

KAJIAN KOMPARASI PENERAPAN ALGORITMA C4.5, NEURAL NETWORK, DAN SVM DENGAN TEKNIK PSO UNTUK PEMILIHAN KARYAWAN TELADAN PT. XYZ

Rudi Apriyadi Raharjo

Program Studi Teknik Informatika, Universitas Indraprasta PGRI

Email: rudiapriyadiaraharjo@gmail.com

Abstrak

Dalam Meningkatkan kualitas sumber daya manusia bagi perusahaan kepada karyawannya sangatlah diperlukan untuk dapat meningkatkan pertumbuhan perusahaan karena kualitas karyawannya. Sistem Teknologi Informasi yang berkembang saat ini sangat membantu dalam pengambilan keputusan yang akurat. PT. XYZ adalah sebuah perusahaan yang bergerak di bidang Software Developer, sangat erat dengan permasalahan Sumber Daya Manusia yang terus berputar mengikuti kebutuhan pengguna jasa layanan, mengakibatkan kinerja karyawan yang kurang efektif. Penelitian ini menggunakan metode pengkajian komparasi algoritma C4.5, Neural Network dan SVM dengan menggunakan Teknik PSO menggunakan software Rapid Miner dengan membandingkan beberapa klasifikasi. Hasil dari penelitian ini adalah menghasilkan sebuah model yang dapat menunjang dalam pengambilan keputusan dalam hal menentukan karyawan teladan secara obyektif.

Kata Kunci: C4.5, neural network, SVM, PSO, Rapid Miner, Karyawan Teladan

Abstract

Improving the quality of human resources for the company to its employees is needed to be able to increase the growth of the company for the quality of its employees. Systems Information Technology is currently developing very helpful in making accurate decisions. PT. XYZ is a company engaged in the field of Software Developer, very closely with Human Resources issues that continue to rotate to follow the needs of service users, resulting in a less effective employee performance. This study uses a comparative assessment of the C4.5 algorithm, Neural Network and SVM using PSO technique using Rapid Miner software by comparing several classifications. The results of this study is to produce a model that can support decision making in terms of objectively determining the model employee.

Keyword : C4.5, neural network, SVM, PSO, Rapidminer, employee nonsuch

Pendahuluan

Berkompetisi adalah hal yang wajar terjadi didalam kehidupan karena manusia dilahirkan sebagai makhluk kompetitif. Sifat ini akan terus ada dalam naluri bahwa sadar manusia. Tanpa sadar, setiap manusia akan suka terhadap kompetisi dalam kehidupan sehari-harinya, termasuk dalam hal pekerjaan seperti menjadi Karyawan Teladan. Karyawan merupakan aset penting bagi setiap perusahaan, karena sangat mempengaruhi banyak aspek penentu keberhasilan kerja dari perusahaan tersebut. Suatu perusahaan akan dapat menjalankan semua proses usahanya dengan baik jika semua karyawannya dapat diorganisir dengan baik pula oleh bagian Sumber Daya Manusia. Namun untuk mengelola karyawan menjadi karyawan yang teladan tidaklah mudah karena terdapat beberapa kendala salah satunya adalah banyaknya karyawan yang dimiliki. Dengan jumlah karyawan yang banyak dalam perusahaan tersebut maka untuk mengelola perencanaan jenjang karir dari setiap karyawan dan kaderisasi atau pergantian jabatan dalam perusahaan menjadi sulit dan menghabiskan banyak waktu, karena manager bagian Sumber Daya Manusia belum tentu mengenal dengan dekat tiap karyawan yang ada pada divisi IT.

Pengambilan keputusan adalah proses menemukan korelasi baru bermakna, pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola serta statistik.

Berdasarkan latar belakang penelitian yang telah diuraikan, maka dapat diidentifikasi permasalahan penelitian sebagai berikut :

Saat ini kondisi pengambilan keputusan untuk karyawan teladan pada PT. XYZ masih dilakukan satu sisi dengan menggunakan perkiraan saja.

Penilaian dan pemilihan karyawan teladan hanya berdasarkan penilaian dari atasan langsung divisi IT.

Tujuan Penelitian ini adalah untuk Membangun klasifikasi dengan membandingkan tingkat akurasi yang dihasilkan oleh teknik atau model algoritma C4.5, *Neurel network* dan SVM dalam menentukan pemilihan karyawan teladan PT. XYZ, Untuk memperoleh informasi tentang karakteristik dan pola kelayakan dalam pemilihan karyawan teladan PT. XYZ, Memformulasikan pengambil keputusan dalam kelayakan proses menentukan karyawan teladan PT. XYZ dengan menggunakan GUI Java Netbeans 6.8.

Manfaat Penelitian ini adalah Manfaat dari penelitian ini diharapkan dapat memberikan kontribusi terhadap intitusi terkait serta memberikan manfaat praktis dan manfaat teoritis: Manfaat praktis yakni dari hasil penelitian ini diharapkan agar dapat digunakan oleh PT. XYZ dalam menentukan seseorang yang berhak menjadi karyawan teladan, Manfaat teoritis dari penelitian ini diharapkan dapat memberikan kontribusi bagi penelitian yang berkaitan dengan komparasi metode klasifikasi algoritma khususnya dalam menentukan karyawan teladan.

Tinjauan Pustaka

Pengambilan Keputusan adalah proses menemukan korelasi baru bermakna, pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola serta statistik. Pengambilan Keputusan adalah bidang interdisipliner yang mengacu pada ilmu komputer (data base, kecerdasan buatan, pembelajaran mesin, grafis dan visualisasi model), statistik dan teknik (pengenalan pola, jaringan saraf). melibatkan analisis besar basis data yang ada dalam rangka untuk menemukan pola dan hubungan dalam data, dan temuan lainnya (tak terduga, mengejutkan, dan berguna). Pengambilan Keputusan adalah gabungan ilmu statistik, kecerdasan buatan, dan penelitian database.

Algoritma C4.5

Pada akhir tahun 1970 sampai awal tahun 1980 J. Ross Quinlan, seorang peneliti di bidang *machine learning*, membuat sebuah algoritma *decision tree* yang dikenal dengan ID3 (*Iterative Dichotomiser*). Quinlan kemudian membuat algoritma C4.5 (sering disebut dengan pohon keputusan) yang merupakan pengembangan dari algoritma ID3. Algoritma ini memiliki kelebihan, yaitu mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar (pohon keputusan). Algoritma C4.5 merupakan algoritma klasifikasi dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, 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:

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

Keterangan:

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi S_i terhadap S

3. Kemudian hitung nilai *gain* menggunakan rumus:

$$Gain(S,A) = Entropy(S) -$$

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

(II.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.

Algoritma Neural Network

Neural network dikenal dengan nama lain yaitu Jaringan Syaraf Tiruan (JST), *Artificial Neural Network* (ANN), disebut juga *Simulated Neural Network* (SNN), atau biasanya hanya disebut *Neural Network*(NN). *Neural Network*(NN) adalah jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan susunan syaraf manusia. JST atau *neural network* merupakan sistem adaptif yang dapat merubah strukturnya untuk memecahkan masalah berdasarkan informasi eksternal maupun internal yang mengalir melalui jaringan tersebut. Secara sederhana, *neural network* adalah alat pemodelan untuk memodelkan hubungan yang kompleks antara *input* dan *output* untuk menemukan pola-pola pada data. Pengertian lain *Neural network* adalah satu set unit *input/output* yang terhubung dimana tiap relasinya memiliki bobot. Menurut Kohonen *neural network* dibuat pertama kali oleh Warren McCulloch dan Walter Pitts (1943) dan dianggap sebagai basis *neural network* modern saat ini. Pada tahun 1950, Rosenblatt dan peneliti lainnya membuat sebuah kelas *neural network* yang disebut *perceptron* yang merupakan model *neuron* biologi tapi hanya bisa menyelesaikan masalah *linearly separable*. Kemudian *perceptron* yang dihasilkan Minsky dan papert (1969) memecahkan masalah *nonlinearly separable* dengan membuat *multilayer network* dengan unit *hidden*, tapi *training* terhadap *network* tidak bisa dilakukan.

Langkah pembelajaran dalam algoritma *backpropagation* adalah sebagai berikut:

1. Inisialisasi bobot jaringan secara acak (biasanya antara -0.1 sampai 1.0)
2. Untuk setiap data pada data *training*, hitung input untuk simpul berdasarkan nilai input dan bobot jaringan saat itu, menggunakan rumus:

$$Input_j = \sum_{i=1}^n O_i w_{ij} + \Theta_j$$

Keterangan:

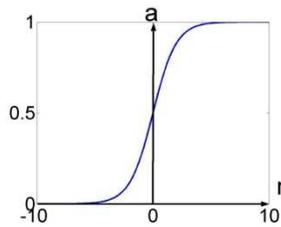
O_i = *Output* simpul i dari layer sebelumnya

w_{ij} = bobot relasi dari simpul i pada *layer* sebelumnya ke simpul j

Θ_j = bias (sebagai pembatas)

3. Berdasarkan input dari langkah dua, selanjutnya membangkitkan output untuk simpul menggunakan fungsi aktivasi sigmoid:

$$\text{Output} = \frac{1}{1+e^{-\text{Input}}}$$



Gambar 1. Fungsi Aktivasi Sigmoid

Algoritma SVM

Klasifikasi SVM juga dikenal dengan Naive Bayes, memiliki kemampuan sebanding dengan pohon keputusan dan neural network. Mudah diinterpretasikan sehingga pengguna yang tidak punya keahlian dalam bidang teknologi klasifikasi pun bisa mengerti. Klasifikasi Bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Kusrini, 2009). Klasifikasi Bayes didasarkan pada SVM, diambil dari nama seorang ahli matematika yang juga menteri *Prebysterian* Inggris, *Thomas Bayes* (1702-1761), yaitu:

$$P(x|y) = \frac{P(y|x) P(x)}{P(y)}$$

keterangan :

- y = data dengan kelas yang belum diketahui
- x = hipotesis data y merupakan suatu kelas spesifik
- $P(x | y)$ = probabilitas hipotesis x berdasar kondisi y (*posteriori probability*)
- $P(x)$ = probabilitas hipotesis x (*prior probability*)
- $P(y | x)$ = probabilitas y berdasarkan kondisi pada hipotesis x
- $P(y)$ = probabilitas dari y
- $P(x|y)$ = $P(y|x) P(x)$

Metodologi Penelitian

Dilihat dari jenis informasi yang dikelola maka jenis penelitian ini terbagi menjadi :

1. Penelitian Experimental
Penelitian experimental merupakan penelitian yang bersifat uji coba, mempengaruhi hal-hal yang terkait dengan seluruh variabel atau atribut.
2. Penelitian Perbandingan atau studi komparasi yakni dengan membandingkan *algoritma C4.5, SVM dan Neural Network* dengan teknik PSO dan kemudian diuji dengan menggunakan tool *Rapid Miner*.

Populasi dalam penelitian ini adalah seluruh karyawan Divisi IT PT. XYZ yang beralamatkan di Jl. Swadaya Gedung Graha Indika Lantai 2-6, Dewi Sartika Jakarta Timur. Dari seluruh jumlah populasi yang ada maka diambil 137 orang sebagai sampel yang merupakan karyawan Divisi IT PT. XYZ. Adapun Pemilihan sampel dilakukan secara apriori yaitu dengan melihat suatu fenomena atau fakta dalam keadaan yang sebenarnya. Pengumpulan data merupakan bagian paling penting dalam sebuah penelitian. Ketersediaan data akan sangat menentukan dalam proses pengolahan dan analisa selanjutnya. Karenanya, dalam pengumpulan data harus dilakukan teknik yang menjamin bahwa data diperoleh itu benar, akurat dan bisa dipertanggungjawabkan sehingga hasil pengolahan dan analisa data tidak bias. Pengumpulan data bersifat teoritis yang berhubungan dengan penelitian ini.

Pengambilan data tersebut dilakukan dengan cara mempelajari literatur-literatur, jurnal-jurnal penelitian, bahan kuliah dan sumber-sumber lain yang ada hubungannya dengan permasalahan yang dibahas. Pengumpulan data dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan data primer

Digunakan untuk mendapatkan data primer, yaitu dengan cara melakukan wawancara langsung dengan pihak PT. XYZ. Terlampir pada lampiran lembar wawancara.

2. Pengumpulan data sekunder

Data sekunder dikumpulkan dengan mengamati data, membaca, mempelajari dan mengutip dari buku literatur, serta sumber-sumber yang berhubungan erat dengan penelitian ini.

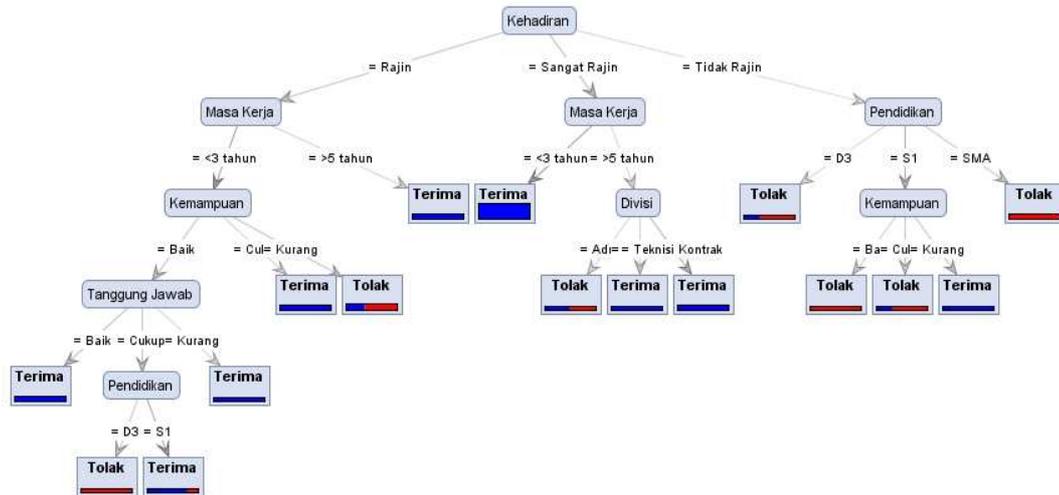
Hasil dan Pembahasan

Perhitungan Algoritma C4.5

Menurut Kusrini dalam bukunya *Algoritma Data Mining* , secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Pada langkah awal dari proses perhitungan algoritma C4.5 ini adalah menentukan terlebih dahulu atribut sebagai akarnya. Atribut yang terpilih adalah yang mempunyai nilai *gain* tertinggi dari nilai *gain-gain* atribut yang lainnya, kemudian dijadikan akar dari pohon. Dari nilai entropy dan gain yang diperoleh tabel 1 maka selanjutnya tentukan simpul berikutnya, yaitu simpul 1.1, dan dilakukan perhitungan entropy dan gain masing-masing atribut dari kehadiran. Jumlah kasus yang dihitung adalah nilai dari simpul kehadiran dan seterusnya sampai semua record dalam simpul, mendapatkan kelas yang sama. Berikut ini adalah hasil uji dengan tools Rapid Miner terhadap kehadiran sehingga langsung ditentukan atribut prediktor merupakan penentu dari seluruh atribut lainnya.



Gambar 2. Hasil menggunakan Rapid Miner

Perhitungan SVM

Perhitungan pemilihan mitra dengan nilai remark diterima dan ditolak terlihat pada Tabel 2 baris pertama. Baris-baris berikutnya adalah hasil perhitungan nilai probabilitas prior, yaitu probabilitas nilai diterima dan ditolak masing-masing atribut terhadap total diterima dan ditolak dari seluruh data. Dalam data training terdapat 137 record dengan 110 kasus diterima

dan 27 kasus ditolak, untuk menentukan prior probability dengan menggunakan rumus (II.3) dan (II.4) :

$$P(\text{yes},n) = 110/137 = 0.8029$$

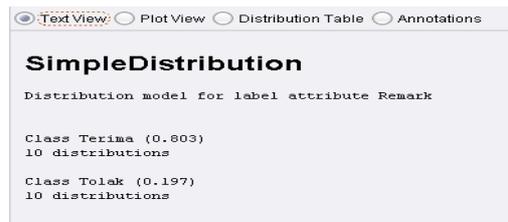
$$P(\text{no},n) = 27/137 = 0.1971$$

Jika terdapat kasus baru yang menuntut untuk mendapatkan hasil pemilihan pemilihan karyawan teladan akan digunakan tabel IV-II untuk menentukan pemilihan karyawan teladan termasuk dalam kelas yang diterima atau tidak diterima (Ditolak) maka dilakukan perhitungan probabilitas posterior yang telah dihitung sebelumnya. Berikut ini probabilitas posterior pada tabel dibawah ini:

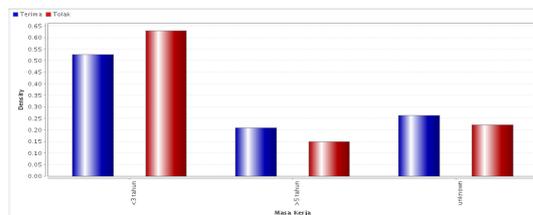
Dari tabel tersebut terdapat beberapa langkah untuk menghitung, yaitu:

- Diketahui $P(X|Ci) = P(\text{kasus}|\text{remark} = \text{diterima}) = 1.234480\text{E-}07$
- Diketahui $P(X|Ci) = P(\text{kasus}|\text{remark} = \text{ditolak}) = 8.35\text{E-}13$
- Lalu hitung $P(X|Ci)P(Ci)$ katagori *diterima* dengan menggunakan rumus $P(X|Ci)P(Ci) = 1.234480\text{E-}07 * 0.029197 = 9.911880\text{E-}08$
- Lalu hitung $P(X|Ci)P(Ci)$ katagori *ditolak* dengan menggunakan rumus $P(X|Ci)P(Ci) = 8.35458\text{E-}13 * 0.1971 = 1.64652\text{E-}13$

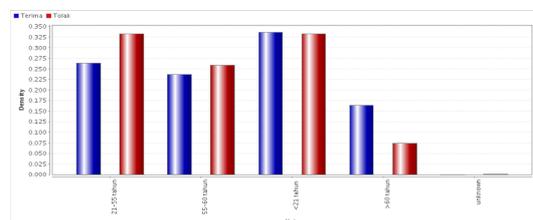
Dari hasil perhitungan diatas, diperoleh nilai $P(X|Ci)$ dan $P(X|Ci) P(Ci)$ lebih besar pada remark *diterima*, sehingga dapat disimpulkan bahwa data baru termasuk dalam klasifikasi *diterima*.



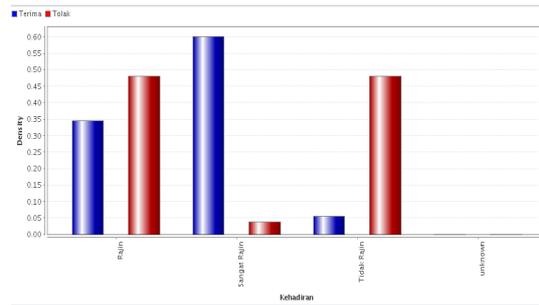
Gambar 3. Remark pemilihan karyawan teladan svm dengan Rapid Miner



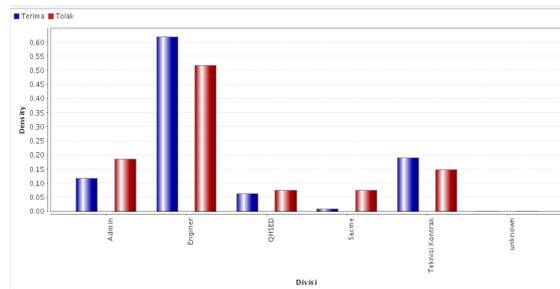
Gambar 4. Plot prediksi masa kerja pemilihan karyawan teladan



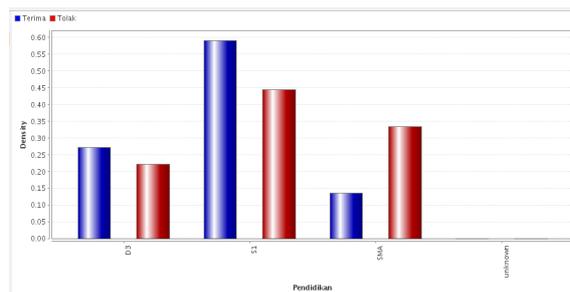
Gambar 5. Plot prediksi usia pemilihan karyawan teladan



Gambar 6. Plot prediksi kehadiran pemilihan karyawan teladan



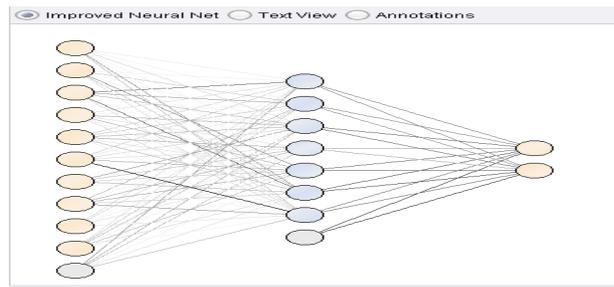
Gambar 7. Plot prediksi divisi pemilihan karyawan teladan



Gambar 8. Plot prediksi pendidikan pemilihan karyawan teladan

Perhitungan Neural Network

Dalam neural network yang menggunakan algoritma back propagation pada 6 (enam) langkah pembelajaran yaitu dengan menghitung atau menginisialisasi nilai bobot awal antara -0.1 sampai dengan 1.0 untuk *input layer*, *hidden layer* dan bias atau *threshold*. Pada simpul bias terdiri dari dua, yaitu: simpul bias pada *input layer* yang terhubung dengan simpul-simpul *hidden layer* dan simpul bias pada *hidden layer* yang menghubungkan pada *output layer*. Setelah hitung input untuk simpul berdasarkan nilai input dan bobot jaringan saat itu, lalu bangkitkan output untuk simpul menggunakan fungsi aktivasi sigmoid. Kemudian tentukan nilai error baru yang pada akhirnya nilai error tersebut digunakan kembali untuk memperbaharui bobot relasi berikutnya. Berikut ini adalah *neural net* yang dihasilkan dari data training dengan menggunakan *multilayerperceptron* pada tools Rapid Miner.



Gambar 9. Neural Net yang Dihasilkan

Setiap data pada data *training*, dihitung input untuk simpul berdasarkan nilai input dan jaringan saat itu. Bobot awal untuk *input layer*, *hidden layer*, dan bias diinisialisasi secara acak (biasanya antara -0.1 sampai dengan 1.0). Simpul bias terdiri dari dua, yaitu pada *input layer* yang terhubung dengan simpul-simpul pada *hidden layer*, dan pada *hidden layer* yang terhubung pada *output layer*. Setelah semua nilai awal diinisialisasi, kemudian dihitung masukan, keluaran, dan *error*. Selanjutnya membangkitkan output untuk simpul menggunakan fungsi aktivasi sigmoid. Setelah didapat nilai dari fungsi aktivasi, hitung nilai *error* antara nilai yang diprediksi dengan nilai yang sesungguhnya. Setelah nilai *error* dihitung, selanjutnya dibalik ke *layer* sebelumnya (*backpropagated*). Nilai *Error* yang dihasilkan dari langkah sebelumnya digunakan untuk memperbarui bobot relasi.

Kurva ROC

Hasil perhitungan divisualisasikan dengan kurva ROC (*Receiver Operating Characteristic*) atau AUC (*Area Under Curve*). ROC memiliki tingkat nilai diagnosa yaitu:

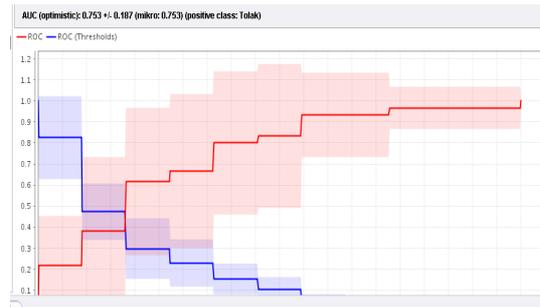
- a. Akurasi bernilai 0.90 – 1.00 = *excellent classification*
- b. Akurasi bernilai 0.80 – 0.90 = *good classification*
- c. Akurasi bernilai 0.70 – 0.80 = *fair classification*
- d. Akurasi bernilai 0.60 – 0.70 = *poor classification*
- e. Akurasi bernilai 0.50 – 0.60 = *failure*

Hasil yang didapat dari pengolahan ROC untuk algoritma C4.5 dengan menggunakan data *training* sebesar 0.979 dapat dilihat pada gambar 10 dengan tingkat diagnosa *excellent classification*.



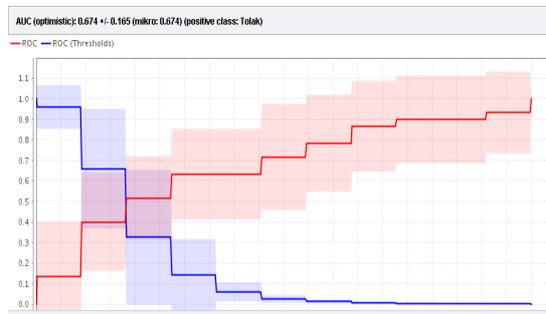
Gambar 10. Kurva ROC untuk Algoritma C4.5

Hasil yang didapat dari pengolahan ROC untuk algoritma *Svm* dengan menggunakan data *training* sebesar 0.753 dapat dilihat pada gambar 11 dengan tingkat diagnosa *fair classification*.



Gambar 11. Kurva ROC Data Algoritma Svm

Hasil yang didapat dari pengolahan ROC untuk algoritma *Neural Network* dengan menggunakan data *testing* sebesar 0.674 dapat dilihat pada gambar 12 dengan tingkat diagnosa *poor classification*.



Gambar 12. Kurva ROC data algoritma Neural Network

Perbandingan hasil perhitungan nilai AUC untuk metode C4.5, *Neural Network*, dan *Svm* dapat dilihat pada Tabel 1.

Tabel 1. Perbandingan Nilai AUC

AUC	C4.5	Svm	Neural Network
	0.979	0.753	0.674

Analisis Evaluasi Komparasi Model

Berdasarkan pengujian dan evaluasi hasil klasifikasi dengan algoritma C4.5, *Neural Network*, dan *Svm*, dapat kita lihat hasilnya yaitu tingkat akurasi pada data *training* yang paling tinggi dengan algoritma C4.5 dengan tingkat akurasi 87.58. Berdasarkan kolom ROC pada tabel IV-6, data *training* algoritma C4.5 memiliki tingkat ROC paling tinggi, yaitu 0.979 termasuk dalam katagori *excellent classification*. Dengan menggunakan perbandingan data *training*, maka perbandingan komparasinya untuk akurasi dapat dilihat dalam tabel 2:

Tabel 2. Perbandingan Akurasi

	C4.5	Svm	Neural network
Accuracy	87.58%	84.67%	77.36%
ROC	0.979	0.753	0.674

Berdasarkan hasil perbandingan akurasi pada tabel 2 Algoritma C4.5 yang memiliki tingkat akurasi yang paling tinggi, sehingga baik digunakan untuk klasifikasi pemilihan karyawan teladan dengan persentase 87,58%.

Penerapan Algoritma Terpilih

Berdasarkan hasil perbandingan akurasi pada tabel 2 algoritma terpilih sebagai algoritma terbaik dalam klasifikasi pemilihan karyawan teladan yaitu algoritma C4.5 yang memiliki tingkat akurasi tertinggi dengan persentase 87,58% dilakukan penerapan pada data baru.

Masa Kerja	Usia	Kehadiran	Divisi	Pendidikan	Kinerja	Kemampuan	Tanggung Jawab	Status Perkawinan	Status Karyawan	Remark	
>5 tahun	<21 tahun	angakat	Rajiji	Admin	S1	Baik	Baik	Baik	lum menikah	Kontrak	Terima
<3 tahun	!1-55 tahun	angakat	Rajiji	Admin	D3	Cukup	Baik	Baik	Menikah	Tetap	Terima
>5 tahun	!1-55 tahun		Rajin	Admin	SMA	Cukup	Cukup	Cukup	Menikah	Tetap	Tolak
3-5 tahun	<21 tahun		Rajin	Engineer	S1	Baik	Baik	Baik	lum menikah	Kontrak	Terima
<3 tahun	<21 tahun		Rajin	Engineer	S1	Kurang	Kurang	Kurang	lum menikah	Kontrak	Tolak
<3 tahun	!1-55 tahun		Rajin	Engineer	D3	Baik	Baik	Baik	Menikah	Tetap	Terima
>5 tahun	i5-60 tahun	angakat	Rajiknisi	Kontr	S1	Kurang	Kurang	Cukup	lum menikah	Tetap	Tolak
3-5 tahun	i5-60 tahun	angakat	Rajiknisi	Kontr	D3	Baik	Cukup	Baik	Menikah	Tetap	Terima
<3 tahun	!1-55 tahun	angakat	Rajiji	Sacme	SMA	Baik	Baik	Baik	Menikah	Tetap	Terima
>5 tahun	>60 tahun	Tidak	Rajin	QHSED	S1	Kurang	Kurang	Kurang	lum menikah	Tetap	Tolak

Gambar 13. Data Baru untuk Penerapan Algoritma Terpilih

Hasil penerapan *rule* algoritma C4.5 terhadap data baru sejumlah 10 record data dimana 6 data diprediksi diterima sesuai dengan hasil klasifikasi dan 1 data diprediksi diterima tetapi ternyata ditolak. Sedangkan prediksi ditolak sejumlah 3 data diprediksi diterima pada kelas ditolak. Hasil dapat dilihat pada gambar 15.

accuracy: 90.00% +/- 30.00% (mikro: 90.00%)			
	true Terima	true Tolak	class precision
pred. Terima	6	1	85.71%
pred. Tolak	0	3	100.00%
class recall	100.00%	75.00%	

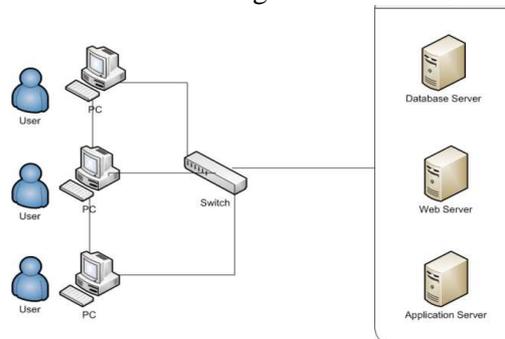
Gambar 14. Confussion Matrix Data Baru dengan Algoritma C4.5

Rule hasil klasifikasi dari algoritma C4.5 diterapkan kedalam pembuatan aplikasi untuk klasifikasi pemilihan karyawan teladan dengan menggunakan Java seperti pada gambar 15.



Gambar 15. Aplikasi untuk Klasifikasi Pemilihan Karyawan Teladan

Pada aplikasi untuk klasifikasi mahasiswa pada gambar 15 dihasilkan klasifikasi diterima dan ditolak Input data karyawan pada program tersebut sesuai dengan atribut yang dibutuhkan, kemudian klik tombol HASIL, maka secara otomatis tampil hasil klasifikasi karyawan diterima atau ditolak. Untuk menginput kembali data baru klik tombol INPUT. Berikut ini desain arsitektur sistem data mining:



Gambar 16. Arsitektur Sistem Data Mining

Simpulan

Dari pengukuran kinerja ketiga algoritma yang telah dilakukan dengan menggunakan teknik PSO dan berdasarkan jumlah data maka dapat disimpulkan bahwa algoritma *C4.5* memiliki kemampuan dalam pengambilan keputusan untuk menentukan Karyawan Teladan yang diterima dan yang ditolak. Maka hasil penelitian dari percobaan yang telah dilakukan dapat disimpulkan bahwa:

- Ketiga algoritma (*C4.5*, *Neural Network*, dan *SVM*) dapat digunakan dalam memutuskan Karyawan Teladan di PT. XYZ. Ketiga algoritma ini dikomparasi kemudian diuji akurasi. Tingkat akurasi tertinggi lah yang digunakan dalam menentukan Karyawan Teladan.
- Algoritma *C4.5* memiliki tingkat akurasi tertinggi diantara ketiga algoritma tersebut. Algoritma inilah yang di implementasikan dalam menentukan Karyawan Teladan.

Dari hasil penelitian yang dibuat sangat besar harapan agar dapat berguna bagi setiap perusahaan yang ingin menerapkan sistem untuk mendapatkan karyawan yang berkualitas dalam kenaikan jabatan. Harapan agar sistem ini dapat menjadi bahan acuan untuk melanjutkan proses penelitian lanjutan bagi penulis.

Saran

Agar penelitian ini bisa ditingkatkan, berikut adalah saran-saran yang diusulkan:

- Penelitian ini diharapkan dapat digunakan pihak Perusahaan sebagai bahan pertimbangan dalam menentukan pemilihan karyawan teladan, sehingga dapat meningkatkan akurasi ketika melakukan pemilihan karyawan teladan pada PT. XYZ.
- Penelitian ini dapat dikembangkan dengan metode optimasi lainnya seperti Ant Colony Optimization (ACO), Genetic Algorithm (GA), dan lainnya.
- Penelitian ini dapat dikembangkan dengan metode klasifikasi data mining lainnya seperti Naive Bayes, KNN dan lainnya untuk melakukan perbandingan
- Implementasi pemilihan karyawan teladan di PT. XYZ memberikan dampak positif terhadap perusahaan. Oleh karena itu data-data yang banyak dapat di olah kemudian di carikan polanya sehingga menemukan pengetahuan baru agar data-data yang banyak tidak terbuang sia-sia. Sehingga data-data tersebut dapat dibuat berubah menjadi sebuah pengetahuan baru melalui data mining.

Kedepannya aplikasi data mining dapat lebih dikembangkan dengan menyesuaikan kasus yang ada serta penyesuaian dengan bahasa pemrograman yang dipilih harus lebih akurat dengan sistematika proses algoritma data mining yang dipilih. Perangkat keras dan perangkat

lunak pendukung pun harus disesuaikan seiring berkembangnya teknologi, serta kaderisasi *operator* dan pembelajaran harus dilaksanakan agar eksistensi sistem data mining khususnya dalam menentukan karyawan teladan dapat terus berjalan dan mampu memberikan solusi pemecahan masalah yang berkaitan dengan karyawan teladan karyawan.

Daftar Pustaka

- Feri S, Dominikus J (2010). *Data Mining Meramalkan Bisnis Perusahaan*. Elex Media Komputindo. Jakarta.
- Hallyana (2012). *Penerapan Algoritma C45 Dalam Mendeteksi Perilaku Nasabah Mikro Kredit Usaha Menggunakan Aplikasi Rapid Miner Studi Kasus PT. Bank Mandiri, Tbk (Persero)*. Tesis., Universitas Budi Luhur.
- Harry, Dhika (2009). *Kajian Komparasi Penerapan Algoritma C4.5, Naïve Bayes, Dan Neural Network Dalam Pemilihan Mitra Kerja Penyedia Jasa Transportasi: Studi Kasus Cv. Kusriani*, E.T. Luthfi, Algoritma Data Mining. Andi Offset, Yogyakarta.
- Larose, Daniel T (2005). *Discovering Knowledge in Data: An Introduction to Data mining*, Jhon Willey & Son Inc., New Jersey.
- Sunjana (ISSN: 1907-5022, juni, 2010):1-6. *Aplikasi Mining Data Mahasiswa Dengan Metode Klasifikasi Decision Tree*. Seminar Nasional Aplikasi Teknologi Informasi.
- Triyani (2011). *Model pengambilan keputusan dalam promosi jabatan karyawan berdasarkan metode analytic network process (anp):Studi kasus PT. Shields security solution*. Tesis., STMIK Nusa Mandiri.
- Turban, E., Aronson, J.E, &Liang, T (2005), *Decision Support Sistem and Intelligent Sistem (Sistem Pendukung Keputusan dan Sistem Cerdas)*, Edisi ketujuh, Andi, Yogyakarta.
- Raharjo., Budi., Imam, H., Arif, H (2007). *Mudah Belajar Java*. Informatika. Bandung, Indonesia.
- Raharjo, B. (2011). *Membuat database menggunakan Mysql* . Informatika Bandung, Bandung.
- Supardi, I. Y. (2007). *Pemrograman Database dengan Java dan MySQL*. Jakarta: PT. Elex Media Komputindo.