

OPTIMASI DECISION TREE MENGGUNAKAN PARTICLE SWARM OPTIMIZATION PADA DATA SISWA PUTUS SEKOLAH

Mirza Yogy Kurniawan¹⁾, Muhammad Edya Rosadi²⁾

^{1,2)}Program Studi Teknik Informatika UNISKA

Jl. Adhiyaksa No. 2 Kayu Tangi - Banjarmasin

e-mail: mirza@fti.uniska-bjm.ac.id¹⁾, edya@fti.uniska-bjm.ac.id²⁾

ABSTRAK

Pendidikan merupakan hak setiap warga, bahkan dijadikan program pemerintah menggalakkan wajib belajar 12 tahun. Putus sekolah menjadi hambatan program pemerintah tersebut dimana terjadinya putus sekolah disebabkan banyak faktor, diantaranya faktor ekonomi, kondisi geografis, dan keinginan siswa itu sendiri. ID3 mampu menghasilkan sebuah pohon keputusan dari suatu kumpulan data yang sangat banyak. Pohon keputusan ini bisa dijadikan sebagai acuan prediksi kemungkinan terjadinya putus sekolah bagi siswa. Agar dapat menjadi acuan yang baik maka klasifikasi yang dihasilkan harus memiliki akurasi yang tinggi. PSO dikenal dapat meningkatkan akurasi berbagai macam klasifikasi data mining. ID3 pada penelitian ini menghasilkan akurasi 72,5% sedangkan setelah dioptimasi dengan PSO maka ID3 akan menghasilkan akurasi 85%.

Kata Kunci: Angka Putus Sekolah, ID3, Decision Tree, Particle Swarm Optimization

ABSTRACT

Education is the right of every citizen, even government makes program to promote the compulsory education of 12 years. Drop out of school has become an obstacle to the government program where the dropout is caused by many factors, including economic factors, geographical conditions, and students' own desires. ID3 is able to generate a decision tree from a very large data set. This decision tree can be used as a reference for possible drop out of students. In order to be a good reference then the resulting classification must have a high accuracy. PSO is known to increase the accuracy of various kinds of data mining classification. ID3 in this study yielded 72.5% accuracy while after optimized with PSO then ID3 will yield 85% accuracy..

Keywords: Drop out rate, ID3, Decision Tree, Particle Swarm Optimization.

I. PENDAHULUAN

Pendidikan adalah pembelajaran pengetahuan, keterampilan, dan kebiasaan sekelompok orang yang diturunkan dari satu generasi ke generasi lainnya melalui pengajaran, pelatihan, atau penelitian. Pendidikan menjadi hak setiap warga dan diakui oleh pemerintah, bahkan pemerintah menggalakkan wajib belajar 12 tahun untuk meningkatkan sumber daya manusia. Putus sekolah adalah proses berhentinya siswa secara terpaksa dari suatu lembaga pendidikan tempat belajar. Angka putus sekolah menjadi permasalahan yang saat ini dihadapi oleh pemerintah terutama pada upaya penggalakkan wajib belajar 12 tahun [1]. Hulu Sungai Selatan Angka putus sekolah usia SMA lebih tinggi dibandingkan dengan angka putus sekolah jenjang pendidikan lain. Sebab putus sekolah ada bermacam-macam diantaranya faktor ekonomi dimana orang tua tidak sanggup membiayai anaknya untuk sekolah, bisa juga faktor membantu usaha orang tua, faktor siswanya sendiri, atau faktor kondisi tempat tinggal yang jauh dari sekolah. Data mining adalah ekstraksi pola yang menarik dari data yang besar, dikatakan menarik apabila pola tersebut tidak sepele, implisit, dan tidak diketahui sebelumnya [2]. Decision Tree adalah pohon terstruktur dari sekumpulan atribut yang dibuat dengan tujuan untuk meramalkan outputnya [3]. ID3 dan C4.5 adalah bagian dari metode klasifikasi dalam data mining yang dapat menghasilkan suatu pohon keputusan dari suatu kumpulan data sehingga diketahui faktor-faktor yang paling mempengaruhi dalam suatu data. ID3 memiliki kinerja (precision, recall, dan accuracy) yang lebih baik dibandingkan C4.5 [4]. ID3 yang diterapkan pada data siswa akan menghasilkan tree yang bisa digunakan untuk membuat perkiraan kemungkinan putus sekolah atau tidaknya seorang siswa. Namun diperlukan tree yang baik dengan akurasi yang tinggi untuk dapat digunakan sebagai prediksi. Untuk meningkatkan akurasinya decision tree bisa dioptimasi dengan metode Particle Swarm Optimization (PSO) untuk menyeleksi atributnya [5]. Berdasarkan latar belakang yang sudah dipaparkan diatas, maka rumusan masalahnya adalah bagaimana menghasilkan suatu tree dengan akurasi yang baik dengan ID3 yang diterapkan pada data siswa. Penelitian ini bertujuan untuk meningkatkan akurasi dari klasifikasi ID3 yang diterapkan pada data siswa. Hasil penelitian ini diharapkan mampu memberikan manfaat bagi sekolah untuk dapat melakukan tindakan-tindakan preventif kepada siswa yang terprediksi putus sekolah. Dalam bidang keilmuan penelitian ini diharapkan berkontribusi pada perkembangan keilmuan tentang data mining pada umumnya dan pada klasifikasi ID3 pada khususnya.

II. LANDASAN TEORI

A. Data Mining

Data mining adalah proses ekstraksi informasi yang tersembunyi, belum pernah diketahui dan berpotensi untuk berguna dari data yang banyak [6]. Pembangunan aplikasi komputer *data mining* dapat membantu manusia untuk menemukan pola dan kebiasaan pada data. Makin kuat polanya maka akan makin tepat prediksinya, hanya saja pola yang kuat biasanya adalah sesuatu yang tidak menarik karena sudah diketahui. Namun dengan data mining bisa ditemukan pola yang belum pernah diketahui sebelumnya [2].

B. ID3

Decision tree adalah salah satu metode dalam data mining yang digunakan untuk melakukan klasifikasi terhadap data. Metode ini menjadi populer karena mudah diinterpretasi oleh manusia [3]. Pada dasarnya decision tree adalah metode untuk mengubah data menjadi sebuah pohon keputusan dengan aturan keputusan. Proses pembentukannya dimulai dari data pada simpul akar yang dilanjutkan dengan pemilihan atribut yang diformulasikan logical test pada atribut tersebut

Algoritma ID3 atau Iterative Dichotomiser 3 (ID3) adalah sebuah model data mining untuk membuat pohon keputusan berdasarkan data yang sudah dimiliki [7]. Model ini menggunakan konsep dari entropi informasi, berikut langkah kerjanya :

1. Information gain dihitung dari setiap atribut dengan menggunakan

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{Nilai}(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

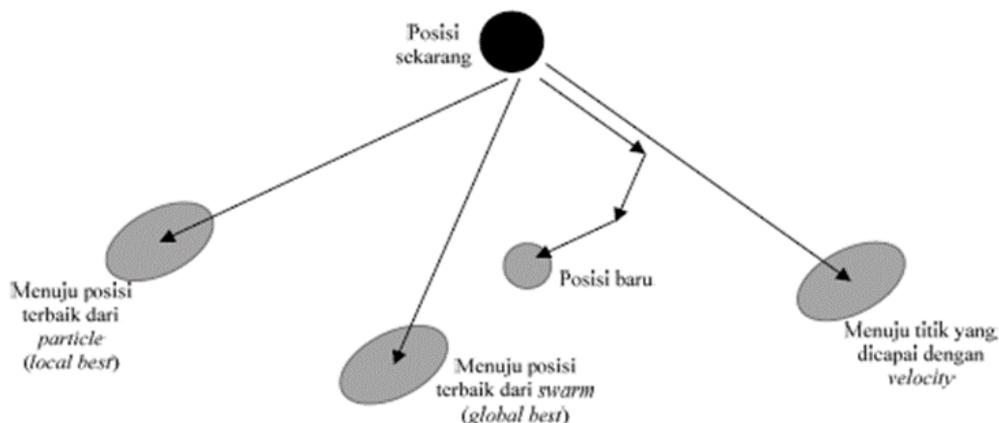
Dimana

$$Entropy(S) = -P_+ \log_2 P_+ - -P_- \log_2 P_-$$

2. Pemilihan atribut yang memiliki nilai information gain terbesar
3. Pembentukan simpul yang berisi atribut tersebut
4. Ulangi proses perhitungan information gain akan terus dilakukan sampai semua data telah termasuk dalam kelas yang sama. Atribut yang telah dipilih tidak diikuti lagi dalam perhitungan nilai information gain

C. PSO

Particle swarm optimization, disingkat sebagai PSO, didasarkan pada perilaku sebuah kawanan serangga, seperti semut, rayap, lebah atau burung. Algoritma PSO meniru perilaku sosial organisme ini. Perilaku sosial terdiri dari tindakan individu dan pengaruh dari individu-individu lain dalam suatu kelompok. Pada dasarnya partikel tidak dapat berpindah secara tiba-tiba, dan bergerak menuju posisi terbaik berdasarkan pengalaman pribadinya maupun pengalaman dari seluruh partikel. Prinsip perpindahannya dapat diilustrasikan pada gambar berikut



Gambar 1 Prinsip perpindahan partikel

Kata partikel menunjukkan, misalnya, seekor burung dalam kawanan burung. Setiap individu atau partikel berperilaku secara terdistribusi dengan cara menggunakan kecerdasannya (intelligence) sendiri dan juga dipengaruhi perilaku kelompok kolektifnya. Dengan demikian, jika satu partikel atau seekor burung menemukan jalan yang tepat atau pendek menuju ke sumber makanan, sisa kelompok yang lain juga akan dapat segera

mengikuti jalan tersebut meskipun lokasi mereka jauh di kelompok tersebut. Metode optimasi yang didasarkan pada swarm intelligence ini disebut algoritma behaviorally inspired sebagai alternatif dari algoritma genetika, yang sering disebut evolution-based procedures. Algoritma PSO ini awalnya diusulkan oleh J. Kennedy and R.C. Eberhart. Dalam konteks optimasi multivariabel, kawanan diasumsikan mempunyai ukuran tertentu atau tetap dengan setiap partikel posisi awalnya terletak di suatu lokasi yang acak dalam ruang multidimensi. Setiap partikel diasumsikan memiliki dua karakteristik: posisi dan kecepatan.

Prosedur Algoritma PSO:

1. Mulai menentukan parameter-parameter
2. Tentukan k_{max} , v_{max} , w_k , p , c_1 dan c_2
3. Inisialisasi posisi dan velocity setiap particle secara random
4. Jika posisi dan velocity setiap particle dinilai layak maka hitung fungsi tujuan setiap particle, apabila tidak layak maka inisialisasikan kembali posisi dan velocity tiap particle
5. Hitung fungsi tujuan setiap particle
6. Perbaharui nilai fungsi tujuan terbaik untuk setiap particle dan nilai terbaik swarm.
7. Perbaharui velocity setiap particle
8. Perbaharui posisi setiap particle
9. Layak atau tidak; apabila tidak layak maka perbaharui kembali velocity dan posisi setiap particle, apabila layak perbaharui nilai w_k
10. Perbaharui nilai w_k
11. Stopping criteria terpenuhi; apabila tidak maka hitung kembali fungsi tujuan setiap particle apabila terpenuhi artinya selesai.

D. Validasi (K-Fold Cross Validation)

Cross validasi adalah metode statistik untuk membandingkan algoritma pembelajaran dan mengevaluasinya. Prosesnya data dibagi menjadi dua bagian yaitu data latih dan data validasi. Data latih dan data validasi harus cross-over secara berturut-turut sebanyak K kali, misalkan pada 10 fold cross validation maka data dibagi menjadi 10 bagian dan dilakukan 10 kali iterasi untuk pengujiannya [2].

E. Confusion Matrix

Confusion matrix adalah metode yang digunakan untuk mengukur kinerja model klasifikasi berdasarkan perhitungan objek testing, dimana data hasil prediksi ada diantara dua kelas yaitu menghasilkan kelas positif dan kelas negative. Untuk proses evaluasi dengan confusion matrix maka akan diperoleh nilai precision, recall, dan accuracy yang didapat dari rumus [7]:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Dimana,

TP : jumlah kasus positif yang diklasifikasikan sebagai positif

FP : jumlah kasus negatif yang diklasifikasikan sebagai positif

TN : jumlah kasus negatif yang diklasifikasikan sebagai negatif

FN : jumlah kasus positif yang diklasifikasikan sebagai negatif

III. METODE PENELITIAN

Metode penelitian yang dilakukan adalah metode penelitian eksperimen, dengan tahapan penelitian seperti berikut:

A. Pengumpulan Data (Data Gathering)

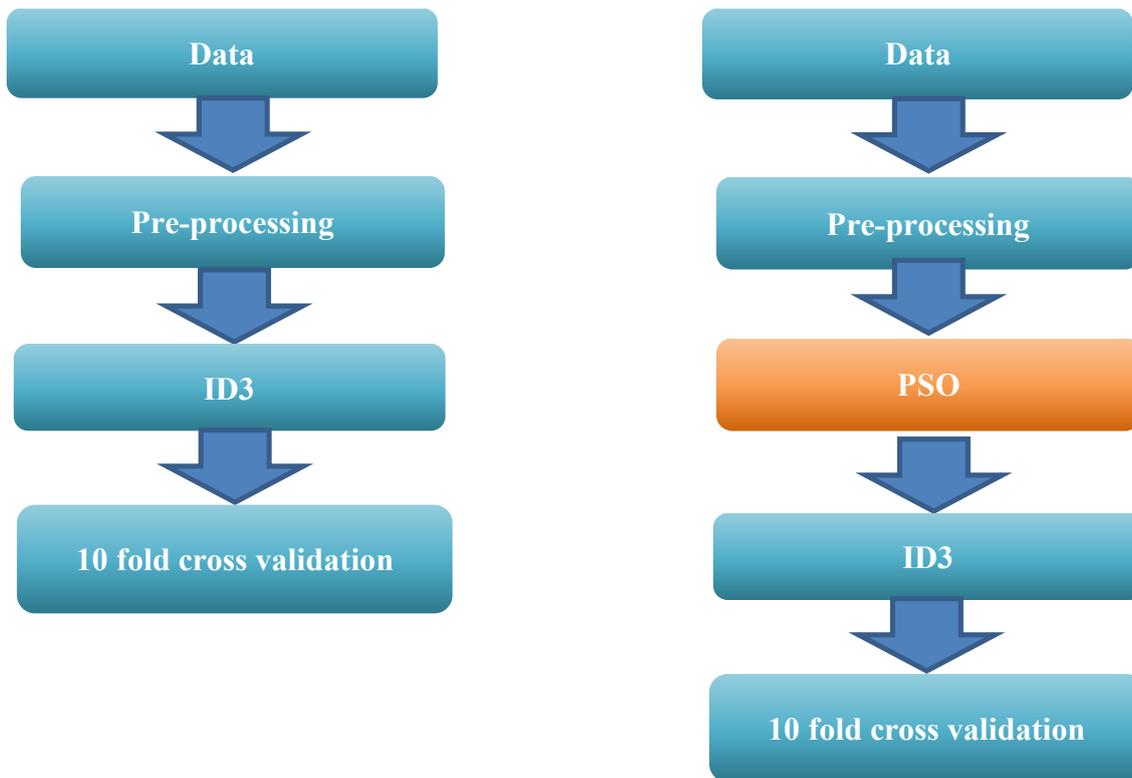
Data yang digunakan adalah data siswa kelas XII pada SMAN 2 Barabai, Hulu Sungai Tengah yang berjumlah 40 orang dimana 13 orang diantaranya putus sekolah. Data kemudian direkapitulasi kedalam bentuk file excel.

B. Pengolahan Data Awal (Data Pre-processing)

Data numerik di beberapa paramater dikonversi menjadi nominal agar dapat diproses dengan model ID3. Proses konversi dilakukan dengan mengelompokkan nilai pada batasan tertentu sehingga bisa dikelompokkan dalam bentuk nominal

C. Desain Eksperimen dan Pengujian Metode (Method Test and Experiment)

Eksperimen dilakukan dengan menerapkan ID3 pada data yang sudah melalui pre-processing dan diuji tingkat akurasi klasifikasinya dengan *confusion matrix*. Kemudian diterapkan pada PSO dan diuji tingkat akurasi untuk dibandingkan dengan hasil sebelumnya yang tanpa menggunakan PSO.



Gambar 2 Desain Eksperimen

D. Analisa Hasil Eksperimen

Pengukuran tingkat akurasi dilakukan dengan menghitung *Precision* dan *Recall* dari *tree* yang dihasilkan, baik ID3 tanpa PSO juga ID3 dengan PSO untuk dianalisa peningkatan akurasi. Semakin tinggi nilai perbaikannya menyatakan semakin baiknya klasifikasi dan mendekati dengan solusi ideal.

IV. HASIL DAN PEMBAHASAN

Rapidminer akan dijalankan pada dua kondisi yang berbeda, kondisi pertama data excel diproses dengan model ID3 dan kemudian diukur performanya dengan metode X-validation, kondisi kedua sama dengan kondisi pertama namun ditambahkan optimasi PSO kemudian diukur dengan X-validation dan dibandingkan dengan hasil pada kondisi pertama.

Tabel 1 Data Siswa

ID	JK	Jarak Rumah-Sekolah (Km)	Pekerjaan Ortu	Absensi	Nilai	Keterangan
0001	L	7	PETANI	0	67	Tidak Putus
0002	P	2	PETANI	0	70	Tidak Putus
0003	L	2	PETANI	0	73	Putus
0004	L	5.5	GURU	0	68	Tidak Putus

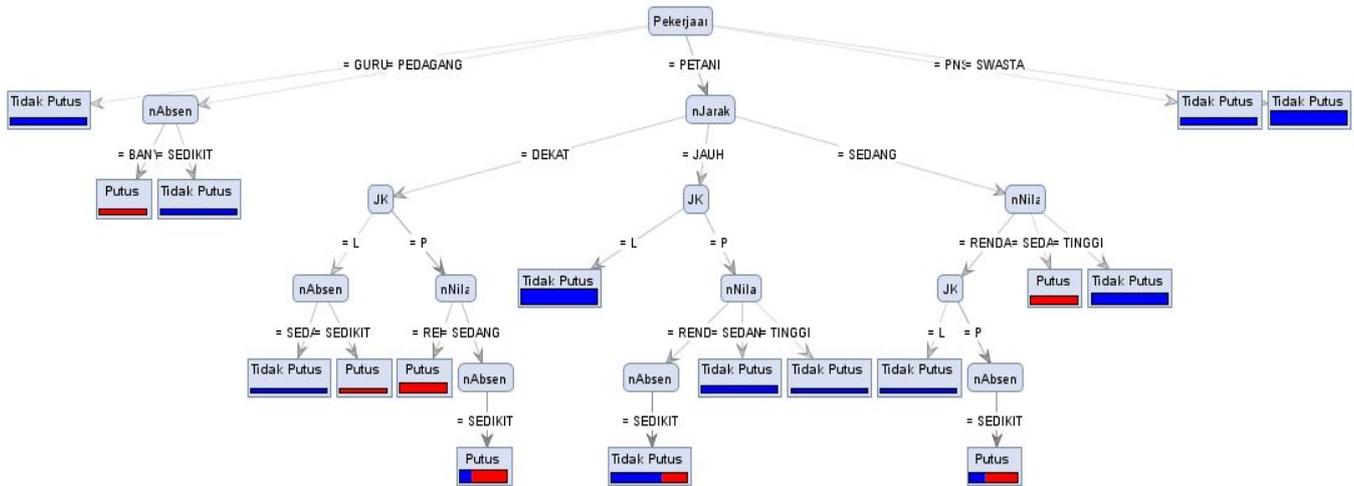
0005	P	4.5	PEDAGANG	3	70	Putus
0006	P	2.5	GURU	0	80	Tidak Putus
0007	P	2.5	PETANI	0	60	Putus
0008	L	12	PETANI	0	76	Tidak Putus
0009	L	3	SWASTA	0	71	Tidak Putus
0010	P	4.5	PETANI	0	64	Putus
0011	P	12	PETANI	0	68	Tidak Putus
0012	P	1.5	PETANI	0	70	Putus
0013	P	2.5	PETANI	1	66	Putus
0014	L	2	PETANI	1	65	Tidak Putus
0015	P	5.5	PETANI	2	73	Putus
0016	P	20	PETANI	0	78	Tidak Putus
0017	P	1.5	PNS	0	85	Tidak Putus
0018	P	4.5	PETANI	0	67	Putus
0019	P	2.5	PETANI	0	74	Putus
0020	L	7	PETANI	0	70	Tidak Putus
0021	P	9.5	PETANI	0	66	Putus
0022	P	5.5	PETANI	0	65	Tidak Putus
0023	P	12	PETANI	4	73	Tidak Putus
0024	L	4	PETANI	0	60	Tidak Putus
0025	L	1.5	SWASTA	0	60	Tidak Putus
0026	P	5	PETANI	0	84	Tidak Putus
0027	P	8	PETANI	0	84	Tidak Putus
0028	P	8.5	SWASTA	0	69	Tidak Putus
0029	P	8.5	SWASTA	1	79	Tidak Putus
0030	L	6	PETANI	0	75	Tidak Putus
0031	P	5	PETANI	0	84	Tidak Putus
0032	P	5.5	PEDAGANG	0	72	Tidak Putus
0033	P	2.5	PETANI	0	78	Putus
0034	P	1	PNS	0	69	Tidak Putus
0035	P	1.5	PETANI	2	63	Putus
0036	P	5.5	SWASTA	0	63	Tidak Putus
0037	P	4.5	PETANI	0	72	Putus
0038	P	3.5	PETANI	0	87	Tidak Putus
0039	P	11	PETANI	0	66	Tidak Putus
0040	L	7	PETANI	0	68	Tidak Putus

Tabel 2 Data Siswa setelah preprocessing

ID	JK	Jarak Ru- mah-Sekolah	Pekerjaan Ortu	Absensi	Nilai	Keterangan
0001	L	JAUH	PETANI	SEDIKIT	RENDAH	Tidak Putus
0002	P	DEKAT	PETANI	SEDIKIT	SEDANG	Tidak Putus
0003	L	DEKAT	PETANI	SEDIKIT	SEDANG	Putus
0004	L	SEDANG	GURU	SEDIKIT	RENDAH	Tidak Putus
0005	P	SEDANG	PEDAGANG	BANYAK	SEDANG	Putus

0006	P	DEKAT	GURU	SEDIKIT	TINGGI	Tidak Putus
0007	P	DEKAT	PETANI	SEDIKIT	RENDAH	Putus
0008	L	JAUH	PETANI	SEDIKIT	SEDANG	Tidak Putus
0009	L	SEDANG	SWASTA	SEDIKIT	SEDANG	Tidak Putus
0010	P	SEDANG	PETANI	SEDIKIT	RENDAH	Putus
0011	P	JAUH	PETANI	SEDIKIT	RENDAH	Tidak Putus
0012	P	DEKAT	PETANI	SEDIKIT	SEDANG	Putus
0013	P	DEKAT	PETANI	SEDANG	RENDAH	Putus
0014	L	DEKAT	PETANI	SEDANG	RENDAH	Tidak Putus
0015	P	SEDANG	PETANI	SEDANG	SEDANG	Putus
0016	P	JAUH	PETANI	SEDIKIT	SEDANG	Tidak Putus
0017	P	DEKAT	PNS	SEDIKIT	TINGGI	Tidak Putus
0018	P	SEDANG	PETANI	SEDIKIT	RENDAH	Putus
0019	P	DEKAT	PETANI	SEDIKIT	SEDANG	Putus
0020	L	JAUH	PETANI	SEDIKIT	SEDANG	Tidak Putus
0021	P	JAUH	PETANI	SEDIKIT	RENDAH	Putus
0022	P	SEDANG	PETANI	SEDIKIT	RENDAH	Tidak Putus
0023	P	JAUH	PETANI	BANYAK	SEDANG	Tidak Putus
0024	L	SEDANG	PETANI	SEDIKIT	RENDAH	Tidak Putus
0025	L	DEKAT	SWASTA	SEDIKIT	RENDAH	Tidak Putus
0026	P	SEDANG	PETANI	SEDIKIT	TINGGI	Tidak Putus
0027	P	JAUH	PETANI	SEDIKIT	TINGGI	Tidak Putus
0028	P	JAUH	SWASTA	SEDIKIT	RENDAH	Tidak Putus
0029	P	JAUH	SWASTA	SEDANG	SEDANG	Tidak Putus
0030	L	JAUH	PETANI	SEDIKIT	SEDANG	Tidak Putus
0031	P	SEDANG	PETANI	SEDIKIT	TINGGI	Tidak Putus
0032	P	SEDANG	PEDAGANG	SEDIKIT	SEDANG	Tidak Putus
0033	P	DEKAT	PETANI	SEDIKIT	SEDANG	Putus
0034	P	DEKAT	PNS	SEDIKIT	RENDAH	Tidak Putus
0035	P	DEKAT	PETANI	SEDANG	RENDAH	Putus
0036	P	SEDANG	SWASTA	SEDIKIT	RENDAH	Tidak Putus
0037	P	SEDANG	PETANI	SEDIKIT	SEDANG	Putus
0038	P	SEDANG	PETANI	SEDIKIT	TINGGI	Tidak Putus
0039	P	JAUH	PETANI	SEDIKIT	RENDAH	Tidak Putus
0040	L	JAUH	PETANI	SEDIKIT	RENDAH	Tidak Putus

Model ID3 menghasilkan tree yang akan diukur performanya terhadap data berikut gambar tree dari hasil permodelan ID3 terhadap data siswa untuk analisa putus sekolah.



Gambar 3 Tree Hasil ID3

Dari pohon keputusan tersebut diperoleh aturan-aturan sebagai berikut

- 1) if Pekerjaan= GURU then class = Tidak Putus
- 2) if Pekerjaan = PEDAGANG and nAbsen = BANYAK then class = Putus
- 3) if Pekerjaan = PEDAGANG and nAbsen = SEDIKIT then class = Tidak Putus
- 4) if Pekerjaan = PETANI and nJarak=DEKAT and JK = L and nAbsen = SEDANG then class = Tidak Putus
- 5) if Pekerjaan = PETANI and nJarak=DEKAT and JK = L and nAbsen = SEDIKIT then class = Putus
- 6) if Pekerjaan = PETANI and nJarak=DEKAT and JK = P and nNilai = RENDAH then class = Putus
- 7) if Pekerjaan = PETANI and nJarak=DEKAT and JK = P and nNilai = SEDANG and nAbsen = SEDIKIT then class = Putus
- 8) if Pekerjaan = PETANI and nJarak=JAUH and JK = L then class = TidakPutus
- 9) if Pekerjaan = PETANI and nJarak=JAUH and JK = P and nNilai = RENDAH and nAbsen = SEDIKIT then class = TidakPutus
- 10) if Pekerjaan = PETANI and nJarak=JAUH and JK = P and nNilai = SEDANG then class = Tidak Putus
- 11) if Pekerjaan = PETANI and nJarak=JAUH and JK = P and nNilai = TINGGI then class = Tidak Putus
- 12) if Pekerjaan = PETANI and nJarak=SEDANG and nNilai = RENDAH and JK = L then class = Tidak Putus
- 13) if Pekerjaan = PETANI and nJarak=SEDANG and nNilai = RENDAH and JK = P and nAbsen = SEDIKIT then class = Putus
- 14) if Pekerjaan = PETANI and nJarak=SEDANG and nNilai = SEDANG then class = Putus
- 15) if Pekerjaan = PETANI and nJarak=SEDANG and nNilai = TINGGI then class = Tidak Putus
- 16) if Pekerjaan = PNS then class = Tidak Putus
- 17) if Pekerjaan = PNS then class = Tidak Putus

Metode pengukuran X-validation menghasilkan *confusion matrix* yang menunjukkan precision, recall, hasil true positive, true negative, false positive, dan false negative.

accuracy: 72.50% +/- 13.46% (mikro: 72.50%)			
	true Tidak Putus	true Putus	class precision
pred. Tidak Putus	21	5	80.77%
pred. Putus	6	8	57.14%
class recall	77.78%	61.54%	

Gambar 4 Confusion Matrix ID3

Kondisi kedua dimana ID3 dioptimasi dengan PSO menghasilkan *confusion matrix* sebagai berikut.

accuracy: 85.00% +/- 16.58% (mikro: 85.00%)			
	true Tidak Putus	true Putus	class precision
pred. Tidak Putus	24	3	88.89%
pred. Putus	3	10	76.92%
class recall	88.89%	76.92%	

Gambar 5 *Confusion Matrix* ID3 dengan PSO

Data siswa putus sekolah diproses dengan permodelan ID3 menghasilkan tree, performa akurasi, *precision* dan *recall* sedangkan ID3 dengan PSO menghasilkan performa akurasi, *precision*, *recall* dan bobot PSO. Tree yang dihasilkan menunjukkan pekerjaan orang tua memiliki pengaruh paling utama dimana jika pekerjaan orang tua Guru, PNS, dan Swasta maka siswa tidak putus sekolah. Sedangkan jika Pedagang maka dilihat jumlah tidak hadirnya, jika Banyak maka putus sekolah. Jika pekerjaan orang tua adalah Petani maka parameter lain dilihat dengan berbagai macam ketentuan. Berikut perbandingan hasil ID3 dengan ID3 yang dioptimasi dengan PSO

Tabel 3 Perbandingan Performa ID3

	ID3	ID3 dengan PSO
<i>Accuracy</i>	72.5	85
<i>Precision</i>	57.14	76.92
<i>Recall</i>	65	75

Dari tabel 4.1 dapat dilihat PSO dapat meningkatkan performa akurasi ID3 pada data siswa putus sekolah dari 72.5% menjadi 85%

V. KESIMPULAN

ID3 pada penelitian ini menghasilkan akurasi 72,5% sedangkan setelah dioptimasi dengan PSO maka ID3 akan menghasilkan akurasi 85%. Dapat dilihat bahwa PSO dapat meningkatkan akurasi dari klasifikasi ID3 yang diterapkan pada data siswa putus sekolah. Untuk pengembangannya ID3 bisa diterapkan pada data siswa jenjang lain, atau ditambahkan parameter perhitungannya, atau dibuatkan aplikasi yang *user friendly* untuk bisa mudah dipakai oleh pihak sekolah.

VI. DAFTAR PUSTAKA

- [1] Purbasari and D. Arlintha, "Permodelan Angka Putus Sekolah Tingkat SLTP dan Sederajat di Jawa Timur Tahun 2012 Menggunakan Metode Analisis Regresi Logistik Ordinal," Institut Teknologi Sepuluh Nopember Surabaya, Surabaya, 2014.
- [2] I. H. Witten, Data Mining Practical Machine Learning Tools and Techniques Third Edition, Amsterdam: Morgan Kauffman Press, 2011.
- [3] R. A. Giovani, "Sistem Pendukung Keputusan Prediksi Kecepatan Studi Mahasiswa Menggunakan Metode ID3," Universitas Atmajaya Yogyakarta, Yogyakarta, 2011.
- [4] S. Defiyanti and D. L. C. Pardede, "Perbandingan Kinerja Algoritma ID3 dan C4.5 dalam Klasifikasi Spam-Mail," Jurusan Sistem Informasi Universitas Gunadarma, Jakarta, 2010.
- [5] L. S. Ramdhani, "Penerapan Particle Swarm Optimization (PSO) untuk Seleksi Atribut dalam Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis dengan Metode Algoritma C4.5," *Swabumi*, vol. IV, no. 1, pp. 1-15, 2016.
- [6] E. Prasetyo, Data Mining - Konsep dan Aplikasi Menggunakan Matlab, Gresik: Andi Offset, 2012.
- [7] N. A. Kartika, "Penerapan Teknik Data Mining Menggunakan Algoritma Iterative Dichotomizer (ID3) untuk Menentukan Posisi Siswa Dalam Dunia Kerja Berdasarkan Kecerdasan Emosi," Universitas Dian Nuswantoro, Semarang, 2014.