



Kombinasi *Analytical Hierarchy Process*, C4.5, dan *Particle Swarm Optimization* pada Klasifikasi Pegawai

Dafiz Adi Nugroho^{a,*}, Catur Edi Widodo^b, Rahmat Gernowo^c

^a Magister Sistem Informasi, Universitas Diponegoro, Semarang, Indonesia

^{b,c} Departemen Fisika, Universitas Diponegoro, Semarang, Indonesia

Naskah Diterima : 22 April 2022; Diterima Publikasi : 18 Oktober 2022

DOI: 10.21456/vol12iss2pp81-88

Abstract

Decision Tree C4.5 is widely implemented in various research fields in determining classification, but there are still weaknesses in Decision Tree C4.5, one of which is that it cannot rank each alternative. In this study, to overcome the weakness of Decision Tree C4.5, a combination of Analytical Hierarchy Process (AHP) methods, Decision Tree C4.5, and Particle Swarm Optimization (PSO) methods is proposed in the case study of employee classification for promotion recommendations. The research begins by determining the criteria and weighting criteria from the interview results which are then processed with AHP to produce employee ratings and eligibility labels for the classification process. The classification process uses the Decision Tree C4.5 method which is optimized with the PSO algorithm so as to produce employee eligibility data for promotions. The results of the combined research of AHP, Decision Tree C4.5, and PSO methods show that AHP can produce employee ratings based on performance and potential criteria, and Decision Tree C4.5 classification and optimization with PSO have better accuracy results, namely 95.80% compared to Decision Tree C4.5 method without PSO optimization is 93.40%. Based on the results of the ranking and classification of this research can be used as a basis for promotion of employees.

Keywords : Data Mining; Classification; AHP; C4.5; PSO.

Abstrak

Decision Tree C4.5 banyak diimplementasikan dalam berbagai bidang penelitian dalam menentukan klasifikasi, tetapi masih terdapat kelemahan pada *Decision Tree C4.5*, salah satunya tidak dapat melakukan pemeringkatan pada setiap alternatif. Pada penelitian ini, untuk mengatasi kelemahan dari *Decision Tree C4.5* tersebut, diusulkan kombinasi metode *Analytical Hierarchy Process* (AHP), *Decision Tree C4.5*, dan *Particle Swarm Optimization* (PSO) pada studi kasus klasifikasi pegawai untuk rekomendasi promosi jabatan. Penelitian dimulai dengan menentukan kriteria dan bobot kriteria dari hasil wawancara yang kemudian diproses dengan AHP untuk menghasilkan peringkat pegawai dan label kelayakan untuk proses klasifikasi. Proses klasifikasi menggunakan metode *Decision Tree C4.5* yang dioptimasi dengan algoritma PSO sehingga menghasilkan data kelayakan pegawai untuk promosi jabatan. Hasil dari penelitian kombinasi metode AHP, *Decision Tree C4.5*, dan PSO menunjukkan bahwa AHP dapat menghasilkan peringkat pegawai berdasarkan kriteria kinerja dan potensi, dan klasifikasi *Decision Tree C4.5* dan optimasi dengan PSO mempunyai hasil akurasi yang lebih baik yaitu 95.80 % dibandingkan dengan metode *Decision Tree C4.5* tanpa optimasi PSO yaitu 93.40 %. Berdasarkan hasil pemeringkatan dan klasifikasi dari penelitian ini dapat digunakan sebagai dasar promosi jabatan pegawai.

Keywords : Penambangan Data; Klasifikasi; AHP; C4.5; PSO.

1. Pendahuluan

Data mining merupakan proses untuk menemukan informasi yang signifikan atau hilang yang tersimpan dalam jumlah data yang mempunyai skala besar (Saha *et al.*, 2019), *data mining* merupakan aktivitas dalam eksplorasi data untuk mendapatkan pengetahuan yang bermanfaat (Lee dan Kim, 2009) dengan menganalisis pola data pada basis data.

Klasifikasi merupakan metode *data mining* atau penambangan data untuk memprediksi kelas atau kategori dari sekumpulan atribut data, atau merupakan

metode *supervised learning* yang memproses dari sekumpulan data dan menghasilkan model klasifikasi yang digunakan untuk mengklasifikasikan *instance* baru (Malik dan Haouassi, 2021). Klasifikasi merupakan metode yang menggunakan teknik *machine-learning* dengan melakukan pembelajaran terhadap pola-pola informasi atau data yang telah mempunyai label sehingga dapat menempatkan informasi atau data baru dengan label yang belum diketahui sebelumnya ke dalam kelas atau kelompok masing-masing.

*) Penulis korespondensi: dafizadi77@gmail.com

Salah satu metode klasifikasi yang sering digunakan dalam membuat klasifikasi pada data adalah *Decision Tree C4.5*. *C4.5* merupakan proses klasifikasi memilih atribut dari dataset sebagai *root node* dan membuat cabang pada setiap nilai atribut yang digunakan dalam membagi kasus pada setiap cabang, kemudian melakukan proses yang sama pada setiap cabang sampai didapatkan kelas yang sama untuk semua kasus pada cabang (Ariyati *et al.*, 2020). Klasifikasi menggunakan *Decision Tree C4.5* telah banyak digunakan di berbagai penelitian. Salah satu penerapan pada bidang pendidikan yaitu proses evaluasi beasiswa pada mahasiswa perguruan tinggi (Wang *et al.*, 2019). Pada bidang energi yaitu melakukan klasifikasi pada penilaian stabilitas tegangan *online* (Meng *et al.*, 2020), tetapi masih terdapat kelemahan pada *Decision Tree C4.5* yaitu tidak dapat melakukan pemeringkatan pada setiap alternatif (Gustian dan Hundayani, 2018), dan perlunya optimasi pada algoritma untuk meningkatkan akurasi pada klasifikasi data.

AHP dapat digunakan untuk perankingan atau mencari alternatif terbaik dan menjadi solusi dari permasalahan seleksi. Salah satu penerapan AHP yaitu dalam bidang manajemen sumber daya manusia yaitu digunakan dalam deskripsi dan analisis kriteria yang digunakan dalam penilaian pegawai (Ayuningtyas *et al.*, 2017). Tetapi dalam penerapan metode AHP mempunyai kelemahan yaitu pada kasus dengan jumlah kriteria dan alternatif yang banyak, AHP tidak bisa berjalan secara efektif dan memerlukan waktu yang banyak dalam prosesnya (Dewi *et al.*, 2018), dan metode AHP tidak dapat melakukan klasifikasi pada setiap alternatif yang dinilai.

Dari kekurangan yang dimiliki oleh AHP dan *Decision Tree C4.5*, perlu adanya kombinasi metode AHP dengan *Decision Tree C4.5* untuk mengatasi kekurangan kedua metode tersebut, yaitu dapat melakukan pemeringkatan dan pembobotan pada kriteria dan alternatif pilihan yang dilakukan oleh metode AHP, dan melakukan klasifikasi yang dilakukan oleh *Decision Tree C4.5*, seperti pada penelitian kombinasi AHP dan *Decision Tree C4.5* pada klasifikasi siswa *drop out* (Gustian dan Hundayani, 2018). Pada klasifikasi *Decision Tree C4.5* dapat ditambahkan metode optimasi untuk meningkatkan nilai akurasi klasifikasi.

Optimasi yang sudah digunakan untuk meningkatkan akurasi klasifikasi *Decision Tree C4.5* salah satunya yaitu PSO. PSO merupakan algoritma pencarian berbasis populasi yang pada prosesnya dilakukan inisialisasi populasi solusi secara acak untuk dapat menyelesaikan masalah optimasi (Grosan *et al.*, 2006). Penggunaan optimasi PSO pada *Decision Tree C4.5* sudah pernah dilakukan dalam penelitian klasifikasi penggunaan pembayaran digital yang dapat meningkatkan nilai akurasi pada klasifikasi (Ariyati *et al.*, 2020).

Pada penelitian ini, kombinasi metode AHP, *Decision Tree C4.5*, dan PSO akan diimplementasikan pada studi kasus klasifikasi pegawai di Pemerintahan Provinsi Jawa Tengah dengan menggunakan data kriteria-kriteria yang digunakan meliputi nilai Sasaran Kinerja Pegawai (SKP) Tahunan, nilai Penilaian Sikap Kinerja dan Perilaku (PKSP), dan presensi pegawai sebagai komponen dari nilai kinerja pegawai, serta nilai Psikometri pegawai dan tingkat pendidikan sebagai nilai potensi pegawai. Sehingga diharapkan dapat melakukan klasifikasi pegawai dan peringkat pegawai yang digunakan untuk rekomendasi dalam promosi jabatan, dan memiliki akurasi yang lebih baik dengan melakukan optimasi.

2. Kerangka Teori

2.1. Analytical Hierarchy Process (AHP)

AHP adalah metode pendukung keputusan yang dikembangkan oleh Thomas L. Saaty yang merupakan seorang profesor matematika University of Pittsburgh kelahiran Irak. Metode AHP digunakan dalam menyelesaikan suatu masalah *Multi Attribute Decision Making* (MADM). AHP menggunakan beberapa kriteria dan alternatif dengan proses analisis bertingkat (Xiulin dan Dawei, 2014), atau suatu metode pengambilan keputusan dengan melakukan penilaian terhadap kriteria-kriteria terkait dalam suatu urutan hierarki (Saaty, 2008) (Sedghiyan *et al.*, 2021).

Pada prinsipnya proses AHP terdapat 3 langkah. Pertama adalah dekomposisi yaitu membagi masalah kedalam tujuan, kriteria, dan alternatif. Kriteria yang digunakan pada penelitian ini diperoleh dari data kepegawaian Badan Kepegawaian Provinsi Jawa Tengah seperti pada tabel 1.

Tabel 1. Sampel Data Kriteria

ID Pegawai	PKSP	SKP	Presensi (hari)	Tingkat Pendidikan	Psikometri
1	90.5	87.82	20	S1	56
2	84.86	94.31	17	SMA	59
3	81.67	87.82	5	S1	70

Kedua, membuat matriks perbandingan berpasangan dari kriteria-kriteria yang digunakan menggunakan skala prioritas saaty dalam menerjemahkan pendapat yang mempunyai nilai skala 1 sampai dengan 9 (Saaty, 2008), yang dapat dilihat pada tabel 2.

Tabel 2. Skala Prioritas Saaty (Saaty, 2008)

Skala	Keterangan
1	Sama pentingnya
3	Sedikit lebih penting daripada yang lainnya
5	Lebih penting daripada yang lainnya
7	Jelas lebih penting daripada yang lainnya
9	Mutlak lebih penting daripada yang lainnya
2, 4, 6, 8	Nilai-nilai antara dua pertimbangan yang berdekatan

Skala pada setiap kriteria pada penelitian ini diperoleh dengan melakukan wawancara dengan pejabat yang mengelola jabatan struktural di Badan Kepegawaian Provinsi Jawa Tengah.

Ketiga, hasil dari matriks perbandingan berpasangan dilakukan pengolahan untuk menentukan alternatif terbaik, dimulai dengan menghitung bobot prioritas, kemudian mengukur konsistensi dengan menghitung *consistency index* atau indeks konsistensi dengan rumus sebagai berikut.

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{1}$$

Keterangan :

- CI : Indeks Konsistensi
- λ_{max} : Jumlah nilai *eigen*
- n : Jumlah kriteria

Selanjutnya menghitung nilai *Consistency Ratio*, dengan rumus sebagai berikut.

$$CR = \frac{CI}{IR} \tag{2}$$

Keterangan :

- CR : *Consistency Ratio*
- IR : Indeks *Random Consistency*

Dengan nilai indeks *Random Consistency* tergantung dengan besar ukuran matriks yang digunakan, seperti yang dijabarkan tabel 3 berikut :

Tabel 3. Indeks *Random Consistency*

Ukuran Matriks	Nilai IR
1 dan 2	0
3	0.58
4	0.90
5	1.12
6	1.24
7	1.32
8	1.41
9	1.45
10	1.49

Terakhir pengambilan keputusan berdasarkan perbandingan hasil penghitungan bobot alternatif.

2.2. Decision Tree C4.5

Algoritma *Decision Tree* C4.5 seperti dengan namanya merupakan metode klasifikasi pohon keputusan. *Decision Tree* C4.5 merupakan proses klasifikasi memilih atribut dari dataset sebagai *root node* dan membuat cabang pada setiap nilai atribut yang digunakan dalam membagi kasus pada setiap cabang, kemudian melakukan proses yang sama pada setiap cabang sampai didapatkan kelas yang sama untuk semua kasus pada cabang (Ariyati *et al.*, 2020). Algoritma ini mempunyai masukan terdiri dari data latih dan sampel. Data latih digunakan untuk membuat pohon keputusan yang telah teruji kebenarannya.

Sedangkan sampel merupakan data yang dijadikan parameter pada klasifikasi data (Tarigan *et al.*, 2019).

Algoritma C4.5 secara umum mempunyai tahapan proses dalam membuat pohon keputusan yaitu sebagai berikut (Ariyati *et al.*, 2020):

- 1 Salah satu atribut dipilih sebagai *root*.
- 2 Membuat cabang untuk setiap nilai.
- 3 Menentukan kelas kasus pada cabang.
- 4 Proses diulang pada setiap cabang sampai semua kasus di cabang memiliki kelas yang sama.

Nilai *gain* tertinggi dari atribut merupakan dasar yang digunakan dalam memilih atribut sebagai *root*. Sebelum melakukan penghitungan nilai *gain* pada atribut, terlebih dahulu dilakukan penghitungan nilai entropi, yaitu dengan rumus sebagai berikut :

$$H(S) = - \sum_{i=1}^n p_i \log_2 p_i \tag{3}$$

Keterangan :

- H : Entropi
- S : kumpulan (set data)
- n : jumlah *record*
- p_i : probabilitas dari ya atau tidak total kasus

Berikut merupakan rumus yang digunakan untuk menghitung *gain* :

$$IG(A) = H(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} H(S_i) \tag{4}$$

Keterangan :

- IG : *Gain*
- H : Entropi
- S : kumpulan (set data)
- A : atribut yang akan digunakan
- n : jumlah atribut partisi A
- $|S_i|$: jumlah kasus pada partisi i
- $|S|$: jumlah kasus di S

2.3. Particle Swarm Optimization (PSO)

PSO merupakan algoritma pencarian berbasis populasi dengan populasi solusi diinisialisasi secara acak serta digunakan untuk menyelesaikan masalah optimasi (Grosan *et al.*, 2006). Dr. Kennedy dan Eberhart memperkenalkan PSO pada tahun 1995 yang merupakan suatu metode optimasi heuristik global berdasarkan penelitian kawanan burung dan ikan (Haupt dan Haupt, 2003), untuk mendapatkan posisi yang tepat untuk mencapai tujuan tertentu (Haupt dan Haupt, 2003). Pada penerapan PSO, setiap partikel memiliki posisi dan bergerak dengan kecepatan yang diperbarui pada setiap iterasi. Setiap partikel dalam suatu populasi memiliki nilai *fitness* yang dihitung dari fungsi *fitness*. Ciri-ciri utama partikel dalam PSO adalah posisi, kecepatan, dan kemampuan bertukar informasi dengan partikel tetangganya, kemampuan mengingat posisi sebelumnya, dan kemampuan menggunakan informasi untuk mengambil keputusan (Chen *et al.*, 2014).

Setiap partikel dalam PSO juga diinterpretasikan oleh kecepatan partikel yang bergerak dengan kecepatan yang secara dinamis disesuaikan dengan perilaku historisnya melalui ruang pencarian. Sehingga, partikel mempunyai kecenderungan bergerak menuju area pencarian yang lebih baik selama proses pencarian (Grosan *et al.*, 2006). Berikut merupakan rumus yang digunakan untuk menghitung kecepatan partikel dan perpindahan posisi :

$$V_i^t = \omega V_i^{t-1} + c_1 r_1 (P_{best\ i} - X_i^t) + c_2 r_2 (G_{best} - X_i^t) \quad (5)$$

$$X_i^t = X_i^{t-1} + V_i^t \quad (6)$$

Keterangan:

V_i^t : kecepatan / *velocity* partikel i pada iterasi t

X_i^t : posisi partikel i pada iterasi t

ω : bobot inersia

c_1, c_2 : koefisien akselerasi

r_1, r_2 : angka acak antara 0 sampai 1

$P_{best\ i}$: posisi terbaik untuk partisan i

G_{best} : posisi terbaik global

Pada penelitian ini, PSO digunakan untuk mendapatkan *feature subset* yang digunakan oleh algoritma C4.5 untuk mendapatkan hasil kinerja klasifikasi terbaik.

3. Metode

Pada penelitian ini terdapat beberapa tahapan yaitu sebagai berikut, tahap pertama adalah analisis masalah yaitu dengan melakukan analisis masalah dalam studi kasus klasifikasi pegawai di Pemerintahan Provinsi Jawa Tengah yang menghasilkan rumusan masalah yang diperlukan dalam klasifikasi pegawai.

Tahap kedua dilakukan studi literatur terhadap metode klasifikasi yang akan digunakan pada studi kasus, hasil dari tahapan ini adalah penggunaan metode kombinasi metode AHP, *Decision Tree* C4.5, dan PSO yang memungkinkan adanya peningkatan kinerja atau akurasi pada klasifikasi dengan dilakukan optimasi, dan pemeringkatan alternatif pegawai.

Tahap ketiga dilakukan pengumpulan data, data kriteria-kriteria yang dikumpulkan meliputi nilai Sasaran Kinerja Pegawai (SKP) Tahunan, nilai Penilaian Sikap Kinerja dan Perilaku (PKSP), dan presensi pegawai sebagai komponen dari nilai kinerja pegawai, serta nilai Psikometri pegawai dan tingkat pendidikan sebagai nilai potensi pegawai yang didapat dari Badan Kepegawaian Daerah Provinsi Jawa Tengah.

Tahap keempat dilakukan data *preprocessing* pada data, dengan melakukan pembersihan pada data tidak konsisten serta data tidak lengkap maupun kosong, serta dilakukan proses transformasi data seperti tabel 4.

Tabel 4. Transformasi Data Presensi

Presensi	
Absen	Nilai Presensi
0	100
$0 < \text{Absen} < 3$	90
$3 \leq \text{Absen} < 4$	80
$4 \leq \text{Absen} < 7$	70
$7 \leq \text{Absen} < 11$	60
$11 \leq \text{Absen} < 21$	50
$21 \leq \text{Absen}$	25

Transformasi data presensi dilakukan dengan mengubah jumlah absen pegawai dalam satu tahun menjadi nilai presensi seperti pada tabel 4 diatas.

Tabel 5. Transformasi Data Tingkat Pendidikan

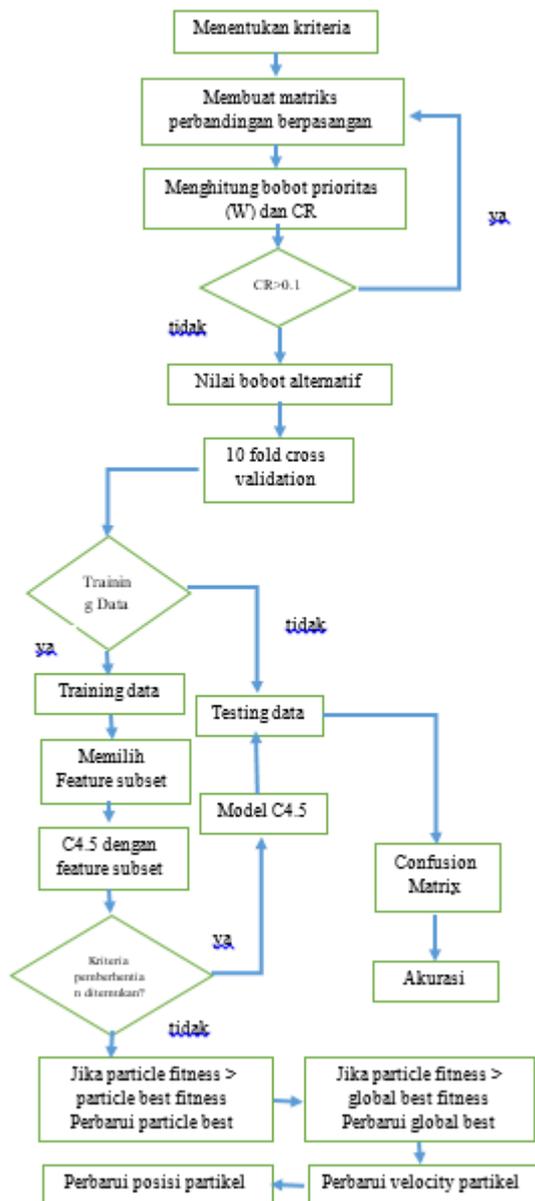
Tingkat Pendidikan	
Tingkat	Nilai
S3	95
S2	85
S1/D4	75
D3	70
SMA	65
SMP	60
SD	50

Transformasi data tingkat pendidikan dilakukan dengan mengubah setiap tingkat pendidikan menjadi nilai yang merepresentasikan tingkat pendidikan seperti pada tabel 5 diatas.

Tahap kelima yaitu pemodelan, pada tahap ini dilakukan pemodelan penelitian yang diusulkan yaitu kombinasi metode AHP, *Decision Tree* C 4.5, dan PSO yang diilustrasikan pada gambar 1 dibawah ini.

Dari ilustrasi gambar 1, proses AHP pada metode yang diusulkan, yang pertama yaitu menentukan kriteria-kriteria yang digunakan, kemudian membuat nilai prioritas pada setiap kriteria dari hasil wawancara dengan pejabat yang mengelola jabatan struktural kemudian menggunakan skala prioritas saaty dalam menerjemahkan pendapat yang mempunyai nilai skala 1 sampai dengan 9 (Saaty, 2008). Setelah menentukan nilai prioritas pada setiap kriteria dibuat matriks perbandingan berpasangan untuk menghitung bobot prioritas, kemudian mengukur konsistensi dengan menghitung *consistency index* atau indeks konsistensi dan menghitung nilai *Consistency Ratio*. Setelah itu dilakukan pemeriksaan konsistensi AHP yaitu apabila $CR > 0.1$ maka proses kembali diulang dengan menghitung matriks perbandingan berpasangan, apabila $CR \leq 0.1$ maka perhitungan sudah konsisten dan dianggap benar. Pada metode AHP dalam usulan penelitian ini, dilakukan modifikasi pada proses AHP

yaitu penghitungan bobot prioritas hanya dilakukan pada nilai bobot kriteria saja, yang kemudian dari nilai bobot kriteria dikalikan dengan nilai pada setiap nilai alternatif. Dari hasil tersebut akan digunakan dalam proses klasifikasi dan pemeringkatan setiap alternatif pegawai.



Gambar 1. Usulan Metode yang digunakan

Pada tahap klasifikasi menggunakan algoritma C4.5 dan optimasi PSO, langkah pertama yang dilakukan adalah membagi data berdasarkan 10 fold cross validation atau validasi silang sepuluh kali lipat. validasi silang sepuluh kali lipat membagi data kedalam 10 set, ukuran set data dibagi 10 kemudian 9 set data pelatihan dan 1 set data pengujian kemudian langkah itu diulang hingga 10 kali iterasi. Data pelatihan digunakan untuk membangun model sedangkan memvalidasi model menggunakan data

pengujian. Kemudian, data pelatihan digunakan untuk pemodelan C4.5 dan PSO dengan penghitungan C4.5 berdasarkan *feature subset* yang dihasilkan. Apabila kriteria pemberhentian belum didapat dalam PSO, seperti nilai *particle fitness* lebih besar dari *particle best fitness* dan *global best fitness*, maka perbarui *particle best fitness* dan *global best fitness*, selain itu dilakukan pembaruan posisi partikel dan *velocity* partikel.

Pembaruan nilai tersebut dilakukan untuk penghitungan algoritma *Decision Tree* C4.5 sampai kriteria pemberhentian terpenuhi, sehingga menghasilkan solusi terbaik (*fitness*) kemudian dilakukan perhitungan *Decision Tree* C4.5. Dari model *Decision Tree* C4.5 terbaik, kemudian dilakukan pengujian terhadap *Decision Tree* C4.5 menggunakan *confusion matrix* untuk menghitung nilai akurasi.

4. Hasil dan Pembahasan

4.1. Preprocessing

Pada penelitian ini, pada tahap *preprocessing* dilakukan pembersihan data sehingga didapat 333 data sampel yang lengkap, kemudian dilakukan transformasi data pada data presensi dan tingkat pendidikan sehingga didapatkan data sampel seperti pada tabel berikut :

Tabel 6. Sampel Data Setelah *Preprocessing*

ID Pegawai	PKSP	SKP	Presensi	Tingkat Pendidikan	Psikometri
1	90.5	87.82	50	75	56
2	84.86	94.31	50	65	59
3	81.67	87.82	60	75	70
4	85.56	87.66	80	65	53
5	83.19	65.78	90	85	61
6	93.48	87.83	90	85	67
7	62.88	65.88	100	75	59
8	90.54	86.73	90	85	60
9	87.24	87.33	90	70	55
10	85.97	87.55	70	75	55

4.2. Hasil AHP

Dari sampel data yang dihasilkan setelah *preprocessing* dilakukan proses AHP yang dimulai dengan membuat matriks berpasangan menggunakan skala prioritas yang sudah ditentukan pada kriteria kinerja seperti pada tabel 7 dan kinerja potensi seperti pada tabel 8.

Tabel 7. Matriks Berpasangan Kinerja

Kriteria Kinerja	SKP	PKSP	Presensi
SKP	1	2	7
PKSP	0.5	1	5
Presensi	0.142857	0.2	1

Tabel 8. Matriks Berpasangan Potensi

Kriteria Potensi	Psikometri	Tingkat Pendidikan
Psikometri	1	5
Pendidikan	0.2	1

Dari penghitungan matriks berpasangan dan konsistensi dari AHP, dihasilkan bobot prioritas dari kriteria kinerja seperti pada tabel 9 dan kriteria potensi seperti pada tabel 10.

Tabel 9. Bobot Prioritas Kinerja

Kriteria Kinerja	Bobot Prioritas
SKP	0.59248
PKSP	0.33307
Presensi	0.07445

Tabel 10. Bobot Prioritas Potensi

Kriteria Potensi	Prioritas
Psikometri	0.833
Pendidikan	0.167

Dari hasil yang didapat pada tabel 9 dan tabel 10 diatas, kemudian dilakukan perkalian nilai bobot prioritas dengan nilai kriteria pada setiap alternatif dan dilakukan penjumlahan dari setiap nilai tersebut sehingga menghasilkan peringkat pegawai berdasarkan kinerja dan potensi sebagai berikut :

Tabel 11. Tabel Peringkat Berdasarkan Kinerja

ID Pegawai	Kinerja
156	95.71
199	94.74
27	94.58
263	94.35
15	94.28
266	93.83
64	93.83
149	93.79
101	93.21
207	92.81

Pada tabel 11, dapat dilihat peringkat pegawai berdasarkan nilai kinerja yaitu peringkat pertama dengan kinerja terbaik yaitu pegawai dengan ID 156 dengan nilai 95.71, peringkat kedua yaitu pegawai dengan ID 199 dengan nilai 94.74, peringkat ketiga yaitu pegawai dengan ID 27 dengan nilai 94.58, dan seterusnya.

Pada tabel 12, dapat dilihat peringkat pegawai berdasarkan nilai potensi yaitu peringkat pertama dengan potensi terbaik yaitu pegawai dengan ID 319 dengan nilai 80.00, peringkat kedua yaitu pegawai dengan ID 320 dengan nilai 79.17, peringkat ketiga

yaitu pegawai dengan ID 305 dengan nilai 78.33, dan seterusnya.

Tabel 12. Tabel Peringkat Berdasarkan Potensi

ID Pegawai	Potensi
319	80.00
320	79.17
305	78.33
144	78.33
155	77.50
313	77.50
271	77.50
44	77.50
266	76.67
321	76.67

Dari hasil peringkat berdasarkan nilai kinerja dan potensi yang diperoleh diatas digunakan untuk sebagai dasar peringkat pegawai dan label klasifikasi kelayakan promosi jabatan.

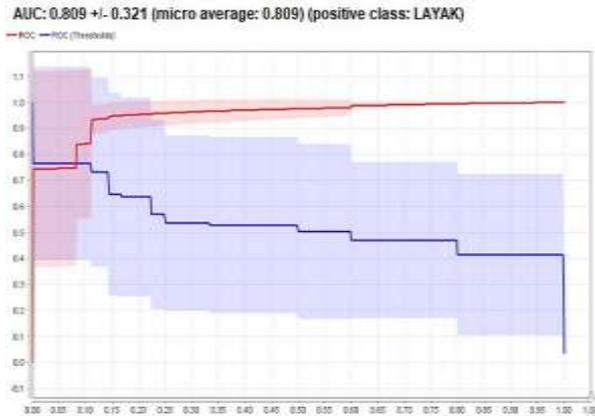
4.3. Klasifikasi dengan C4.5

Data gabungan dari sampel data dengan nilai kinerja dan potensi dari proses AHP didapatkan label klasifikasi yang kemudian data tersebut diproses menggunakan metode C4.5 menggunakan *tools* aplikasi RapidMiner. Parameter yang diatur pada metode C4.5 antara lain *criterion* yaitu *information gain*, *maximal depth* yaitu 5, *minimum leaf size* yaitu 2, *minimum size for split* yaitu 4, *minimum gain* yaitu 0, dan menggunakan validasi silang sepuluh kali lipat untuk membagi data menjadi latih dan data uji. Dari model dan parameter C 4.5 tersebut didapatkan hasil didapatkan nilai akurasi dari model C4.5 yaitu sebesar 93,40 % seperti pada tabel 13.

Tabel 13. Akurasi Model C4.5

	True TIDAK LAYAK	True LAYAK	Class precision
Pred. TIDAK LAYAK	55	14	79.71%
Pred. LAYAK	8	256	96.97%
Class recall	87.30%	94.81%	
Accuracy : 93.70% +/- 3.69% (micro average :93.39%)			

Nilai AUC yang dihasilkan dari model C4.5 didapatkan hasil sebesar 0.809 yang dapat dilihat pada gambar 2. Dengan hasil AUC sebesar 0.809 dapat disimpulkan bahwa model C4.5 yang digunakan merupakan model klasifikasi dengan kategori yang baik.



Gambar 2. AUC Model C4.5

Hasil dari *performance vector* yang dihasilkan dari model C4.5, didapat hasil nilai recall sebesar 94.69 %, nilai *precision* sebesar 96.96 %, nilai AUC (*optimistic*) sebesar 0.888 dan nilai AUC (*pessimistic*) sebesar 0.809 seperti pada gambar 3.

PerformanceVector

```
PerformanceVector:
accuracy: 93.80% +/- 3.63% (micro average: 93.33%)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   14
LAYAK:  4   260
precision: 96.96% +/- 3.79% (micro average: 96.97%) (positive class: LAYAK)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   14
LAYAK:  4   260
recall: 94.69% +/- 3.31% (micro average: 94.69%) (positive class: LAYAK)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   14
LAYAK:  4   260
AUC (optimistic): 0.888 +/- 0.312 (micro average: 0.888) (positive class: LAYAK)
AUC: 0.809 +/- 0.321 (micro average: 0.809) (positive class: LAYAK)
AUC (pessimistic): 0.836 +/- 0.27% (micro average: 0.836) (positive class: LAYAK)
```

Gambar 3. Performance Vector Model C4.5

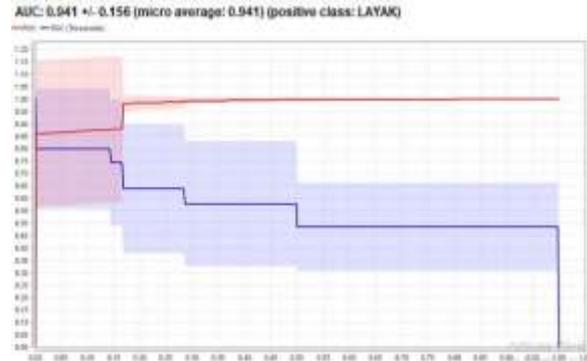
4.4. Klasifikasi dengan C4.5 dan PSO

Data dari proses AHP juga dilakukan pemrosesan dengan metode yang sudah diusulkan yaitu model C4.5 dan PSO menggunakan *tools* aplikasi Rapid Miner. Parameter yang digunakan pada metode C4.5 dan PSO antara lain pada PSO diatur parameter *population size* yaitu 10, *maximum number of generations* yaitu 30, bobot inersia yaitu 1, *local best weight* yaitu 1, *global best weight* yaitu 1, *minimal weight* yaitu 0 dan *maximal weight* yaitu 1. Pada C4.5 diatur parameter *criterion* yaitu *information gain*, *maximal depth* yaitu 5, *minimum gain* yaitu 0, *minimum leaf size* yaitu 2, *minimum size for split* yaitu 4, dan menggunakan validasi silang sepuluh kali lipat untuk membagi menjadi data latih dan data uji. Dari model dan parameter C4.5 dan PSO tersebut didapatkan hasil yaitu sebesar 95,80 % seperti pada tabel 14.

Tabel 14. Akurasi Model C4.5 dan PSO

	True TIDAK LAYAK	True LAYAK	Class precision
Pred. TIDAK LAYAK	59	10	85.51%
Pred. LAYAK	4	260	98.48%
Class recall	93.65%	96.30%	
Accuracy : 95.80% +/- 3.47% (micro average :95.80%)			

Nilai AUC yang dihasilkan dari model C4.5 dan PSO didapatkan hasil sebesar 0.941 yang dapat dilihat pada gambar 4. Dengan hasil AUC sebesar 0.941 dapat disimpulkan bahwa model C4.5 dan PSO yang digunakan merupakan model klasifikasi dengan kategori baik sekali.



Gambar 4. AUC Model C4.5 dan PSO

Hasil dari *performance vector* yang dihasilkan dari model C4.5 dan PSO, didapat hasil nilai *recall* sebesar 96.30 %, nilai *precision* sebesar 98.53 %, nilai AUC (*optimistic*) sebesar 0.997 dan nilai AUC (*pessimistic*) sebesar 0.969 seperti pada gambar 5.

PerformanceVector

```
PerformanceVector:
accuracy: 95.80% +/- 3.47% (micro average: 95.60%)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   10
LAYAK:  4   260
precision: 98.53% +/- 2.53% (micro average: 98.46%) (positive class: LAYAK)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   10
LAYAK:  4   260
recall: 96.30% +/- 3.02% (micro average: 96.30%) (positive class: LAYAK)
ConfusionMatrix:
True:  TIDAK LAYAK  LAYAK
TIDAK LAYAK:  59   10
LAYAK:  4   260
AUC (optimistic): 0.997 +/- 0.008 (micro average: 0.997) (positive class: LAYAK)
AUC: 0.941 +/- 0.156 (micro average: 0.941) (positive class: LAYAK)
AUC (pessimistic): 0.969 +/- 0.03% (micro average: 0.969) (positive class: LAYAK)
```

Gambar 5. Performance Vector Model C 4.5 dan PSO

4.5. Perbandingan Hasil

Dari hasil metode C4.5 dan PSO serta metode C4.5 didapat perbandingan pada kinerja klasifikasi seperti pada tabel 15 berikut.

Tabel 15. Perbandingan Hasil klasifikasi C4.5 dan PSO dengan C4.5

Metode	Akurasi	Precision	Recall	AUC
C4.5	93.40 %	96.96 %	94.69 %	0.809
C4.5 dan PSO	95.80 %	98.53 %	96.30 %	0.941

Dari Tabel 15, dapat dijelaskan bahwa kinerja yang dihasilkan dari metode C4.5 dan PSO mempunyai kinerja yang lebih baik dengan nilai akurasi 95.80 % dibanding dengan metode C4.5 yang mempunyai akurasi 93.40 %.

5. Kesimpulan

Berdasarkan hasil dari percobaan klasifikasi data pegawai menggunakan usulan metode kombinasi AHP, C4.5, dan PSO, terbukti dengan menggunakan metode AHP dapat melakukan pemeringkatan pegawai berdasarkan kinerja dan potensi yang tidak dapat dilakukan oleh metode C4.5, dan pada klasifikasi menggunakan C4.5 dengan optimasi PSO terbukti dapat meningkatkan akurasi hasil dibandingkan tanpa menggunakan optimasi PSO. Akurasi yang dihasilkan metode C4.5 dengan optimasi PSO sebesar 95.80 % lebih baik dari metode C4.5 sebesar 93.40 % dengan selisih akurasi sebesar 2.40 %, dan metode C4.5 dengan optimasi PSO mempunyai nilai AUC 0.941 yang merupakan klasifikasi kategori baik sekali, lebih baik daripada metode C4.5 yang mempunyai nilai AUC 0.809 yang merupakan klasifikasi kategori baik. Sehingga dari data hasil tersebut, disimpulkan kombinasi AHP, C4.5, dan PSO mempunyai kinerja yang lebih baik dan dapat digunakan sebagai dasar promosi jabatan pegawai berdasarkan klasifikasi kelayakan pegawai dan peringkat pegawai. Saran untuk penelitian selanjutnya yaitu perlu diperbandingkan dengan metode klasifikasi lainnya atau dengan penggunaan metode optimasi selain PSO agar dapat diketahui metode klasifikasi yang mempunyai kinerja yang paling baik.

Daftar Pustaka

- Ariyati, I., Rosyida, S., Ramanda, K., Riyanto, V., Faizah, S., Ridwansyah, 2020. Optimization of the decision tree algorithm used particle swarm optimization in the selection of digital payments. *Journal of Physics: Conference Series*, Vol. 1641. IOP Publishing Ltd.
- Ayuningtyas, A.K., Saleh, C., Noor I., 2017. Employee Promotion planning in analytical hierarchy process perspective: study on national public procurement agency. *Russian Journal of Agricultural and Socio-Economic Sciences* 70(10), 97–106. doi: 10.18551/rjoas.2017-10.16.
- Chen, K.H., Wang, K.J., Wang, K.M., Angelia, M.A., 2014. Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data. *Applied Soft Computing Journal* 24, 773–80. doi: 10.1016/j.asoc.2014.08.032.
- Dewi, R.K., Hanggara, B.T., Pinandito, P., 2018. A Comparison between AHP and Hybrid AHP for Mobile Based Culinary Recommendation System. *International Journal of Interactive Mobile Technologies* 12(1), 133–40. doi: 10.3991/ijim.v12i1.7561.
- Grosan, C., Abraham, A., Chis, M., 2006. Swarm Intelligence in Data Mining. *Studies in Computational Intelligence* 34, 1–20. doi: 10.1007/978-3-540-34956-3_1.
- Gustian, D., Hundayani, R.D., 2018. Combination of AHP method with C4.5 in the level classification level out students. 3rd International Conference on Computing, Engineering, and Design, ICCED 2017, Vols. 2018-March. Institute of Electrical and Electronics Engineers Inc.
- Haupt, R.L., Haupt, S.E., 2003. *Practical Genetic Algorithms*. Wiley.
- Lee, D., Kim, M.H., 2009. Chapter 2: Data Mining. *Database and Data Communication Network Systems*, Vol. 1 1, 41–47. doi: 10.1007/978-88-470-1163-2.
- Malik, M.M., Haouassi, H., 2021. Efficient sequential covering strategy for classification rules mining using a discrete equilibrium optimization algorithm. *Journal of King Saud University - Computer and Information Sciences*. doi: 10.1016/j.jksuci.2021.08.032.
- Meng, X., Zhang, P., Xu, Y., Xie, H., 2020. Construction of decision tree based on C4.5 algorithm for online voltage stability assessment. *International Journal of Electrical Power and Energy Systems* 118. doi: 10.1016/j.ijepes.2019.105793.
- Saaty, T.L., 2008. *Decision Making with the Analytic Hierarchy Process*, Vol. 1.
- Saha, A., Tasdid, M.N., Rahman, M.R., 2019. Mining Semantic Web Based Ontological Data. in 2018 21st International Conference of Computer and Information Technology, ICCIT 2018.
- Sedghiyan, D., Ashouri, A., Maftouni, N., Xiong, Q., Rezaee, E., Sadeghi, S., 2021. Prioritization of renewable energy resources in five climate zones in Iran Using AHP, Hybrid AHP-TOPSIS and AHP-SAW Methods. *Sustainable Energy Technologies and Assessments* 44. doi: 10.1016/j.seta.2021.101045.
- Tarigan, D.M., Rini, D.P., Sukemi. 2019. Particle swarm optimization-based on decision tree of C4.5 algorithm for Upper Respiratory Tract Infections (URTI) Prediction. in *Journal of Physics: Conference Series*, Vol. 1196.
- Wang, X., Zhou, C., Xu, X., 2019. Application of C4.5 Decision Tree for Scholarship Evaluations. in *Procedia Computer Science*, Vol. 151.
- Xiulin, S., Dawei, L., 2014. An improvement analytic hierarchy process and its application in teacher evaluation. in *Proceedings - 2014 5th International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2014*.