Mining

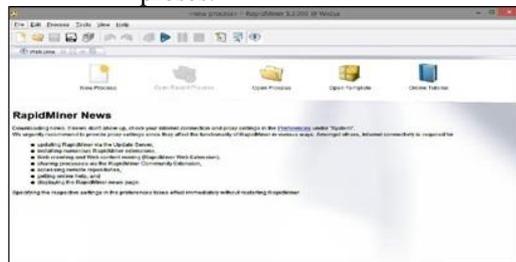
Pada tahap ini metode data mining diterapkan untuk menemukan pengetahuan tersembunyi dan berharga dari data. Metode yang digunakan adalah klasifikasi dengan algoritma C4.5. Berikut penerapan algoritma C4.5 memakai tool RapidMiner.

1. Langkah awal yaitu membuka aplikasi RapidMiner akan muncul loading tampilan seperti berikut :



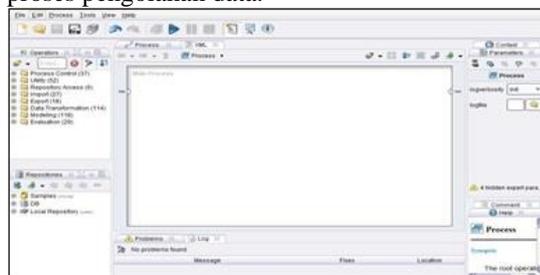
Gambar 5.2 Tampilan awal RapidMiner

Setelah loading selesai akan muncul tampilan menu utama sebelum proses.



Gambar 5.3 Tampilan Menu utama sebelum proses

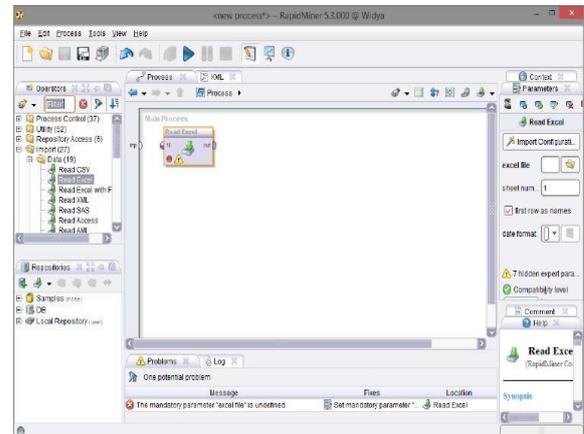
Setelah muncul tampilan menu utama, tekan New Process, Blank untuk memulai proses pengolahan data.



Gambar 5.4 Tampilan Menu proses

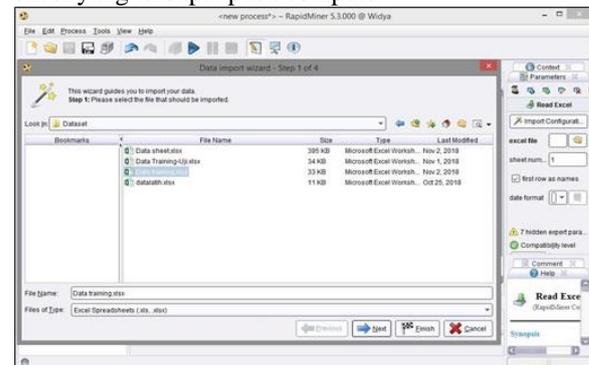
Setelah muncul tampilan Process selanjutnya pada menu operator cari *import* kemudian pilih data klik dan drag *read Excel* ke Main process hingga

muncul kotak *read Excel* pada *process*.



Gambar 5.5 Tampilan *Process read Excel*

Kemudian klik *Import Configuration Wizard* untuk mengambil data *training* dengan format data excel yang terdapat pada computer.



Gambar 5.6 Data *Import Wizard*

### 2. Mengolah data *training*

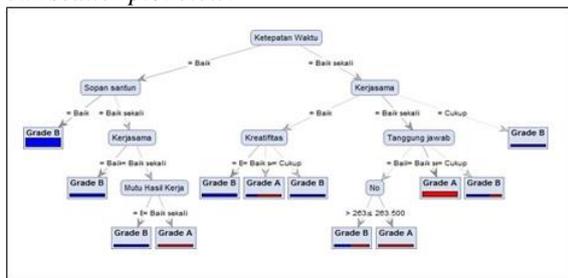
Menyiapkan data *training* dari data penilaian karyawan, karna banyaknya *record* maka penulis hanya akan menguji untuk karyawan yang mendapatkan hasil grade A dan grade B. Pada data *import wizard* klik data *training* lalu klik *next*.



grade A true grade A adalah 84.44%.

- b. Prediksi grade B, dengan true grade B mencapai 119 orang dan true grade A mencapai 7 orang, dengan pencapaian class precision 94.44%. Sedangkan untuk class recall grade B true grade B adalah 93.70% dan class recall grade A true grade A adalah 84.44%.

Hasil proses klasifikasi predicate kelulusan dengan metode C4.5 atau pohon keputusan ditunjukkan seperti gambar 5.13 diatas dapat dilihat bahwa atribut yang memiliki pengaruh paling tinggi untuk menentukan klaifikasi predikat karyawan terbaik adalah ketepatan waktu yang mana atribut ini menjadi node akar pertama. Selain menampilkan hasil C4.5 berupa graph atau tampilan pohon keputusan. RapidMiner juga menyediakan tool untuk menampilkan hasil berupa text view, table dan scatter plot view.



```

Tree
Ketepatan Waktu = Baik
| Sopan santun = Baik: Grade B {Grade B=87, Grade A=0}
| Sopan santun = Baik sekali
| | Kerjasama = Baik: Grade B {Grade B=13, Grade A=0}
| | Kerjasama = Baik sekali
| | | Mutu Hasil Kerja = Baik: Grade B {Grade B=4, Grade A=0}
| | | Mutu Hasil Kerja = Baik sekali: Grade A {Grade B=0, Grade A=3}
Ketepatan Waktu = Baik sekali
| Kerjasama = Baik
| | Kreatifitas = Baik: Grade B {Grade B=14, Grade A=0}
| | Kreatifitas = Baik sekali: Grade A {Grade B=1, Grade A=2}
| | Kreatifitas = Cukup: Grade B {Grade B=3, Grade A=0}
| Kerjasama = Baik sekali
| | Tanggung jawab = Baik
| | | No > 263.500: Grade B {Grade B=1, Grade A=1}
| | | No <= 263.500: Grade A {Grade B=0, Grade A=3}
| | Tanggung jawab = Baik sekali: Grade A {Grade B=0, Grade A=35}
| | Tanggung jawab = Cukup: Grade B {Grade B=2, Grade A=1}
| Kerjasama = Cukup: Grade B {Grade B=2, Grade A=0}
    
```

Gambar 5.18 Hasil dari text view

Seperti gambar 5.17 dapat dilihat penjelasan RapidMiner menggunakan text view untuk menjelaskan berapa jumlah ketepatan waktu yang baik sekali, baik, cukup dan kurang.

**Daftar Pustaka**

Aprilla C, Dennis., Baskoro, Donny Aji., Ambarwati, Lia., Wicaksana, I Wayan Simri. (2013). *Belajar Data Mining*

Dengan RapidMiner, 5(4), 1–5. <https://doi.org/10.1007/s13398-014-0173-7.2>

Cahyadi, D. (2014). *Analisis Parameter Operasi Pada Proses Plastik Injection Molding Untuk Pengendalian Cacat Produk*. Jurnal Mesin Teknologi, 8(2), 8–

16. Retrieved from <https://jurnal...>

Dai, Wei., Ji, W. (2014). *A Map Reduce Implementation of C4.5 Decision Tree Algorithm*, 7(1), 49–60.

Dewi, Shanty Kusuma. (2012). *Minimasi Defect Produk Dengan Konsep Six Sigma*. Jurnal Teknik Industri, 13, 43–50.

Gullo, Francesco. (2015). *From Patterns in Data to Knowledge Discovery: What Data Mining Can Do*. Physics Procedia, 62, 18–22. <https://doi.org/10.1016/j.phpro.2015.02.005>

Han, Jiawei., Kamber, Micheline ., Pei, Jian. (2012). *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0>

Kotu, Vijay., Deshpanda, Bala. (2015). *Predictive Analytics and Data Mining*. In *Predictive Analytics and Data Mining* (pp. 1–16). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-801460-8.00001-X>

Mardi, Yuli. (2017). *Data Mining : Klasifikasi Menggunakan Algoritma C4.5*.

Jurnal Edik Informatika, 2(2), 213–219.

Novandya, Adhika., Oktria, Isni. (2017). *Penerapan Algoritma Klasifikasi Data Mining C4.5 Pada Dataset Cuaca Wilayah Bekasi*. Jurnal Format, 6(2), 98–106.

Raharja, Yosoa Putra. (2014). *Rancang Bangun Sistem Rekomendasi Basiswa Menggunakan ALgoritma Klasifikasi C4.5 pada Universitas DianNuswantoro*. Undinus, 1–4. Retrieved from <http://eprints.dinus.ac.id/13408/>

Rohman, Abdul. (2016). *Komporsi Metode Klasifikasi Data Mining Untuk Prediksi Penyakit Jantung*. Neo Teknika, 2(2), 21–28.

Saifudin, Aries., Wahono, Romi Satrio. (2015). *Penerapan Teknik Ensemble untuk Menangani Ketidakseimbangan Kelas pada Prediksi Cacat Software*. Journal of Software Engineering, 1(1), 28–37. <https://doi.org/10.1016/S1896->

1126(14)00030-3

- Saleh, Hamsir. (2017). *Prediksi Kebangrutan Perusahaan Menggunakan Algoritma C4.5 Berbasis Forward Selection*, 9, 173–180.
- Silalahi, Kristiani Desri., Murfi, Hendri., Satria, Yudi. (2017). *Studi Perbandingan Pemilihan Fitur untuk Support Vector Machine pada Klasifikasi Penilaian Risiko Kredit*, 1(2), 119–136.
- Sinambela, Mutiara Shany., Mayadewi, RA Paramita., Rosely, Ely. (2016). *Sistem Pendukung Keputusan Peminatan Jurusan Siswa Di SMA Menggunakan Algoritma Klasifikasi Data Mining C4.5*, 2(3), 858–866.
- Singh, Sonia., Gupta, Prinyanka. (2014). *Comparative Study ID3 , CART And C4.5 Decision Tree Algorithm : A Survey*, 27(27), 97–103.
- Sunge, Aswan Supriyadi. (2018). *Prediksi Kompetensi Karyawan Menggunakan Algoritma C4.5 ( Studi Kasus: PT Hankook Tire Indonesia )*, 2018(Sentika), 23–24.
- Suyanto. (2017). *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Penerbit: Informatika Bandung. (ISBN): 978-602-6232-36-6.
- Wahyono., Nugroho, Agung. (2018). *Penerapan Algoritma C4.5 Untuk Prediksi Tingkat Kompetensi Karyawan PT Multistrada Arah Sarana*, 3(1), 1–6.
- Witten, Ian H., Frank, Eibe. (2011). *Data Mining: Practical Machine Learning Tools and Techniques* (Google eBook). Complementary literature None. <https://doi.org/0120884070>, 9780120884070