PERANCANGAN APLIKASI PEMILIHAN MITRA KERJA

Harry Dhika[1)], Tri Yani Akhirina[2)], Surajiyo[3)]

[1), 2),3)] Teknik Informatika Universitas Indraprasta PGRI
Jl Nangka No. 58B, Jakarta, DKI Jakarta 12530
Email : dhikatr@yahoo.com[1)], azizahputriku@gmail.com[2)], drssurajiyo@yahoo.co.id[3)]

Abstrak

Aplikasi pemilihan mitra kerja dibuat dengan metode Knowledge Discovery in Database (KDD) yang merupakan framework dari data mining. Algoritma yang diterapkan dalam KDD adalah algoritma C4.5. algoritma berupa pohon keputusan yang menentukan berdasarkan data yang diinput apakah mitra kerja dapat dipilih atau tidak. Aplikasi ini juga bersifat open source dibuat dengan tools Java Netbeans. Data dianalisis menggunakan perangkat lunak Rapidminer. Tingkat akurasi data yang diperoleh adalah 94,37% dengan class recall mencapai 90.91% terhadap nilai True Very Good. Untuk nilai class recall true good nilai akurasi mencapai 94.12% sedangkan untuk class recall true bad nilainya mencapai 100%.rule yang tercipta dari pengukuran tingkat akurasi data diterapkan dalam aplikasi pemilihan mitra kerja.

Kata kunci: *Algoritma C4.5, Data Mining, KDD, Aplikasi Pemilihan Mitra Kerja.*

1. Pendahuluan

Saat ini semakin banyak berkembang perusahaan yang bergerak pada bidang jasa pengiriman barang untuk menjangkau pulau-pulau tersebut. Banyak perusahaan transportasi yang bergerak dibidang jasa pengiriman barang. Perusahaan ini selalu melayani pelanggannya ketika pelanggan tersebut membutuhkan jasa untuk pengiriman barang ke daerah atau wilayah lain atau pulau-pulau lain. Dalam melakukan pengiriman barang, perusahaan transportasi membutuhkan sarana yang dapat menjangkau wilayah-wilayah pengiriman tujuan barang. Oleh karena itu, perusahaan ini melakukan kerja sama dengan perusahaan-perusahaan penyedia jasa transportasi baik darat, laut maupun udara.

Karena pada tahun-tahun sebelumnya, perusahaan tersebut selalu menggunakan cara konvensional dalam menggunakan alat transportasi terhadap semua pemesanan jasa pengiriman barang dan selalu menimbulkan masalah yaitu kendala kecelakaan atau kurang layaknya alat transportasi yang dapat menyebabkan barang hilang bahkan rusak.

Melihat masalah-masalah yang ada maka digunakan metode pengambilan keputusan dengan algoritma C4.5 untuk menetukan pemilihan mitra kerja penyedia jasa transportasi jasa pengiriman barang yang baik pada perusahaan transportasi di Jakarta.

Pemilihan mitra kerja penyedia jasa transportasi barang yang baik berdasarkan pengalaman kerja, harga jasa pengiriman, legalitas perusahaan mitra, jumlah armada yang dimiliki, dan intensitas pelayanan kerja mitra dalam melayani jasa pengiriman barang.

Pengambilan keputusan menggunakan menggunakan metode algoritma C4.5 dalam pengolahan data akan dilakukan menggunakan software pendukung yaitu Rapidminer versi 5.2.008. Menerapkan algoritma data mining C4.5 pada data mitra kerja penyedia jasa transportasi. Pada tahapan ini dilakukan kajian mendalam terhadap data dari mitra kerja penyedia jasa transportasi, pengumpulan data dilakukan pada sepuluh perusahaan yang menyediakan fasilitas pengiriman barang. Dilakukan pengolahan data, menyusun data yang diperoleh dalam bentuk tabel dan menggunakan algoritma sebagai prinsip pengolahan data.

Algoritma C4.5

Algoritma C4.5 merupakan algoritma klasifikasi dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Kelebihan ini misalnya dapat mengolah data numerik (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diintrepetasikan dan tercepat diantara algoritma-algoritma yang lain. Keakuratan prediksi yaitu kemampuan model untuk dapat memprediksi label kelas terhadap data baru atau yang belum diketahui sebelumnya dengan baik. Dalam hal kecepatan atau efisiensi waktu komputasi yang diperlukan untuk membuat dan menggunakan model. Kemampuan model untuk memprediksi dengan benar walaupun data ada nilai dari atribut yang hilang. Dan juga skalabilitas yaitu kemampuan untuk membangun model secara efisien untuk data berjumlah besar (aspek ini akan mendapatkan penekanan). Terakhir interpretabilitas yaitu model yang dihasilkan mudah dipahami. Selain itu dapat juga dikatakan bahwa Algoritma merupakan kumpulan perintah yang tertulis secara sistematis guna menyelesaikan permasalahan logika dari matematika. Pengertian Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [4], yaitu:

1. Menyiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
2. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus:

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

Keterangan:
S = himpunan kasus
n = jumlah partisi S
p_i = proporsi S_i terhadap S

3. Kemudian hitung nilai gain menggunakan rumus:

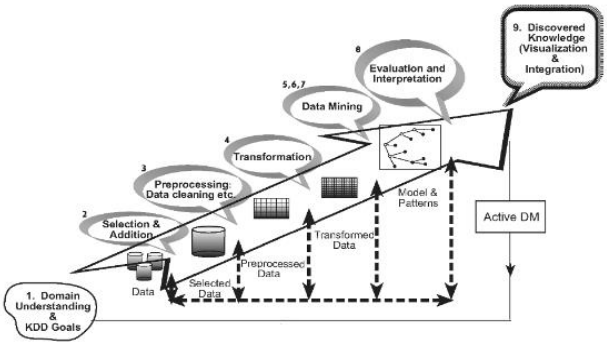
$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|si|}{|s|} * Entropy\ (Si) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Keterangan:
S = himpunan kasus
A = fitur
n = jumlah partisi atribut A
$\left|\frac{S_i}{S}\right|$ = proporsi S_i terhadap S
$|S|$ = jumlah kasus dalam S

4. Ulangi langkah ke-2 hingga semua record terpartisi.
5. Proses partisi pohon keputusan akan berhenti saat :
   a. Semua record dalam simpul N mendapat kelas yang sama.
   b. Tidak ada atribut di dalam record yang dipartisi lagi.
   c. Tidak ada record di dalam cabang yang kosong.

Metode Penelitian

Obyek Penelitian ini adalah perusahaan pengiriman barang di Indonesia khususnya di DKI Jakarta. Dari perusahaan yang ada di DKI Jakarta penelitian ini akan mengabil sample pada 10 (sepuluh) perusahaan di wilayah DKI Jakarta yang membutuhkan mitra kerja penyedia jasa transportasi.

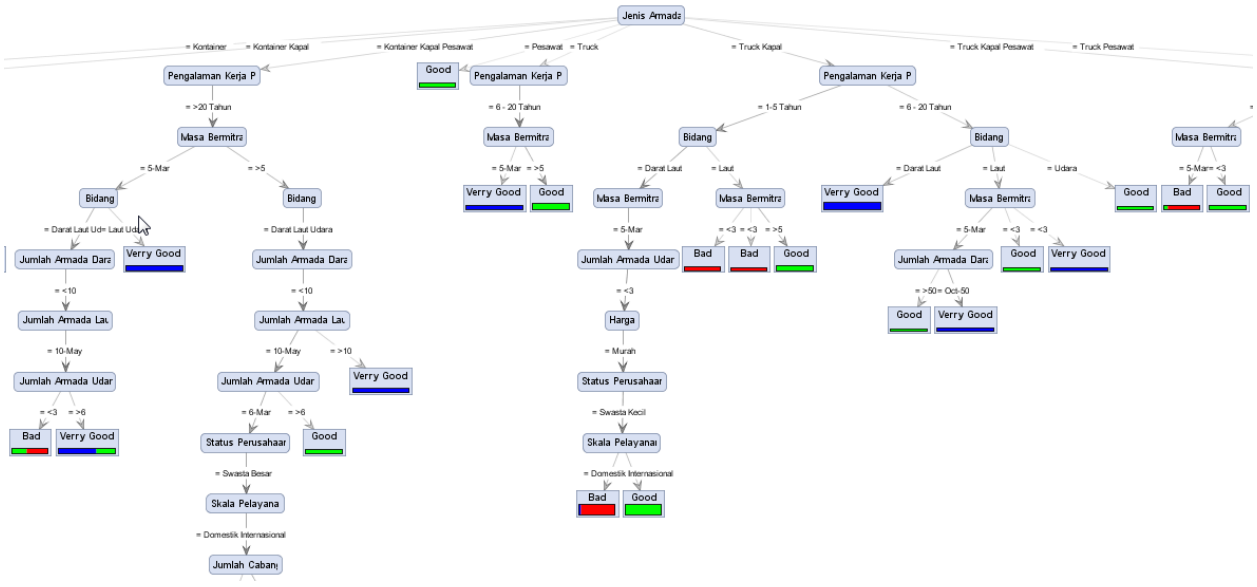


Gambar 1. *Proses KDD (Knowledge Discovery In Databases)*

Penelitian dilakukan dengan menggunakan konsep dari data mining menggunakan *Knowledge Discovery in Database* (KDD) yakni sebuah kerangka kerja sebagai bahan acuan dalam pengkajian dibidang data mining [5].

2. Pembahasan

Berikut kriteria penggolongan data yang dilakukan terhadap perusahaan dalam menentukan mitra kerja penyedia jasa transportasi. Setelah dilakukan pengolahan dan penyesuaian data berdasarkan algoritma C4.5 maka diperoleh data seperti pada gambar 3, jenis armada merupakan simpul tertinggi yang diperoleh dari nilai gain index kemudian simpul pengalaman kerja hingga selanjutnya remark *Verry Good*, *Good* dan *Bad*. Berikut ini adalah rule yang dihasilkan oleh algoritma C4.5:



Gambar 2.*Rule Pohon Keputusan Pengolahan data dengan Algoritma C4.5*

Pada pohon keputusan gambar 2. Jika dibuat rule menghasilkan algaritma sebagai berikut:

Jenis Armada = Kontainer
| Pengalaman Kerja PT = >20 Tahun
| | Masa Bermitra = 3-5: Verry Good {Verry Good=10, Good=0, Bad=0}
| | Masa Bermitra = >5: Good {Verry Good=0, Good=28, Bad=0}
Jenis Armada = Kontainer Kapal
| Pengalaman Kerja PT = 1-5 Tahun: Bad {Verry Good=0, Good=0, Bad=2}
| Pengalaman Kerja PT = >20 Tahun
| | Masa Bermitra = 3-5: Good {Verry Good=0, Good=6, Bad=0}
| | Masa Bermitra = >5
| | | Bidang = Darat Laut: Bad {Verry Good=5, Good=0, Bad=22}
| | | Bidang = Laut: Good {Verry Good=0, Good=6, Bad=0}
Jenis Armada = Kontainer Kapal Pesawat
| Pengalaman Kerja PT = >20 Tahun
| | Masa Bermitra = 3-5
| | | Bidang = Darat Laut Udara
| | | | Jumlah Armada Darat = <10
| | | | | Jumlah Armada Laut = 5-10
| | | | | | Jumlah Armada Udara = <3: Bad {Verry Good=0, Good=5, Bad=7}
| | | | | | Jumlah Armada Udara = >6: Verry Good {Verry Good=8, Good=4, Bad=0}
| | | Bidang = Laut Udara: Verry Good {Verry Good=12, Good=0, Bad=0}
| | Masa Bermitra = >5
| | | Bidang = Darat Laut Udara
| | | | Jumlah Armada Darat = <10
| | | | | Jumlah Armada Laut = 5-10
| | | | | | Jumlah Armada Udara = 3-10
| | | | | | | Status Perusahaan = Swasta Besar
| | | | | | | | Skala Pelayanan = Domestik Internasional
| | | | | | | | | Jumlah Cabang = 5-10: Verry Good {Verry Good=9, Good=0, Bad=0}
| | | | | | | | | Jumlah Cabang = <5: Good {Verry Good=0, Good=13, Bad=0}
| | | | | | Jumlah Armada Udara = >6: Good {Verry Good=0, Good=8, Bad=0}
| | | | | Jumlah Armada Laut = >10: Verry Good {Verry Good=8, Good=0, Bad=0}
Jenis Armada = Pesawat: Good {Verry Good=0, Good=7, Bad=0}
Jenis Armada = Truck
| Pengalaman Kerja PT = 6 - 20 Tahun
| | Masa Bermitra = 3-5: Verry Good {Verry Good=7, Good=0, Bad=0}
| | Masa Bermitra = >5: Good {Verry Good=0, Good=14, Bad=0}
Jenis Armada = Truck Kapal
| Pengalaman Kerja PT = 1-5 Tahun
| | Bidang = Darat Laut
| | | Masa Bermitra = 3-5
| | | | Jumlah Armada Udara = <3
| | | | | Harga = Murah
| | | | | | Status Perusahaan = Swasta Kecil
| | | | | | | Skala Pelayanan = Domestik: Bad {Verry Good=2, Good=0, Bad=29}
| | | | | | | Skala Pelayanan = Domestik Internasional: Good {Verry Good=0, Good=30, Bad=0}
| | Bidang = Laut
| | | Masa Bermitra = <3: Bad {Verry Good=0, Good=0, Bad=6}
| | | Masa Bermitra = <3 : Bad {Verry Good=0, Good=0, Bad=3}
| | | Masa Bermitra = >5: Good {Verry Good=0, Good=12, Bad=0}
| Pengalaman Kerja PT = 6 - 20 Tahun
| | Bidang = Darat Laut: Verry Good {Verry Good=17, Good=0, Bad=0}
| | Bidang = Laut
| | | Masa Bermitra = 3-5
| | | | Jumlah Armada Darat = >50: Good {Verry Good=0, Good=3, Bad=0}
| | | | Jumlah Armada Darat = 10-50: Verry Good {Verry Good=7, Good=0, Bad=0}
| | | Masa Bermitra = <3: Good {Verry Good=0, Good=5, Bad=0}
| | | Masa Bermitra = <3 : Verry Good {Verry Good=3, Good=0, Bad=0}
| | Bidang = Udara: Good {Verry Good=0, Good=5, Bad=0}
Jenis Armada = Truck Kapal Pesawat
| Pengalaman Kerja PT = 1-5 Tahun
| | Masa Bermitra = 3-5: Bad {Verry Good=0, Good=1, Bad=7}
| | Masa Bermitra = <3: Good {Verry Good=0, Good=7, Bad=0}
| Pengalaman Kerja PT = 6 - 20 Tahun
| | Masa Bermitra = 3-5
| | | Bidang = Laut Udara
| | | | Jumlah Armada Darat = >50: Good {Verry Good=3, Good=5, Bad=0}
| | | | Jumlah Armada Darat = 10-50
| | | | | Jumlah Armada Laut = 5-10
| | | | | | Jumlah Armada Udara = 3-10: Good {Verry Good=0, Good=7, Bad=0}
| | | | | | Jumlah Armada Udara = <3: Verry Good {Verry Good=4, Good=0, Bad=0}
| | | | | Jumlah Armada Laut = >10: Verry Good {Verry Good=11, Good=0, Bad=0}
| | | Bidang = Udara: Good {Verry Good=0, Good=4, Bad=0}
Jenis Armada = Truck Pesawat: Verry Good {Verry Good=4, Good=0, Bad=0}

Dari pohon keputusan pada gambar 2 setalah diuji coba kembali menghasilkan tingkat akurasi yang cukup tinggi seperti pada Tabel 1 berikut ini:

Tabel 1.*Pengukuran Akurasi dengan Algoritma C4.5*

accuracy: 94.37% +/- 4.40% (mikro: 94.38%)				
	True Verry Good	**True Good**	**True Bad**	**Class Precision**
Pred. Verry Good	100	4	0	96.15%
Pred. Good	3	160	0	98.16%
Pred. Bad	7	6	76	85.39%
Class Recall	90.91 %	94.12 %	100.00 %	

Pada tabel 1 diperoleh nilai akurasi sebesar 94.37% melalui pengukuran menggunakan alogoritma C4.5. Selanjutnya dilakukan pengukuran tingkat akurasi dengan *confusion matrix* dengan *True positives* adalah jumlah *record* positif yang diklasifikasikan sebagai positif, *false positives* adalah jumlah *record* negatif yang diklasifikasikan sebagai positif, *false negatives* adalah jumlah *record* positif yang diklasifikasikan sebagai negatif, *true negatives* adalah jumlah *record* negatif yang diklasifikasikan sebagai negative, kemudian masukkan data uji. Setelah data uji dimasukkan ke dalam *confusion matrix*, hitung nilai-nilai yang telah dimasukkan tersebut untuk dihitung jumlah *sensitivity* (*recall*), *specificity*, *precision* dan *accuracy*. *Sensitivity* digunakan untuk membandingkan jumlah TP terhadap jumlah *record* yang positif sedangkan *specificity* adalah perbandingan jumlah TN terhadap jumlah *record* yang negatif.

$$sensitivity = \frac{\mathrm{TP}}{\mathrm{P}} \quad \text{.....(3)}$$

$$specificity = \frac{\mathrm{TN}}{\mathrm{N}} \quad \text{.....(4)}$$

$$precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad \text{.....(5)}$$

$$accuracy = sensitivity\frac{\mathrm{P}}{(\mathrm{P+N})} + specificity\frac{\mathrm{N}}{(\mathrm{P+N})} \quad \text{.(6)}$$

Hasil pengukuran menggunakan *confusion matrix* yakni sebagai berikut:

Tabel 2.*Pengukuran Confusion Matrix*

True:	**Verry Good**	**Good**	**Bad**
Verry Good:	100	4	0
Good:	3	160	0
Bad:	7	6	76
kappa:	0.912 +/- 0.068 (mikro: 0.912)		

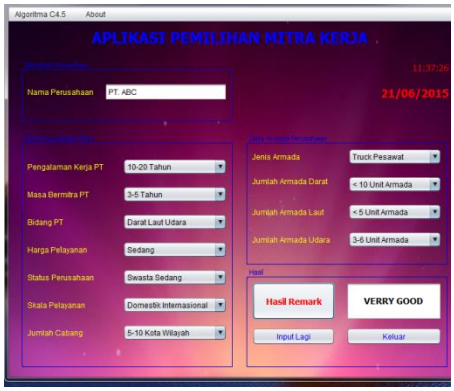
Untuk hasil pengukuran tingkat akurasi dalam data mining, nilai tersebut dapat dibagi menjadi beberapa kelompok (Gorunescu, 2011).

a. 0.90-1.00 = klasifikasi sangat baik
b. 0.80-0.90 = klasifikasi baik
c. 0.70-0.80 = klasifikasi cukup
d. 0.60-0.70 = klasifikasi buruk
e. 0.50-0.60 = klasifikasi salah

Berdasarkan pengelompokkan di atas dan tabel 1. serta tabel 2. maka dapat disimpukan bahwa metode pemilihan mitra kerja penyedia jasa transportasi dengan algoritma C4.5 termasuk klasifikasi **sangat baik** karena memiliki nilai akurasi antara 0.90-1.00.

Kemudian dari algoritma C4.5 tersebut yang diperoleh dari pohon keputusan dilakukan perancangan sistem antar muka sebagai berikut dan diterapkan terhadap mitra kerja penyedia jasa transportasi sebagai berikut:

Gambar 3. *Halaman Utama*



Gambar 4. *Uji Coba Profil Verry Good*

Gambar 5. *Uji Coba Profil Good*

Gambar 6. *Uji Coba Profil Bad*

3. Kesimpulan

Penerapan algoritma C4.5 dilakukan pada pemilihan mitra kerja jasa transportasi untuk memilih mitra kerja yang baik, dari data tersebut terbentuk pohon keputusan seperti pada gambar 2. Tingkat akurasi yang dihasilkan dengan pengolahan data tersebut mencapai 94.37%. Penerapan algoritma C4.5 dalam memilih mitra kerja dilakukan untuk mencari mitra kerja terbaik dan terpercaya.

Daftar Pustaka

[1] Contoh Kasus Prediksi layak tidaknya Indonesian menjadi Juara AFF, http://dennistorio.com/DM/bola.php diakses tanggal 29 juni 2014.
[2] Han, J.,& Kamber, M. 2006. *Data Mining Concept and Tehniques*. San Fransisco: Morgan Kauffman.
[3] Karismariyanti, Magdalena. 2011. *Jurnal: Simulasi Pendukung Keputusan Beasiswa menggunakan Metode Composite Performance Index.*
[4] Kusrini & Luthfi, E. T. 2009. *Algoritma Data Mining*. Yogyakarta: Andi Publishing.
[5] Larose, D. T. 2005. *Discovering Knowledge in Data*. New Jersey: John Willey & Sons, Inc.
[6] Liao. 2007. *Recent Advances in Data Mining of Enterprise Data: Algorithms and Application.* Singapore: World Scientific Publishing
[7] Maimon, Oded & Rokach, Lior. 2005. *Data Mining and Knowledge Discovey Handbook*. New York: Springer
[8] Sumathi, & S., Sivanandam, S.N. 2006. *Introduction to Data Mining and its Applications*. Berlin Heidelberg New York: Springer
[9] Witten, I. H., Frank, E., & Hall, M. A. 2011. *Data Mining: Practical Machine Learning and Tools*. Burlington: Morgan Kaufmann Publisher.
[10] Wu, Xindong & Kumar, Vipin. 2009. *The Top Ten Algorithms in Data Mining*. Boca Raton: CRC Press
[11] Heiat, Abbas. (2011). Modeling Consumer Credit Scoring Through Bayes Network. *World Journal of Social Sciences*, Volume 1, Number 3, pp.132-141, July 2011
[12] Leidiyana, Henny. (2011). *Tesis: Komparasi Algoritma Data Mining Dalam Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. Juli 2011.*

Biodata Penulis

Harry Dhika, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika Universitas Indraprasta PGRI, lulus tahun 2009. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Teknik Informatika STMIK Nusa Mandiri, lulus tahun 2012. Saat ini menjadi Dosen di Univ. Indraprassta PGRI.

Tri Yani Akhirina, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika Universitas Indraprasta PGRI, lulus tahun 2009. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Teknik Informatika STMIK Nusa Mandiri, lulus tahun 2011. Saat ini menjadi Dosen di Univ. Indraprassta PGRI.

Surajiyo, memperoleh gelar Sarjana Sosial (S.Pd), Jurusan Ilmu Filsafat, Universitas Gadjah Mada, lulus tahun 1987. Memperoleh gelar Magister Komunikasi (M.Si) Program Pasca Sarjana Magister Ilmu Komunikasi, lulus tahun 2009. Saat ini menjadi Dosen di Univ. Indraprassta PGRI.

ISSN : 2302-3805