

## OPTIMASI ALGORITMA C4.5 BERBASIS PARTICLE SWARM OPTIMIZATION (PSO) UNTUK MENENTUKAN WHOLESAL PENJUALAN

Yosef Mulyanto Dawa<sup>1</sup>, Abdul Aziz<sup>2</sup>, Moh Ahsan<sup>3</sup>

Universitas PGRI Kanjuruhan Malang<sup>1,2,3</sup>

email: [dawamulyanto@gmail.com](mailto:dawamulyanto@gmail.com), [abdul.aziz@unikama.ac.id](mailto:abdul.aziz@unikama.ac.id), [ahsan@unikama.ac.id](mailto:ahsan@unikama.ac.id)

### Abstrak

Algoritma C4.5 yaitu algoritma yang digunakan untuk membentuk Decision Tree. Pada Algoritma C4.5 pasti memiliki sebuah kelebihan dan kekurangan. Kelebihan algoritma C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik, dan dalam sebuah kelebihan pasti ada kekurangan. Kekurangan Algoritma C4.5 yaitu ketidakstabilan dalam menentukan akurasi. Jumlah data yang digunakan ialah sebanyak 1000 dengan 7 atribut. Data dianalisa menggunakan Particle Swarm optimization pada C4.5. Karena akurasi yang dihasilkan C4.5 masih rendah maka dioptimasi dengan Particle Swarm optimization. Akurasi pada C4.5 81% setelah digunakan optimasi akurasinya meningkat 86%. Pengolahan data menggunakan pemrograman python dan pengujian akurasi menggunakan *Confusion Matrix* untuk membandingkan hasil akurasi.

**Kata kunci :** *Klasifikasi; C4.5; Particle Swarm optimization (PSO)*

### Abstract

*Algorithm C4.5 is an algorithm used to form a Decision Tree. The C4.5 Algorithm definitely has advantages and disadvantages. The features of the C4.5 algorithm can create decision trees that are easy to version, dominate the level of acceptable accuracy, efficient in managing effective category attributes and can set attributes of discrete and numeric types, and in an advantage there are definitely drawbacks. The weakness of the C4.5 algorithm is the instability in determining accuracy. The amount of data used is 1000 with 7 attributes. Data were analyzed using Particle Swarm optimization in C4.5. Because the accuracy produced by C4.5 is still low, it is optimized with Particle Swarm optimization. Accuracy on C4.5 is 81% after using optimization the accuracy increases by 86%. Data processing uses Python programming and accuracy testing uses the Confusion Matrix to compare accuracy results.*

**Keyword :** *Classification; C4.5; Based Particle Swarm Optimization (PSO)*

*Optimasi Algoritma C4.5 Berbasis Particle Swarm Optimization (PSO) Untuk Menentukan Wholesales Penjualan*

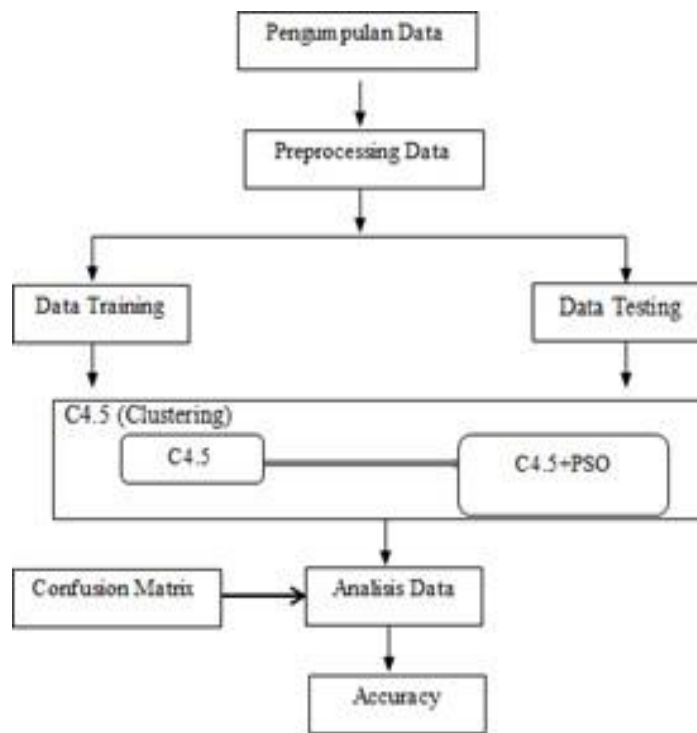
## 1. PENDAHULUAN

Sebuah perusahaan perlu menggunakan sebuah teknik pada bidang teknologi informasi untuk memprediksi. Data mining ialah sebuah proses dengan mengetahui embaran dan wawasan yang terikat melalui beraneka macam database yang superior. Proses kodifikasi sistem C4.5 pada Data Mining adalah sebuah metode yang dapat digunakan dalam memprediksi nilai penjualan barang.

Metode klasifikasi untuk memprediksi ialah pohon keputusan yang berfungsi untuk mengeksplorasi data dan mendapatkan tautan yang tertanam antar sejumlah calon variabel input dalam sebuah variabel target. Pada Algoritma C4.5 memiliki sebuah kelebihan dan kekurangan. Dengan kelebihan dapat menciptakan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik, dan dari sebuah kelebihan pasti ada kekurangan. Kekurangan Algoritma C4.5 yaitu ketidakstabilan dalam menentukan akurasi. Untuk menutupi sebuah permasalahan pada kekurangan Algoritma C4.5 kita membutuhkan sebuah Algoritma Pengoptimasi untuk meningkatkan akurasi.

Metode Particle Swarm Optimization (PSO) dipublikasikan melalui (Kennedy & Eberhart, 1995) yakni salah satu dari teknik komputasi evolusioner, yang mana asosiasi pada PSO didasarkan pada pencarian algoritma dan didahului dengan suatu kelompok yang random yang disebut dengan particle. Tujuan penelitian untuk meningkatkan akurasi dan efisiensi algoritma C4.5 dalam menentukan wholesales penjualan yang optimal dengan mengaplikasikan teknik *Particle Swarm Optimization (PSO)*. Dengan deskripsi di atas bahwa pengguna menerapkan tema tugas akhir ialah, ”Optimal metode C4.5 Berbasis *Particle Swarm Optimization* Untuk Menentukan Wholesales Penjualan”.

Pada percobaan ini implementasi data mining dengan metode C4.5 menggunakan Particle swarm optimization untuk prediksi data penjualan mobil di PT.Gaikindo. Terdapat beberapa tahapan Adapun tahapan umum pada penelitian ini adalah seperti ditunjukkan pada gambar berikut:



Gambar 1. Metode Penelitian

### 3. HASIL DAN PEMBAHASAN

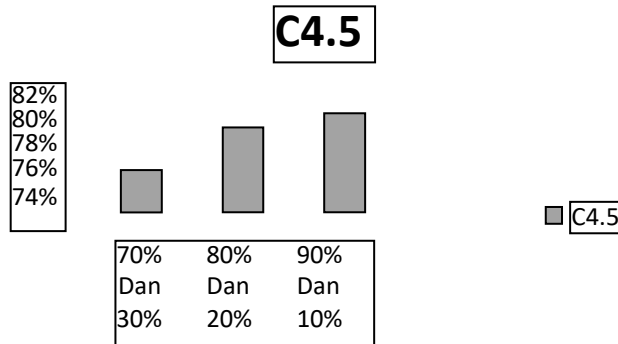
#### Mencari Nilai Entropy dan Gain

Tabel 4. 1 nilai Entropy dan Gain

Node		Keterangan	Jumlah Kasus	Laris	Kurang Laris	Entropy	Gain
1	Total		1000	514	482	-1,00102	-92,0233566
	Brand	BMW	110	14	96	0,549911	
		CHEVROLET	14	8	6	0,062537	
		HONDA	91	61	30	0,140952	
		TOYOTA	179	83	96	-0,03206	
		AUDI	19	0	19	0	
		HAZDA	3	1	2	-0,13835	
		LEXUS	17	5	12	-0,16457	
		NISSAN	53	26	27	-0,00835	
		DAIHATSU	68	54	14	0,733538	
		DATSUN	7	6	1	0,210429	
		DFSK	23	7	16	-0,33163	
		MINI	42	9	33	-0,20286	
		HYUNDAI	10	3	7	0,881291	
		MAZDA	13	1	12	-0,17805	
	SUZUKI	104	85	19	-0,68593		
2	Wheel Base	LuarNegri	600	335	265	0,051264	1,060558332
		DalamNegri	400	233	167	0,071956	
3	CC	>2000	258	198	60	-0,20426	0,39676642
		2000	364	221	143	0,092464	
		<2000	378	183	195	-0,99927	

**Pengujian Algoritma C4.5**

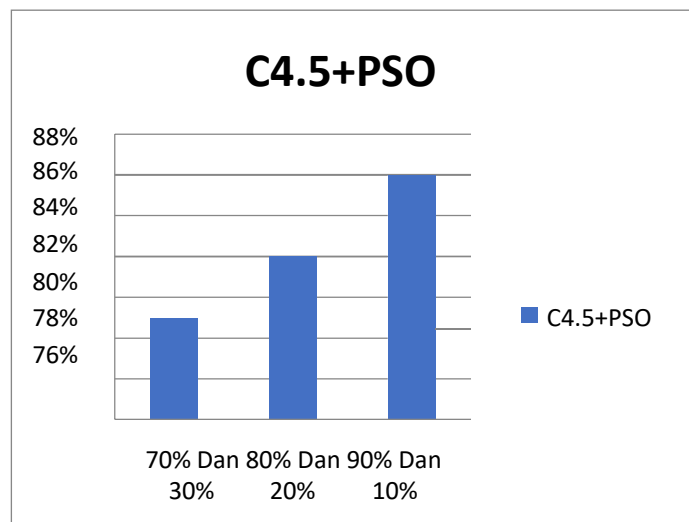
Merupakan pembelajaran mesin yang memiliki pohon keputusan dan dapat digunakan untuk klasifikasi. Pada C4.5, kita dapat menggunakan data pelatihan (training data) untuk membangun pohon keputusan, dan kemudian menguji model tersebut menggunakan data pengujian (testing data). Dalam kasus ini, kita memiliki 3 rasio data pelatihan dan data pengujian yaitu 70:30,80:20,90:10 dan mendapatkan hasil akurasi sebagai berikut:



**Gambar 2 Grafik Hasil Pengujian C4.5**

**Pengujian C4.5 Menggunakan Algoritma Particle Swarm Optimization (PSO)**

Pengujian dengan data dari hasil C4.5. Tahapan ini akan ditambahkan operator Particle Swarm Optimization. Dan kemudian menguji model tersebut menggunakan data pengujian (testing data). Dalam kasus ini, memiliki tiga rasio model data pelatihan dan data pengujian yaitu 70:30,80:20,90:10



**Gambar 3 grafik Akurasi C4.5+PSO**

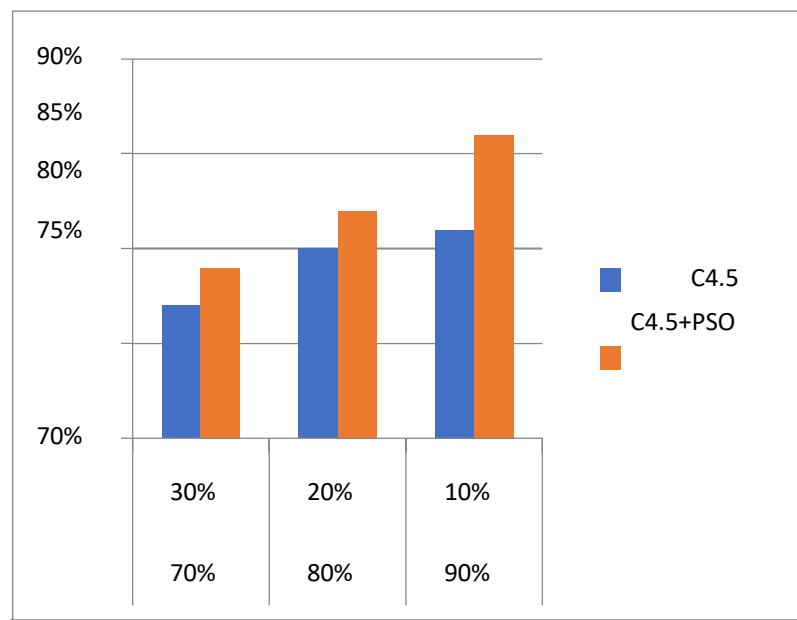
**Perbandingan Metode C4.5 Menggunakan Particle Swarm Optimization**

Optimasi algoritma dengan Particle Swarm Optimization menggunakan data

wholesales dengan jumlah 1000 data. Data dipisahkan menggunakan data training dan testing. dengan presentase berbeda, atribut data jumlah 7 atribut dan satu atribut hasil.

**Tabel.5.2 Hasil perbandingan Akurasi Algoritma C4.5 dengan PSO**

Training	Testing	C4.5	PSO
70%	30%	77%	79%
80%	20%	80%	82%
90%	10%	81%	86%



**Gambar 4 Grafik Hasil Pengujian C4.5 + PSO**

**4. KESIMPULAN**

Setelah melakukan analisis dan pengujian pada klasifikasi data wholesales pada Optimasi C4.dengan Particle Swarm Optimization maka dapat diperoleh kesimpulan sebagai berikut: C4.5 menetapkan data training 70% dan data testing 30% memperoleh hasil akurasi 77%. Dioptimal dengan Particle Swarm Optimization mendapatkan nilai akurasi yang meningkat dengan akurasi 79%. C4.5 dengan data training 80% dan 20% data testing memperoleh hasil akurasi 80%. Dioptimal dengan Particle Swarm Optimization mendapatkan nilai akurasi meningkat dengan akurasi 82%. C4.5 dengan data training 90% dan 10% data testing memperoleh hasil akurasi 81%. Dioptimal dengan Particle Swarm Optimization medapatkan nilai akurasi yang meningkat dengan akurasi 86%.

**Saran**

Untuk saran bagi penelitiselanjutnya guna mengembangkan C4.5 dalam memprediksi :

1. Untuk meningkatkan nilai akurasi dapat menggunakan teknik optimasi lainnya selain Particle Swarm Optimization seperti *Growworm Swarm Optimization (GSO)*.
2. Jumlah data sebaiknya ditingkatkan agar mendapatkan hasil yang maksimal, karena semakin banyak data maka akurasi yang diperoleh semakin tinggi.

## 5. REFERENSI

- [1] Andini, T. I., Witanti, W., & Renaldi, F. (2016). Prediksi Potensi Pemasaran Produk Baru dengan Metode Naïve Bayes Classifier dan Regresi Linear. Seminar Nasional Aplikasi Teknologi Informasi (SNATi), 27–32.
- [2] Junia, A., Pelita, S., & S, I. (2021). Implementasi Algoritma C4 . 5 Untuk Mengukur Tingkat Kepuasan BPJS Ketrnaga Kerjaan. 2(2), 12–15.
- [3] Kennedy, J., & Eberhart, R. (1995). Particle Swarm Optimization, Purdue School of Engineering and Technology. EEE.
- [4] Muslim, M. A., Prasetyo, B., Harum, E. L., Juli, A., Hardiyanti, S. M., & Zurzahputra, A. (2019). Data Mining Algoritma C4.5 Disertai contoh kasus dan penerapannya dengan program computer. Noor, H. (2018). Optimasi Model Klasifikasi C4.5 Dan Particle Swarm Optimization Untuk Prediksi Siswa Bermasalah. *Technologia: Jurnal Ilmiah*, 9(4), 228.
- [5] Saprudin. (2017). Penerapan Particle Swarm Optimization (PSO) untuk Klasifikasi dan Analisis Kredit dengan Menggunakan Algoritma C4.5. *Jurnal Informatika Universitas Pamulang*, 2(4), 214
- [6] Tukino. (2019). Penerapan Algoritma C4 . 5 Untuk Memprediksi Keuntungan. *Jurnal Sistem Informasi Bisnis*, 09(1), 39–46. Wajhillah, R. (2014). Particle Swarm Optimization Untuk Prediksi Penyakit Jantung. 1(1), 26–36
- [7] Wijaya, H. D., & Dwiasnati, S. (2020). Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat. *Jurnal Informatika*, 7(1), 1–7. <https://doi.org/10.31311/ji.v7i1.6203>
- [8] Yulianti, I., Amegia Saputra, R., Sukrisno Mardiyanto, M., & Rahmawati, A. (2020). Accuracy Optimization of C4.5 Algorithm Based on Particle Swarm Optimization with Bagging Technique on Prediction of Chronic Kidney Disease. *Techno.COM*, 19(4), 411–421. <https://archive.ics.uci.edu/ml/>