

IMPLEMENTASI PSO UNTUK OPTIMASI BOBOT ATRIBUT PADA ALGORITMA C4.5 DALAM PREDIKSI KELULUSAN MAHASISWA

Sharazita Dyah Anggita¹⁾, Ikma²⁾

^{1, 2)}Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta
Jl Ringroad Utara Condong Catur Depok Sleman Yogyakarta, Indonesia
e-mail: sharazita@amikom.ac.id¹⁾, ikmahdarwan01@amikom.ac.id²⁾

ABSTRAK

Ketepatan penyelesaian masa studi mahasiswa merupakan salah satu faktor yang banyak disoroti oleh perguruan tinggi. Algoritma C4.5 merupakan salah satu metode yang dapat digunakan dalam memprediksi kelulusan mahasiswa. Salah satu masalah yang terjadi pada data mining adalah semakin banyak penggunaan atribut, akan dapat memengaruhi tingkat akurasi dan kinerja dari algoritma. Penelitian ini akan dilakukan dengan melakukan improvement untuk dapat mengoptimasi nilai akurasi pada Algoritma C4.5. Improvement akan berfokus pada bobot yang berperan sebagai faktor kelulusan mahasiswa. Optimasi pembobotan atribut akan dilakukan menggunakan Algoritma PSO Melalui model yang diusulkan, peneliti akan melihat hasil perbandingan nilai akurasi penerapan Algoritma C4.5 dengan C4.5 yang dioptimasi dengan PSO. Hasil akurasi dari penerapan algoritma C4.5 dalam menentukan status kelulusan mahasiswa sebesar 90.73%. Proses optimasi yang dilakukan untuk membobotkan atribut pada algoritma C4.5 menggunakan algoritma PSO menghasilkan nilai akurasi sebesar 97.78%. Peningkatan performa didapat dari perbandingan hasil pengujian kedua algoritma dengan peningkatan akurasi sebesar 7.05%.

Kata Kunci: C4.5, PSO, prediksi kelulusan

ABSTRACT

The accuracy of completing the student's study period is one of the factors highlighted by universities. The C4.5 algorithm is one method that can be used to predict student graduation. One of the problems that occur in data mining is that the more use of attributes, will be able to affect the level of accuracy and performance of the algorithm. This research will be carried out by making improvements to optimize the accuracy value in the C4.5 Algorithm. Improvements will be focused on the weights that factor into student graduation. Attribute weighting optimization will be done by using PSO Algorithm. Through the proposed model, researchers will see the results of the comparison of the accuracy values of the C4.5 and C4.5 algorithms optimized by PSO. The result of the accuracy of the application of the C4.5 algorithm in determining the student's graduation status is 90.73%. The optimization process carried out for attribute weighting in the C4.5 algorithm using the PSO algorithm produces an accuracy value of 97.78%. Improved performance is obtained from the comparison of the test results of the two algorithms with an increase in accuracy of 7.05%.

Keywords: C4.5, PSO, graduation prediction

I. PENDAHULUAN

Ketepatan penyelesaian masa studi mahasiswa merupakan salah satu faktor yang banyak disoroti oleh perguruan tinggi. Prediksi kelulusan mahasiswa merupakan satu hal yang dapat digunakan untuk mengetahui status kelulusan mahasiswa di masa mendatang [1]. Universitas yang telah menghasilkan banyak lulusan dapat diteliti untuk menghasilkan sebuah prediksi suatu keputusan tertentu. Algoritma C4.5 merupakan salah satu metode yang dapat digunakan dalam memprediksi kelulusan mahasiswa. Penelitian [2] menghasilkan 10 rule prediksi menggunakan algoritma C4.5 dengan tingkat akurasi 65.98% dengan nilai AUC 0.874. Penelitian [3][4] yang melakukan perbandingan nilai akurasi algoritma C4.5 dengan algoritma klasifikasi lainnya menyimpulkan bahwa performa terbaik dihasilkan pada algoritma C4.5. Penerapan Algoritma C4.5 pada penelitian [5] [6] yang dilakukan mampu menghasilkan akurasi yang baik di atas 75%.

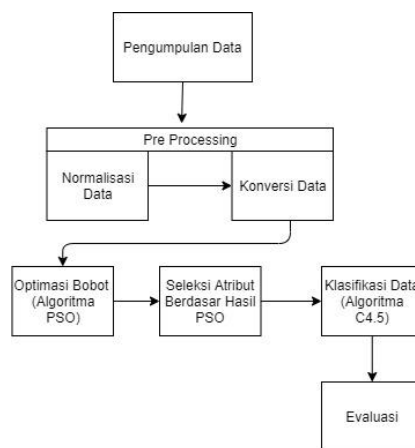
Fokus dari penelitian ini adalah melakukan *improvement* untuk dapat mengoptimasi nilai akurasi pada Algoritma C4.5. Algoritma PSO berfungsi untuk mencari partikel terbaik dengan proses iterasi untuk mencapai solusi optimal [7]. Algoritma PSO akan diterapkan untuk dapat membobotkan atribut pada klasifikasi algoritma C4.5. Penelitian [8] yang dilakukan dengan menerapkan algoritma Multi Objective Particle Swarm Optimization dinilai mampu menghasilkan pohon keputusan lebih akurat. Penelitian [9] yang dilakukan menggunakan Algoritma Naïve Bayes dan PSO mampu menghasilkan peningkatan akurasi diukur dengan confusion matrix dan kurva ROC dari 90.50% menjadi 96.92%. Penerapan Algoritma PSO sebagai seleksi fitur pada Algoritma Naïve Bayes Classifier

(NBC) mampu menghasilkan akurasi hingga 99,87% [10]. Selain itu pada penelitian [11] membuktikan bahwa Algoritma PSO dapat meningkatkan nilai akurasi sebesar 1.91 %.

Salah satu masalah yang terjadi pada data mining adalah semakin banyak penggunaan atribut, akan dapat memengaruhi tingkat akurasi dan kinerja dari algoritma [12]. Berdasarkan hasil dari beberapa penelitian yang sudah dilakukan, Improvement Algoritma C4.5 akan dilakukan pada penelitian ini dengan mengoptimasi bobot atribut yang digunakan sebagai faktor kelulusan mahasiswa menggunakan Algoritma PSO. Pada proses optimasi dilakukan juga pengujian untuk membandingkan nilai performa yang dihasilkan menggunakan parameter *inertia weight*. Hasil dari penelitian ini adalah perbandingan nilai akurasi dari penerapan algoritma C4.5 dengan penerapan optimasi bobot atribut menggunakan algoritma PSO pada algoritma C4.5.

II. METODE PENELITIAN

Penelitian dilakukan dengan melakukan pembagian tahap sebagai dasar implementasi sampai dengan mencapai tujuan penelitian. Tahapan metode penelitian dimodelkan pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

Tahap awal penelitian dilakukan dengan pengumpulan data. Data yang digunakan pada penelitian ini diambil dari data lulusan mahasiswa. Selanjutnya dataset akan dibagi menjadi dua bagian yaitu data training dan data testing. Preprocessing data dilakukan untuk menormalisasi dataset yang sudah disediakan, dimana proses yang akan dilakukan diantaranya penghapusan atribut yang tidak digunakan. Setiap atribut dengan nilai numeric dilakukan konversi data sesuai dengan kategori yang sudah ditentukan.

Atribut yang sudah ditentukan pada proses preprocessing dilakukan optimasi bobot menggunakan algoritma PSO sebagai bentuk optimasi algoritma klasifikasi yang akan diimplementasikan. Melalui hasil pembobotan algoritma PSO akan didapatkan bobot masing-masing atribut yang paling optimal sebagai dasar seleksi atribut yang akan digunakan dalam penentuan status kelulusan mahasiswa.

Tahap klasifikasi data dilakukan dengan mengimplementasikan algoritma C4.5 yang akan membentuk pohon keputusan dalam penentuan status kelulusan mahasiswa sesuai dengan atribut predictor yang telah ditentukan di awal. Hasil pengujian yang dihasilkan berupa tabel akurasi prediksi kelulusan mahasiswa dengan membandingkan klasifikasi Algoritma C4.5 saja dengan Optimasi Bobot Atribut menggunakan Particle Swarm Optimization pada Algoritma C4.5

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah data lulusan mahasiswa. Dataset yang didapatkan dari data wisuda tahun 2020 yang digabungkan dengan data pribadi dan akademik mahasiswa yang selanjutnya akan dijadikan atribut penentu status kelulusan mahasiswa.

B. Preprocessing

Dataset pada Tabel 1 dilakukan beberapa tahap proses eliminasi dengan menghapuskan atribut yang tidak digunakan sebagai penentu kelulusan.

TABEL I
DATASET AWAL

NO	KODE PT	KODE PRODI	NAMA PRODI	NIM	JENIS KELAMIN	TANGGAL MASUK	PERIODE PENDAFTARAN	JUMLAH SKS	IPK	JENIS PENDAFTARAN	TANGGAL YUDISIUM
1	51024	57201	Sistem Informasi S1	13127298	L	1/9/2013	2013/2014 Ganjil	144	3.38	Peserta didik baru	7/25/2020
2	51024	57201	Sistem Informasi S1	13127444	L	1/9/2013	2013/2014 Ganjil	144	2.97	Peserta didik baru	7/25/2020
3	51024	57201	Sistem Informasi S1	13127560	P	1/9/2013	2013/2014 Ganjil	144	3.17	Peserta didik baru	7/25/2020
4	51024	57201	Sistem Informasi S1	13127691	L	1/9/2013	2013/2014 Ganjil	144	2.74	Peserta didik baru	7/25/2020
5	51024	57201	Sistem Informasi S1	14128256	L	1/9/2014	2014/2015 Ganjil	146	2.6	Peserta didik baru	7/25/2020
6	51024	57201	Sistem Informasi S1	15128349	L	1/9/2015	2015/2016 Ganjil	144	3.43	Peserta didik baru	7/25/2020
7	51024	57201	Sistem Informasi S1	15128561	L	1/9/2015	2015/2016 Ganjil	144	3.15	Peserta didik baru	7/25/2020
8	51024	57201	Sistem Informasi S1	16129043	L	1/9/2016	2016/2017 Ganjil	144	3.15	Peserta didik baru	7/25/2020
9	51024	57201	Sistem Informasi S1	16129141	P	1/9/2016	2016/2017 Ganjil	144	3.78	Peserta didik baru	7/25/2020
10	51024	57201	Sistem Informasi S1	16129328	P	1/9/2016	2016/2017 Ganjil	144	3.75	Peserta didik baru	7/25/2020
11	51024	57201	Sistem Informasi S1	16129342	P	1/9/2016	2016/2017 Ganjil	144	3.61	Peserta didik baru	7/25/2020
12	51024	57201	Sistem Informasi S1	16129432	L	1/9/2016	2016/2017 Ganjil	144	3.71	Peserta didik baru	7/25/2020
13	51024	57201	Sistem Informasi S1	16129498	L	1/9/2016	2016/2017 Ganjil	144	3.72	Peserta didik baru	7/25/2020
14	51024	57201	Sistem Informasi S1	16129542	L	1/9/2016	2016/2017 Ganjil	144	3.08	Peserta didik baru	7/25/2020
15	51024	57201	Sistem Informasi S1	18222067	L	1/9/2018	2018/2019 Ganjil	148	3.82	Alih Jenjang	7/25/2020
16	51024	57201	Sistem Informasi S1	18222083	L	1/9/2018	2018/2019 Ganjil	146	3.9	Alih Jenjang	7/25/2020
17	51024	57201	Sistem Informasi S1	13127135	P	1/9/2013	2013/2014 Ganjil	150	2.21	Peserta didik baru	8/31/2020
18	51024	57201	Sistem Informasi S1	13127151	L	1/9/2013	2013/2014 Ganjil	144	3	Peserta didik baru	8/31/2020
19	51024	57201	Sistem Informasi S1	13127180	L	1/9/2013	2013/2014 Ganjil	144	3.04	Peserta didik baru	8/31/2020
20	51024	57201	Sistem Informasi S1	13127182	L	1/9/2013	2013/2014 Ganjil	144	2.4	Peserta didik baru	8/31/2020
21	51024	57201	Sistem Informasi S1	13127184	P	1/9/2013	2013/2014 Ganjil	144	3.33	Peserta didik baru	8/31/2020
22	51024	57201	Sistem Informasi S1	13127190	L	1/9/2013	2013/2014 Ganjil	144	3.29	Peserta didik baru	8/31/2020
23	51024	57201	Sistem Informasi S1	13127206	L	1/9/2013	2013/2014 Ganjil	144	2.58	Peserta didik baru	8/31/2020
24	51024	57201	Sistem Informasi S1	13127221	P	1/9/2013	2013/2014 Ganjil	144	2.5	Peserta didik baru	8/31/2020
25	51024	57201	Sistem Informasi S1	13127255	L	1/9/2013	2013/2014 Ganjil	144	2.61	Peserta didik baru	8/31/2020
26	51024	57201	Sistem Informasi S1	13127303	L	1/9/2013	2013/2014 Ganjil	144	3.08	Peserta didik baru	8/31/2020
27	51024	57201	Sistem Informasi S1	13127309	L	1/9/2013	2013/2014 Ganjil	146	2.84	Peserta didik baru	8/31/2020
28	51024	57201	Sistem Informasi S1	13127320	L	1/9/2013	2013/2014 Ganjil	144	3.69	Peserta didik baru	8/31/2020
29	51024	57201	Sistem Informasi S1	13127338	P	1/9/2013	2013/2014 Ganjil	144	3.46	Peserta didik baru	8/31/2020
30	51024	57201	Sistem Informasi S1	13127351	L	1/9/2013	2013/2014 Ganjil	144	2.78	Peserta didik baru	8/31/2020
...
356	51024	57201	Sistem Informasi S1	13127376	L	1/9/2013	2013/2014 Ganjil	146	2.58	Peserta didik baru	8/31/2020

Melalui proses eliminasi atribut tahap awal ditetapkan sepuluh atribut yang akan digunakan dalam preprocessing selanjutnya pada Tabel 2. Berdasarkan data kelulusan mahasiswa tahun 2020 terdapat 2 status kelulusan yaitu lulus tepat waktu dan tidak tepat waktu. Prosentase perbandingan mahasiswa yang lulus tepat waktu sebanyak 25% dan yang tidak lulus tepat waktu sebanyak 75%.

TABEL 2
DATASET HASIL PREPROCESSING TAHAP 1

JENIS KELAMIN	JUMLAH SKS	IPK	JENIS PENDAFTARAN	UMUR MASUK	MASA STUDI	STATUS
L	144	3.38	Peserta didik baru	17	7	Tidak Tepat
L	144	2.97	Peserta didik baru	18	7	Tidak Tepat
P	144	3.17	Peserta didik baru	17	7	Tidak Tepat
L	144	2.74	Peserta didik baru	19	7	Tidak Tepat
L	146	2.6	Peserta didik baru	18	6	Tidak Tepat
L	144	3.43	Peserta didik baru	18	5	Tidak Tepat
L	144	3.15	Peserta didik baru	17	5	Tidak Tepat
L	144	3.15	Peserta didik baru	21	4	Tidak Tepat
P	144	3.78	Peserta didik baru	17	4	Tepat
P	144	3.75	Peserta didik baru	16	4	Tepat
P	144	3.61	Peserta didik baru	18	4	Tepat
L	144	3.71	Peserta didik baru	19	4	Tepat
L	144	3.72	Peserta didik baru	17	4	Tepat
L	144	3.08	Peserta didik baru	17	4	Tidak Tepat
P	150	2.21	Peserta didik baru	17	7	Tidak Tepat
L	144	3	Peserta didik baru	17	7	Tidak Tepat
...
L	144	2.4	Peserta didik baru	19	7	Tidak Tepat

Dataset hasil dari *preprocessing* tahap 1 selanjutnya dikonversi sesuai dengan kebutuhan atribut yang akan digunakan pada proses klasifikasi data menggunakan algoritma C4.5. Kategori atribut dijelaskan pada Tabel 3. Selanjutnya bentuk dataset final setelah dilakukan konversi terdapat pada Tabel 3. Konversi data sesuai kategori dilakukan untuk dapat menyederhanakan nilai yang terdapat pada setiap atribut [13].

TABEL 3
DATASET FINAL

JENIS KELAMIN	JUMLAH SKS	IPK	UMUR MASUK	MASA STUDI	STATUS
L	Lengkap	Sedang	Muda	Lama	Tidak Tepat
L	Lengkap	Sedang	Muda	Lama	Tidak Tepat
P	Lengkap	Sedang	Muda	Lama	Tidak Tepat
L	Lengkap	Kecil	Muda	Lama	Tidak Tepat
L	Kurang	Kecil	Muda	Lama	Tidak Tepat
L	Lengkap	Sedang	Muda	Lama	Tidak Tepat
L	Lengkap	Sedang	Muda	Lama	Tidak Tepat
L	Lengkap	Sedang	Tua	Cepat	Tidak Tepat
P	Lengkap	Besar	Muda	Cepat	Tepat
P	Lengkap	Besar	Muda	Cepat	Tepat
P	Lengkap	Besar	Muda	Cepat	Tepat
L	Lengkap	Besar	Muda	Cepat	Tepat
L	Lengkap	Besar	Muda	Cepat	Tepat
L	Lengkap	Sedang	Muda	Cepat	Tidak Tepat
L	Kurang	Besar	Tua	Cepat	Tidak Tepat
L	Kurang	Besar	Tua	Cepat	Tidak Tepat
P	Kurang	Kecil	Muda	Lama	Tidak Tepat

TABEL 4
KONVERSI KATEGORI ATRIBUT

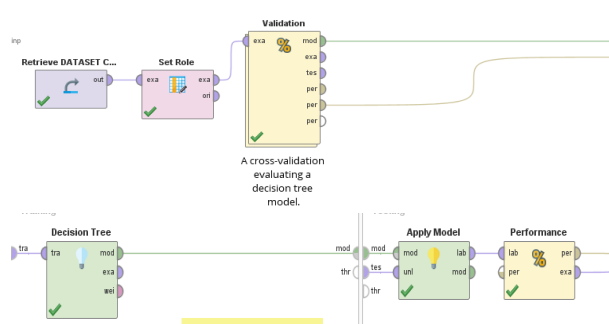
NO	ATRIBUT	KETERANGAN	KATEGORI
1	Jumlah Sks	≥ 144 sks	Lengkap
		< 144 sks	Kurang
2	IPK	≥ 3.5	Besar
		3.5 sd 2.75	Sedang
		≤ 2.75	Kecil
3	Umur Masuk	≥ 20 tahun	Tua
		< 20 tahun	Muda
4	Masa Studi	≤ 4 tahun	Cepat
		> 4 tahun	Lama

Atribut yang selanjutnya akan digunakan pada Tabel 4 dalam penerapan metode adalah Jenis kelamin, Jumlah sks, IPK, Umur masuk, Masa studi dan Status

C.Penerapan Model

1. Klasifikasi Data C4.5

Klasifikasi data awal dilakukan dengan menggunakan algoritma C4.5. Tahap awal yang harus dilakukan sebelum melakukan penerapan algoritma adalah menentukan atribut predictor sebagai akar pohon keputusan. Gambar 2 merupakan model penerapan algoritma C4.5. Atribut predictor ditentukan menggunakan operator Set Role, dalam kasus ini atribut yang dipilih adalah atribut Status. Operator *cross validation* digunakan sebagai proses evaluasi yang akan memisahkan data menjadi data latih dan uji[14]. Hasil dari proses evaluasi berupa model pohon keputusan dan nilai performa algoritma yang diujikan.



Gambar 2. Model Penerapan Algoritma C4.5

Penerapan model di atas menghasilkan nilai performa pada Tabel 4 dengan nilai akurasi sebesar 97.75%, nilai precision sebesar 91.94% dan nilai recall sebesar 95%.

TABEL 4
HASIL PERFORMA PENERAPAN MODEL C4.5

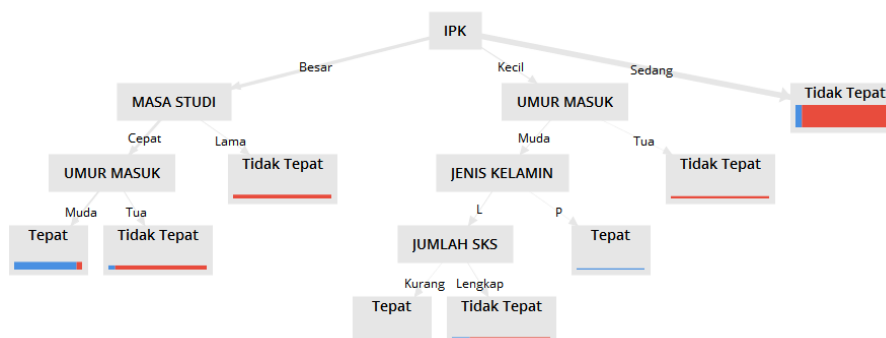
AKURASI (%)	PRECISION (%)	RECALL (%)
90.73	91.94	95

Berdasarkan data di Tabel 5 disimpulkan bahwa : a) Jumlah data Tepat yang benar diprediksi Tepat = 58 data b) Jumlah data Tepat yang diprediksi Tidak Tepat = 11 data c) Jumlah data Tidak Tepat yang diprediksi Tepat = 22 data d) Jumlah data Tidak Tepat yang benar diprediksi Tidak Tepat = 265 data.

TABEL 5
HASIL PREDIKSI PENERAPAN C4.5

	True Tepat	True Tidak Tepat
pred Positive	58	11
pred Negative	22	265

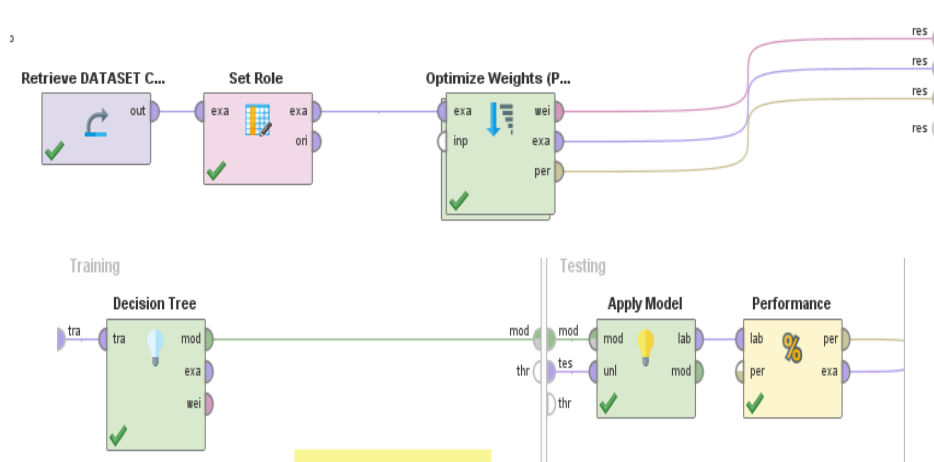
Pohon keputusan yang dihasilkan dari penerapan model di atas terdapat pada Gambar 3.



Gambar 3. Pohon Keputusan Penerapan Algoritma C4.5

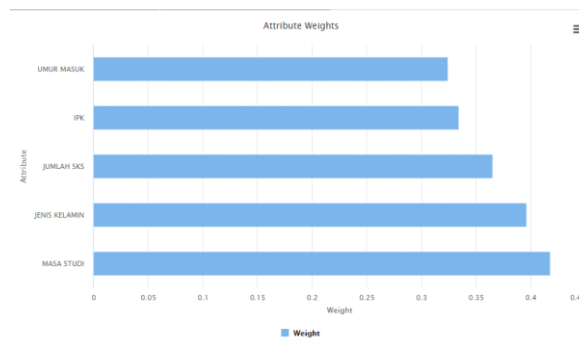
2. Optimasi Bobot Menggunakan PSO pada Algoritma C4.5

Penerapan algoritma PSO pada C4.5 pada penelitian [15][16][17] mampu menghasilkan peningkatan nilai akurasi lebih dari 5% dari penerapan algoritma C4.5 saja, hal ini didukung dengan nilai akurasi yang dihasilkan lebih dari 95%. Penerapan optimasi bobot menggunakan PSO pada algoritma C4.5 pada Gambar 4. Operator Optimize Weight digunakan untuk menghasilkan nilai pembobotan dari lima atribut yang digunakan sebagai penentu status kelulusan mahasiswa. Dengan menggunakan operator tersebut tingkat akurasi yang dihasilkan dari klasifikasi menggunakan algoritma C4.5 akan lebih optimal.



Gambar 4. Penerapan PSO pada C4.5

Penerapan algoritma PSO pada pembobotan atribut menghasilkan nilai bobot dari kelima atribut yang digunakan pada Gambar 5. Atribut jenis kelamin menghasilkan bobot nilai 0.396, jumlah sks 0.365, ipk 0.334, umur masuk 0.324 dan masa studi 0.418. Berdasarkan hasil pembobotan diputuskan bahwa semua atribut tetap digunakan pada proses klasifikasi melihat gap tiap bobot atribut yang tidak terlalu signifikan.



Gambar 5. Hasil Pembobotan Atribut Menggunakan PSO

Penerapan algoritma PSO pada pembobotan atribut di algoritma C4.5 menghasilkan nilai performa pada Tabel 6. Tingkat akurasi yang dihasilkan sebesar 97.78%, nilai precision sebesar 93.21% dan nilai recall sebesar 95%.

TABEL 6
 HASIL PERFORMA PENERAPAN MODEL C4.5 PSO

AKURASI (%)	PRECISION (%)	RECALL (%)
97.78	93.21	95

Berdasarkan data di Tabel 7 disimpulkan bahwa : a) Jumlah data Tepat yang benar diprediksi Tepat = 291 data b) Jumlah data Tepat yang diprediksi Tidak Tepat = 3 data c) Jumlah data Tidak Tepat yang diprediksi Tepat = 5 data d) Jumlah data Tidak Tepat yang benar diprediksi Tidak Tepat = 57 data.

TABEL 7
HASIL PREDIKSI PSO PADA MODEL C4.5 PSO

Accuracy = 97.78% +/- 2.87% (mikro average : 97.75%)		
	True Tepat	True Tidak Tepat
pred Positive	291	3
pred Negative	5	57

Proses pengujian dilanjutkan dengan melihat pengaruh parameter inertia weight terhadap tingkat performa yang dihasilkan. Pengujian dilakukan dengan menggunakan 5 nilai inertia weight pada Tabel 8.

TABEL 8
PERFORMA ALGORITMA BERDASARKAN PARAMETER INERTIA WEIGHT

INERTIA WEIGHT	AKURASI (%)	PRECISION (%)	RECALL (%)
1	97.78	92.38	95
2	97.78	92.64	95
3	97.78	92.14	95
4	97.78	93.21	95
5	97.78	93.21	95

Berdasarkan hasil pengujian pada Tabel 8, parameter inertia weight mampu mempengaruhi performa algoritma yang dimodelkan khususnya pada nilai precision. Terdapat perbedaan nilai precision yang dihasilkan dengan penerapan parameter tersebut. Sedangkan untuk nilai akurasi dan recall tidak mengalami perubahan nilai yang dapat disimpulkan bahwa penerapan parameter inertia weight tidak mempengaruhi tingkat akurasi dan recall.

D. Evaluasi

Tahap evaluasi dilakukan untuk dapat melihat secara jelas perubahan nilai performa dari penerapan algoritma C4.5 saja dengan penambahan optimasi bobot menggunakan algoritma PSO.

TABEL 9
PERBANDINGAN NILAI AKURASI

ALGORITMA	AKURASI (%)	PRECISION (%)	RECALL (%)
C4.5	90.73	91.94	95
C4.5 + PSO	97.78	93.21	95
Peningkatan	7.05	1.27	0

Pengujian algoritma C4.5 pada penentuan status kelulusan mahasiswa menghasilkan nilai akurasi sebesar 90.73% dengan jumlah data lulus Tepat yang benar diprediksi Tepat 291 data. Sedangkan data lulus Tidak Tepat yang benar diprediksi Tidak Tepat sejumlah 57 data.

Optimasi dilakukan dengan menerapkan algoritma PSO pada pembobotan seluruh atribut di algoritma C4.5. Model optimasi tersebut menghasilkan nilai akurasi sebesar 97.78%. Data lulus Tepat yang dihasilkan sesuai prediksi Tepat berjumlah 58 data. Data lulus Tidak Tepat yang benar diprediksi Tidak Tepat sejumlah 265 data. Berdasarkan data pada Tabel 9 peningkatan nilai performa terdapat pada nilai akurasi sebesar 7.05% dan nilai precision sebesar 1.27%.

IV. KESIMPULAN

Penerapan algoritma C4.5 dalam menentukan status kelulusan mahasiswa mampu menghasilkan nilai akurasi sebesar 90.73%. Proses optimasi yang dilakukan untuk membobotkan atribut pada algoritma C4.5 menggunakan algoritma PSO menghasilkan nilai akurasi sebesar 97.78%. Berdasarkan proses pengujian keduanya diketahui

bahwa penerapan PSO pada pembobotan atribut di algoritma C4.5 mampu meningkatkan performa algoritma C4.5. Peningkatan performa didapat dari peningkatan akurasi sebesar 7.05% dan nilai *precision* 1.27%. Pengujian dengan melakukan perbandingan parameter *inertia weight* menghasilkan bahwa parameter tersebut hanya mempengaruhi nilai *precision* serta tidak mempengaruhi nilai akurasi dan *recall*.

DAFTAR PUSTAKA

- [1] E. Etriyanti, D. Syamsuar, and N. Kunang, "Implementasi Data Mining Menggunakan Algoritma Naive Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa," *Telematika*, vol. 13, no. 1, pp. 56–67, 2020, doi: 10.35671/telematika.v13i1.881.
- [2] A. Rohman *et al.*, "Implementasi Data Mining Dengan Algoritma Decision Tree C4.5 Untuk Prediksi Kelulusan Mahasiswa Di Universitas," pp. 134–139, 2019.
- [3] A. Nurzahputra and M. A. Muslim, "Peningkatan Akurasi Pada Algoritma C4.5 Menggunakan Adaboost Untuk Meminimalkan Resiko Kredit," *Pros. SNATIF*, no. 1, pp. 243–247, 2017, [Online]. Available: <https://media.neliti.com/media/publications/173704-ID-none.pdf>.
- [4] E. N. Azizah, U. Pujianto, E. Nugraha, and Darusalam, "Comparative performance between C4.5 and Naive Bayes classifiers in predicting student academic performance in a Virtual Learning Environment," *2018 4th Int. Conf. Educ. Technol. ICET 2018*, no. 1, pp. 18–22, 2018, doi: 10.1109/ICEAT.2018.8693928.
- [5] R. P. S. Putri and I. Waspada, "Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, p. 1, 2018, doi: 10.23917/khif.v4i1.5975.
- [6] G. Kostopoulos, N. Fazakis, S. Kotsiantis, and K. Sgarbas, "Multi-objective Optimization of C4.5 Decision Tree for Predicting Student Academic Performance," *10th Int. Conf. Information, Intell. Syst. Appl. IISA 2019*, pp. 1–4, 2019, doi: 10.1109/IISA.2019.8900771.
- [7] S. Sundaramurthy and P. Jayavel, "A hybrid Grey Wolf Optimization and Particle Swarm Optimization with C4.5 approach for prediction of Rheumatoid Arthritis," *Appl. Soft Comput. J.*, vol. 94, p. 106500, 2020, doi: 10.1016/j.asoc.2020.106500.
- [8] I. S. Damanik, A. P. Windarto, A. Wanto, Poningsih, S. R. Andani, and W. Saputra, "Decision Tree Optimization in C4.5 Algorithm Using Genetic Algorithm," *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012012.
- [9] F. Pramono, Didi Rosiyadi, and Windu Gata, "Integrasi N-gram, Information Gain, Particle Swarm Optimization di Naive Bayes untuk Optimasi Sentimen Google Classroom," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 383–388, 2019, doi: 10.29207/resti.v3i3.1119.
- [10] E. Junianto and D. Riana, "Penerapan PSO Untuk Seleksi Fitur Pada Klasifikasi Dokumen Berita Menggunakan NBC," vol. 4, no. 1, pp. 38–45, 2017.
- [11] S. Antar, B. Vol, V. I. No, and E. Supriyadi, "Metode SVM Berbasis PSO untuk Meningkatkan Prediksi Ketepatan Waktu Kelulusan Mahasiswa," no. 2, pp. 113–120, 2017.
- [12] M. Hilman Aprilian Nurjaman and M. Syahrul Mubarak, "Analisis Sentimen Pada Ulasan Buku Berbahasa Inggris Menggunakan Information Gain Dan Support Vector Machine Sentiment Analysis on the English Book Reviews Using Information Gain and Support Vector Machine," vol. 4, no. 3, pp. 4900–4906, 2017.
- [13] B. Ferdiansyah and L. Goeirmanto, "Prediksi Loyalitas dalam Keterikatan Karyawan terhadap Perusahaan Menggunakan Algoritma C4.5* (Studi Kasus PT.XYZ)," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 1, p. 87, 2020, doi: 10.26418/justin.v8i1.33606.
- [14] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm," *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [15] M. Wasil, "濟無No Title No Title No Title," vol. 2, no. 2, pp. 1–23, 2016.
- [16] I. Yulianti, R. A. Saputra, M. S. Mardiyanto, and A. Rahmawati, "Optimasi Akurasi Algoritma C4.5 Berbasis Particle Swarm Optimization dengan Teknik Bagging pada Prediksi Penyakit Ginjal Kronis," *Techno.Com*, vol. 19, no. 4, pp. 411–421, 2020, doi: 10.33633/tc.v19i4.3579.
- [17] E. Karyadiputra *et al.*, "Penerapan Algoritma Decision Tree C4.5 Berbasis Particle Swarm Optimization (Pso) Untuk Prediksi Tingkat Kepuasan Pelayanan Obat," *Technologia*, vol. 12, no. 2, pp. 84–88, 2021.