 ISSN:2302-8149
e-ISSN:2540-9719

Optimasi *Decision Tree* Menggunakan *Particle Swarm Optimization* untuk Klasifikasi Kesuburan pada Pria

Tyas Widyani Pratiwi*, Toni Arifin
Sistem Informasi, Fakultas Teknologi Informasi, Universitas Adhirajasa Reswara Sanjaya,
Jalan Sekolah Internasional No 1-2 Antapani, Bandung.
*e-mail: tyaswidyanii@gmail.com

(***received:*** 4 Agustus 2020, ***revised:*** 25 September 2020, ***accepted:*** 23 Oktober 2020)

Abstrak

Keturunan adalah hal yang sangat diharapkan pada setiap pasangan suami istri, maka dari itu tingkat kesuburan pada pria adalah salah satu faktor penting. Faktor yang pempengaruhi tingkat kesuburan itu sendiri seperti hormon, penyakit bawaan, pernah atau tidaknya dioperasi. Salah satu cara untuk menganalisis kesuburan pada pria dapat dilakukan dengan teknik *data mining*. *Data Mining* dapat digunakan untuk mengklasifikasi ataupun prediksi. Penelitian ini, bertujuan untuk mendapatkan metode klasifikasi terbaik yang dapat menghasilkan tingkat dari nilai akurasi yang tinggi bila dikombinasikan dengan metode *Particle Swarm Optimization (*PSO). Pada penelitian ini, algoritma *Decision Tree* dengan *Particle Swarm* Optimization meningkatkan nilai akurasi sebesar 93.33% dan nilai AUC sebesar 0,793 termasuk kedalam kategori *Fair classification.*

Kata Kunci: *data mining, decision tree,* kesuburan, *particle swarm optimization*

Abstract

Heredity is something that is expected in every married couple, therefore the level of fertility in men is one important factor. Factors that influence the level of fertility itself such as hormones, congenital diseases, surgery or not. One way to analyze fertility in men can be done with data mining techniques. Data Mining can be used to classify or predict. This study aims to obtain the best classification method that can produce a high level of accuracy when combined with the Particle Swarm Optimization (PSO) method. In this study, the Decision Tree algorithm with Particle Swarm Optimization increased the accuracy value by 93.33% and the AUC value by 0.793 was included in the Fair classification category.

Keywords: *data mining, decision tree, fertility, particle swarm optimization*

1 Pendahuluan

Keturunan adalah hal yang sangat diharapkan pada setiap pasangan suami istri, maka dari itu tingkat kesuburan pada pria adalah salah satu faktor penting [1]. Masa subur pria secara umum sangat dipengaruhi oleh usia yang dimiliki pria tersebut. Jika pria memasuki usia 40 tahun maka kemampuan sperma dalam membuahi sel telur akan menurun dengan drastis, sedangkan kualitas sperma bisa berada dalam kondisi yang sangat buruk ketika pria menginjak usia 55 tahun. Adapun alasan usia 55 tahun sebagai kondisi terburuk bagi pria, karena semakin tua usia pria akan semakin melemah pergerakan spermanya, hal ini menyebabkan sperma tersebut sulit mencapai sel telur, oleh karena itu kesuburan pada pria menjadi sebuah masalah di beberapa dekade ini dalam bidang kesehatan [2].

Di Indonesia sendiri angka tingkat *infertilitas* pasangan suami-istri yang mengalami kesulitan untuk mendapatkan keturunan sekitar 10%. Dari hasil riset yang dilakukan oleh WHO mendapatkan 50% penyebab *infertilitas* adalah pihak pria dan sisanya dipengaruhi oleh faktor semen (*sperma*) [3]. Faktor yang mempengaruhi tingkat kesuburan itu sendiri seperti hormon, penyakit bawaan, pernah atau tidaknya dioperasi. Adapun faktor lain yang mempengaruhi tingkat kesuburan adalah gaya hidup yang tidak sehat seperti merokok, mengkonsumsi minuman beralkohol, *obesitas*, dan duduk terlalu lama [1], [2].

Health Day News melakukan sebuah penelitian bahwa merokok dapat merusak sistem reproduksi seseorang dan mengurangi peluang untuk memiliki anak yang sehat. Merokok juga dapat menurunkan

 ISSN:2302-8149
e-ISSN:2540-9719

kualitas dan kuantitas *spermatozoa* manusia (konsentrasi, *motilitas*, dan *morfologi spermatozoa*) [4]. Alkohol juga dapat mempengaruhi fungsi liver, yang dapat menyebabkan peningkatan *estrogen* dalam tubuh sehingga akan memepengaruhi produksi *sperma* [5].

Analisis untuk tingkat kesuburan pada pria sangat penting untuk evaluasi potensi [6]. Salah satu cara untuk menganalisis kesuburan pada pria dapat dilakukan dengan teknik *data mining* [7]. *Data Mining* adalah tahapan proses untuk menggali nilai tambah berisikan informasi yang selama ini tidak dapat diketahui secara manual dari suatu data [8]. *Data Mining* merupakan penyaringan data dengan memanfaatkan kumpulan data yang ukurannya cukup besar melalui tahapan proses untuk mendapatkan informasi dari data tersebut [9].

Data Mining dapat digunakan untuk mengklasifikasi ataupun prediksi [7]. Dalam klasifikasi sendiri terdapat target variabel kategori. Metode-metode atau model-model yang telah dikembangkan oleh peneliti untuk menyelesaikan kasus klasifikasi [10]. Klasifikasi pada *data mining* untuk memprediksi label *class* dan untuk mengklasifikasi data didasarkan pada *data training* dan nilai label *class* dalam mengklasifikasikan *attribute* dan data baru [11].

Salah satu algoritma yang dapat digunakan pada *data mining* klasifikasi ini adalah *Decision Tree*. *Decision Tree* adalah pohon terstruktur dari sekumpulan *attribute* yang akan diuji dengan tujuan meramalkan output-nya [12]. *Decision tree* adalah struktur *flowchart* yang mempunyai pohon (*tree*), dimana setiap simpul internal menandakan suatu tes *attribute*, cabang menjelaskan hasil tes, dan simpul daun menjelaskan *class* atau distribusi kelas, alur pada *decision tree* ditelusuri dari simpul ke akar simpul daun yang memegang prediksi *class* [1]. Dan juga pada metode ini sangat populer karena dapat melakukan klasifikasi sekaligus menunjukkan hubungan antar *attribute* [13].

Algoritma *Decision Tree* memiliki kelebihan, diantaranya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang satu dengan cepat dibandingkan dengan algoritma lain [14]. Adapun kelemahan yang dimiliki oleh algoritma *Decision Tree* ini termasuk dalam kelompok teknik klasifikasi tunggal. Klasifikasi tunggal ini sangat rentan pada data yang tidak seimbang, dikarenakan pada teknik klasifikasi tunggal lebih mengutamakan *class* mayoritas dan cenderung mengabaikan informasi pada *class* minoritas [15].

Algoritma *Decision Tree* termasuk algoritma yang lemah, karena algoritma *decision tree* bergantung pada pembuatan desain dari pohon tersebut. Apabila desain pohon tersebut tidak terstruktur, misalnya ada satu atribut yang tidak sesuai maka kualitas pohon keputusan tersebut akan lemah dan buruk dan akan terjadi *error* dalam pengakumulasian dari setiap tingkat pada sebuah pohon keputusan yang besar. Oleh karena itu optimasi atau peningkatan akurasi perlu dilakukan [16] . Untuk memaksimalkan hasil klasifikasi yang akan diperoleh, dengan cara meningkatkan nilai akurasinya menggunakan metode *Particle Swarm Optimization* (PSO) [17].

Particle Swarm Optimization (PSO) dikemukakan pertama kali oleh James Kennedy dan Russell C. Eberhart pada tahun 1995. Metode optimasi ini terinspirasi dari perilaku gerak hewan seperti ikan dan burung dalam mencari mangsa [18]. *Particle Swarm Optimimization* (PSO) merupakan algoritma pencarian berbasis populasi dan diinisialisasi dengan populasi solusi acak yang biasa disebut partikel [19]. Hal ini efisien untuk memecahkan masalah optimasi [20].

Pada penelitian ini, penulis akan melakukan optimasi *Decision Tree* menggunakan *Particle Swarm Optimization* (PSO) untuk klasifikasi kesuburan pria dengan *dataset* yang digunakan yaitu *Fertility Dataset* yang diambil dari UCI *Repository*. Atribut pada *fertility dataset* yang akan dioptimasi diantaranya *Season in which the analysis was performed, Age, Childish diseases, Accident or serious trauma, Surgical intervention, High fevers in the last year, Frequency of alcohol consumption, Smoking habit, Number of hours spent sitting per day*, yang akan dioptimasi pada atribut tersebut yaitu bobot atributnya. Adapun tujuan dari penelitian ini adalah untuk mengetahui nilai akurasi dari algoritma *Decision Tree* dan mengetahui peningkatan performa dari algoritma *decision tree* setelah dilakukan optimasi menggunakan *Particle Swarm Optimization* dalam klasifikasi kesuburan pada pria.

2 Tinjauan Literatur

Penelitian tentang klasifikasi kesuburan pria dengan metode yang beragam telah banyak dilakukan oleh para peneliti terdahulu. Seperti dalam penelitian [21], [1], [22], [23], [24]. Pada

 ISSN:2302-8149
e-ISSN:2540-9719

penelitian Budianita et al., [21] melakukan penelitian tentang kesuburan pria menggunakan metode *Hopfield*, dan menghasilkan nilai akurasi sebesar 88,51% dengan vektor inisialisasi N (0,1,0,0,1,-1) dan O (1,1,1,0,0,-1). Penelitian Amrulloh & Wibowo, [1] melakukan penelitian tentang kesuburan pria menggunakan metode *Decision Tree* menghasilkan nilai akurasi sebesar 92%.

Penelitian Nurelasari [22] melakukan penelitian tentang kesuburan pria untuk menguji kemampuan antara algoritma *Naive Bayes* dengan *Support Vector Machine* berbasis *Particle Swarm Optimization.* Untuk mengukur tingkat akurasi dan AUC (*Area Under Cureve*) dari penentuan prediksi kesuburan menggunakan *cross validation.* Hasil dari komparasi kedua data tersebut menghasilkan nilai akurasi untuk metode *Support Vector Machine* berbasis *Particle Swarm Optimization* memperoleh 88.00% dengan AUC sebesar 0.49 sedangkan menggunakan metode *Naive Bayes* akurasi yang diperoleh sebesar 85.00% dengan nilai AUC sebesar 0.52.

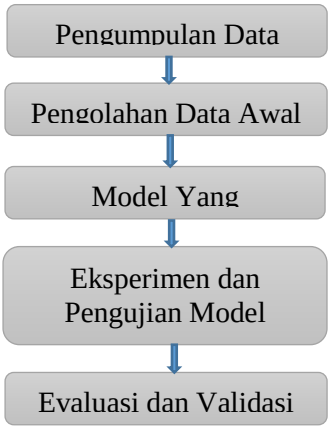
Penelitian Siswa & Prihandoko [23] memiliki tujuan untuk mengevaluasi perbandingan penerapan optimasi terhadap metode klasifikasi data mining algoritma C4.5 dan *Naïve Bayes* yang berbasis *Particle Swarm Optimization* (PSO) untuk mendeteksi kanker payudara menggunakan pengukuran *confusion matrix*, AUC dan t-*Test*. Hasil yang diperoleh dari penelitian ini menyatakan bahwa optimasi *Particle Swarm Optimization* (PSO) dapat meningkatkan nilai akurasi C4.5 dari 90,19% menjadi 94,29% dan *Naive Bayes* 97,65% menjadi 97,96%.

Penelitian Rifai & Aulianita [24] adalah perbandingan dua algoritma yaitu algoritma C4.5 yang dikombinasikan dengan algoritma PSO dan algoritma *Naive Bayes* yang dikombinasikan dengan algoritma PSO untuk menentukan resiko kredit menggunakan data nasabah koperasi. Hasil dari penelitian ini menunjukan algoritma C4.5 yang dikombinasikan dengan algoritma PSO memiliki nilai akurasi sebesar 87.61%, AUC sebesar 0.860, *recall* sebesar 96.32% dan *precision* sebesar 88.96% sedangkan menggunakan algoritma *Naive Bayes* yang dikombinasikan dengan algoritma PSO memiliki nilai akurasi sebesar 87.31%, AUC sebesar 0.851, *recall* sebesar 96.75 % dan *precision* sebesar 88.56%.

Pada penelitian ini, penulis melakukan optimasi *Decision Tree* menggunakan *Particle Swarm Optimization* (PSO) untuk klasifikasi kesuburan pria dengan *dataset* yang digunakan yaitu *Fertility Dataset.* Yang membedakan dengan penelitian sebelumnya di atas yaitu penggunaan metode yang berbeda pada *dataset* yang sama dan juga metode yang sama pada *dataset* yang berbeda.

3 Metode Penelitian

Menurut Dawson [25], secara umum metode yang biasa digunakan baik oleh individu atau gabungan ada empat metode yaitu penelitian tindakan, eksperimen, studi kasus, survei. Dalam penelitian ini menggunakan metode penelitian eksperimen dengan pendekatan kuantitatif. Penelitian ini dilakukan dengan tahapan yang ada pada Gambar 1.



Gambar 1. Tahapan Penelitian

 ISSN:2302-8149
e-ISSN:2540-9719

3.1. Pengumpulan Data

Pada tahapan ini, menentukan data yang akan diteliti. Mengintegrasikan semua data ke dalam *dataset*, termasuk menentukan variabel yang diperlukan [26].

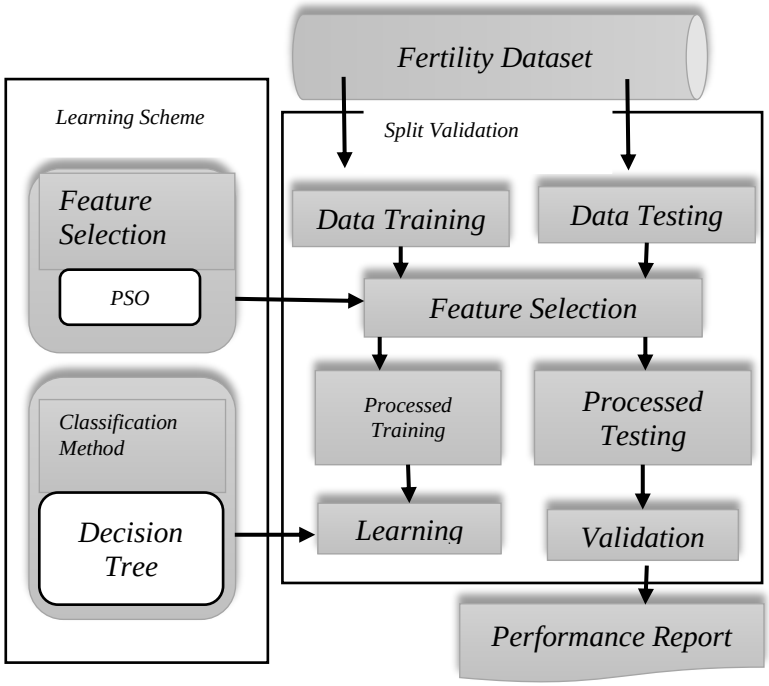
Dataset yang digunakan dalam penelitian ini yaitu *Fertility Dataset* yang merupakan data publik atau biasa disebut data sekunder. Data sekunder adalah data yang bukan dihasilkan dari penelitian melainkan didapat dari sumber lain. *Fertility Dataset* adalah hasil dari penelitian yang dilakukan oleh David Gil, Jose Luis Girela, Joaquin De Juan, M. Jose Gomez-Torres, Magnus Johnsson pada tahun 2012. *Dataset* ini dipublikasikan melalui situs web UCI *Learning Machine Repository*. *Fertility Dataset* ini terdiri atas 100 *record* dengan 10 atribut yang dapat dilihat pada Tabel 1. Label *dataset* ini terdiri atas 2 *class* antara lain N (normal) dan O (*altered*). Dimana 88 *record* pada *class* N (normal) dan 12 *class* pada *class* O (*altered*). Pengambilan *dataset* ini langsung dari halaman situs web UCI *Learning Machine Repository* dengan nama *dataset Fertility Dataset.*

3.2. Pengolahan Data Awal

Pada tahapan ini, membutuhkan eksplorasi atau pendalaman terhadap *Fertility Dataset.* Eksplorasi dilakukan dengan tujuan untuk menunjukkan pada semua atribut dan *class* dalam *dataset* tersebut *valid*, sehingga bisa digunakan untuk objek penelitian yang baik. Maka dari itu, tujuan untuk mengetahui hasil klasifikasi yang terbaik dari *Fertility Dataset*.

3.3. Model yang Diusulkan

Fertility Dataset adalah data sekunder yang telah siap untuk diproses dalam *data mining.* Model yang digunakan pada proses ini yaitu model yang belum pernah dilakukan oleh para peneliti sebelumnya dalam penggunaan *fertility dataset.* Model yang akan digunakan dalam penelitian ini yaitu dengan metode klasifikasi *Decision Tree* dan *Particle Swarm Optimization* sebagai metode optimasi. Bisa dilihat pada Gambar 2 adalah model yang diusulkan.



Gambar 2. Model yang Diusulkan

Model yang diusulkan pada Gambar 2 dilakukan dalam tahapan sebagai berikut :

1. Tahapan pertama yaitu pengumpulan *dataset fertility* yang diambil dari UCI *Repository*.
2. Berikutnya pembagian data menggunakan *split validation*, yaitu membagi data menjadi *data training* dan *data testing.*

 ISSN:2302-8149
e-ISSN:2540-9719

3. Pada tahap *learning scheme* menggunakan *Particle Swarm Optimization* untuk optimasi, sehingga akan mengoptimasi nilai bobot pada setiap atribut. *Particle Swarm Optimization* (PSO) adalah algoritma pencarian berbasis populasi dan diinisialisasi dengan populasi solusi acak yang disebut partikel [27]. Modifikasi kecepatan dan posisi partikel dapat di hitung menggunakan jarak pbesti, gbestd seperti ditunjukan persamaan (1) dan (2) berikut [17] , [28]

   vi,d = w * vi,d + c1 * R * (pbesti,d - xi,d) + c2 * R * (gbestd - xi,d) (1)

   xid = xi,d + vi,d (2)

   Keterangan :
   Vi, d = Kecepatan partikel ke-i pada iterasi ke- i
   w = Faktor bobot inersia
   c1, c2 = Konstanta akselerasi (learning rate)
   R = Bilangan random (0-1)
   Xi, d = Posisi saat ini dari partikel ke-i pada iterasi ke-i
   pbesti = Posisi terbaik sebelumnya dari partikel ke-i
   gbesti = Partikel terbaik di antara yang terbaik pada semua partikel dalam satu kelompok atau populasi.

   Berikutnya algoritma *decision tree* untuk mengklasifikasi kesuburan pada pria dan menghasilkan pohon keputusan. Algoritma *decision tree* adalah salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi bersifat prediktif [29]. Adapun tahapan dalam pembuatan sebuah pohon keputusan menggunakan algoritma *Decision Tree* adalah sebagai berikut [30], [8] :
   a. Pilihlah atribut sebagai akar.
   b. Buat cabang untuk masing-masing nilai.
   c. Bagi kasus dalam cabang.
   d. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

   Dalam pemilihan atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang sudah ada. Perhitungan *gain* dan *entropy* dapat dilihat pada persamaan (3) dan (4) berikut :

   $$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} Entropy(S_i) \quad (3)$$

   Keterangan :
   S : Himpunan Kasus
   A : Atribut
   N : Jumlah partisi atribut A
   |Si| : Jumlah kasus pada partisi ke i
   |S| : Jumlah kasus dalam S

   Entropy dapat dilakukan dengan rumus :
   $$Entropy(S) = \sum_{i=1}^{n} -p_i \times \log_2 p_i \quad (4)$$

   Keterangan :
   S : Himpunan Kasus
   N : Jumlah partisi S
   Pi : Proporsi dari Si terhadap S

4. Berikutnya yaitu dilakukan validasi menggunakan *Confusion matrix* dan kurva ROC.

ISSN:2302-8149
e-ISSN:2540-9719

3.4. Eksperimen dan Pengujian

Dalam tahapan ini, dilakukan eksperimen terhadap model data yang akan diproses menggunakan metode yang telah diusulkan menggunakan sebuah *software*. Eksperimen pada penelitian ini dilakukan terhadap *Fertility Dataset*, setelah itu *dataset* akan dibagi menjadi dua bagian yaitu *data training* sebesar 70% sedangan *data testing* sebesar 30%. Jumlah dalam *data training* dan *data testing* ditentukan dengan mengubah metode *split* dimana awalnya *relative* menjadi *absolute*. Kemudian akan diterapkan algoritma optimasi menggunakan *Particle Swarm Optimization* untuk mengoptimasi nilai bobot setiap atribut, sehingga akan meningkatkan nilai akurasi dari *dataset* tersebut. Model ini akan menampilkan nilai akurasi dengan tampilan *Confusion Matrix* serta diukur menggunakan *Area Under Curve* (AUC).

3.5. Evaluasi dan Validasi

Tahapan ini dilakukan evaluasi dari model yang akan digunakan. Proses evaluasi ini dilakukan menggunakan *Confusion Matrix* dan kurva ROC (*Receiver Operating Characteristic)*. Sedangkan proses validasi yang akan dilakukan dengan menggunakan *Split Validation*, yaitu data yang telah dibagi menjadi *data training* dan *data testing*. Tahap evaluasi dengan menggunakan *Software Rapidminer* dengan secara otomatis ketika proses sedang dijalankan. Dengan performa model yang akan dibandingkan yaitu perfoma algoritma *Decision Tree* dengan algoritma *Decision Tree* yang sudah dioptimasi menggunakan *Particle Swarm Optimization.*

4 Hasil dan pembahasan

Dalam penerapan algoritma *Decision Tree* dengan *Particle Swarm Optimization* pada *Fertility Dataset* dengan tujuan untuk mengetahui dan mendapatkan hasil nilai akurasi lebih baik pada klasifikasi kesuburan pada pria dari penelitian sebelumnya yang dilakukan oleh [1], [21], [22]. *Dataset* yang sudah diolah selanjutnya akan dilakukan perhitungan menggunakan *Software RapidMiner* untuk mengetahui hasil nilai akurasi metode yang digunakan pada penelitian ini.

4.1. Hasil Eksperimen *Decision Tree*

Eksperimen ini dilakukan pada *rapidminer*, dengan pembagian *dataset* menggunakan *split validation*, yaitu pembagian jumlah *data training* dan *testing* ditentukan secara manual. Pada penelitian ini pembagian *data training* dan *testing* adalah 70%:30%, yaitu 70% untuk *data training* atau 70 data dan 30% untuk *data testing* atau 30 data. *Data training* ini untuk pengembangan pada model dan *data testing* ini untuk pengujian model. Adapun pembagian *data training* dan *data testing* sebesar 70:30, karena nilai akurasinya lebih besar dibandingkan dengan pembagian data lainnya. Kemudian dihitung menggunakan *Decision Tree* dan menghasilkan *confusion matrix* seperti Tabel 1.

Tabel 1. *Confusion Matrix Decision Tree*

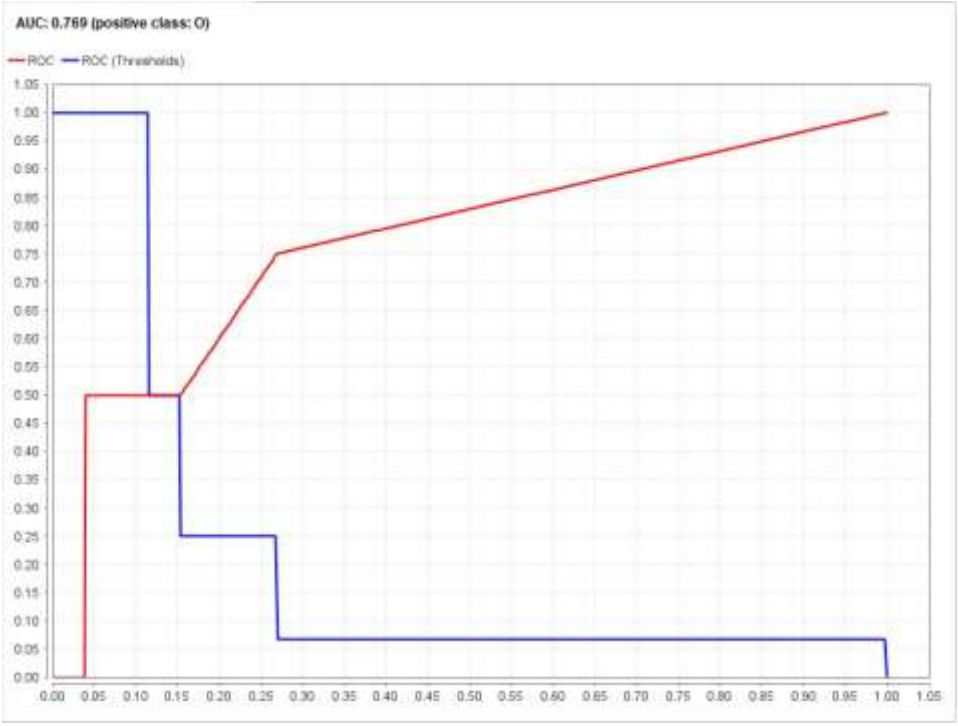
	True N	***True O***	***Class Precision***
Pred. N	25	2	92.59%
Pred. O	1	2	66.67%
Class Recall	96.15%	50.00%	

Dari *confusion matrix* pada Tabel 1. dapat diukur tingkat akurasi dari klasifikasi sebagai berikut:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

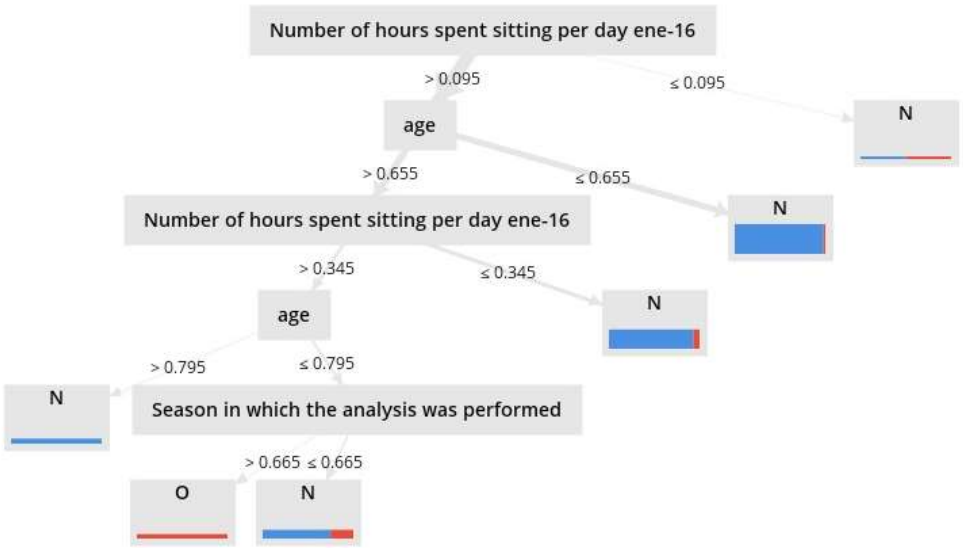
$$= \frac{25+2}{25+2+1+2} \times 100\% = 90{,}00\%$$

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 3. Grafik ROC pada Eksperimen Metode *Decision Tree*

Selain dapat menghasilkan *confusion matrix*, pengujian tersebut juga menghasilkan kurva ROC (*Receiver Operating Characteric*) seperti yang ada pada Gambar 3 dengan nilai AUC sebesar 0,769 dan termasuk ke dalam kategori *Fair classification.*



Gambar 4. Hasil Pohon Keputusan *Fertility Dataset*

Akar pohon diambil dari atribut yang terpilih yaitu *Number of hours spent sitting per day*, dilanjut dengan atribut *Age* seperti terlihat pada Gambar 4.

 ISSN:2302-8149
e-ISSN:2540-9719

4.2. Hasil Eksperimen Menggunakan *Decision Tree* dan *Particle Swarm Optimization*

Pada tahap ini akan dijelaskan hasil eksperimen menggunakan metode *Particle Swarm Optimation* sebagai optimasi bobot atribut yang dimaksud untuk menambah nilai bobot dari masing-masing atribut pada suatu *dataset* untuk klasifikasi kesuburan pria akan dilakukan perhitungan menggunakan *RapidMiner* dan menghasilkan *weight* atau nilai bobot seperti Tabel 2.

Tabel 2. Hasil Optimasi Bobot Atribut

Attribute	***Weight***
Season in which the analysis was performed.	1
Age	1
Childish diseases	0.646
Accident or serious trauma	0.136
Surgical intervention	0.001
High fevers in the last year	0.930
Frequency of alcohol consumption	0.907
Smoking habit	0
Number of hours spent sitting per day	0.563

Hasil dari optimasi bobot atribut ini mendapatkan 1 atribut terendah yaitu dengan *weight* 0 pada atribut *Smoking Habit* yang menunjukan bahwa atribut tersebut tidak memiliki pengaruh terhadap nilai akurasi.

Bobot dari atribut yang bernilai 1 dapat memungkinkan mempengaruhi hasil akurasi secara signifikan. Tedapat 2 atribut yang bobotnya 1 yaitu *Season In Which The Analysis Was Performed* dan *Age.*

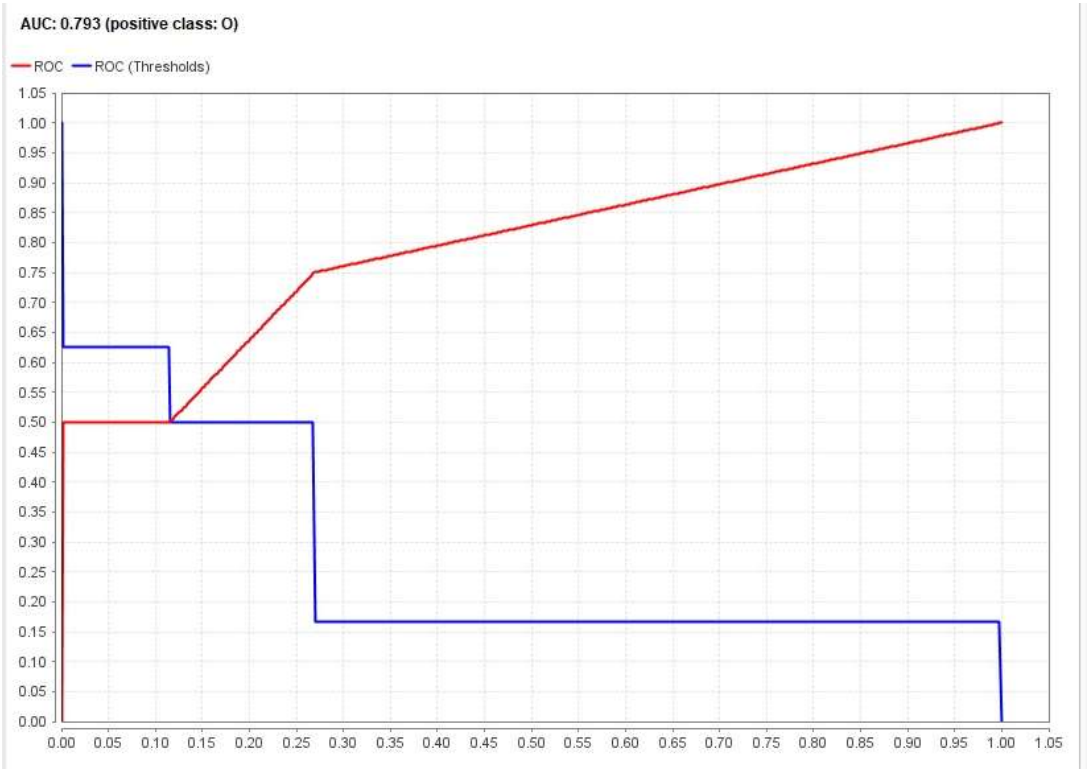
Tabel 3. *Confusion Matrix Decision Tree* dengan *Particle Swarm Optimization*

	True N	***True O***	***Class Precision***
Pred. N	26	2	92.86%
Pred. O	0	2	100.00%
Class Recall	100.00%	50.00%	

Dari *confusion matrix* pada Tabel 3. dapat diukur tingkat akurasi dari klasifikasi sebagai berikut :

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$
$$= \frac{26 + 2}{26 + 2 + 0 + 2} \times 100\% = 93{,}33\%$$

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 5. Grafik ROC Pada Eksperimen Metode *Decision Tree* Dan *Particle Swarm Optimization*

Kurva ROC (*Receiver Operating Characteric*) seperti yang ada pada Gambar 5 dengan nilai AUC sebesar 0,793 dan termasuk ke dalam kategori *Fair classification.*

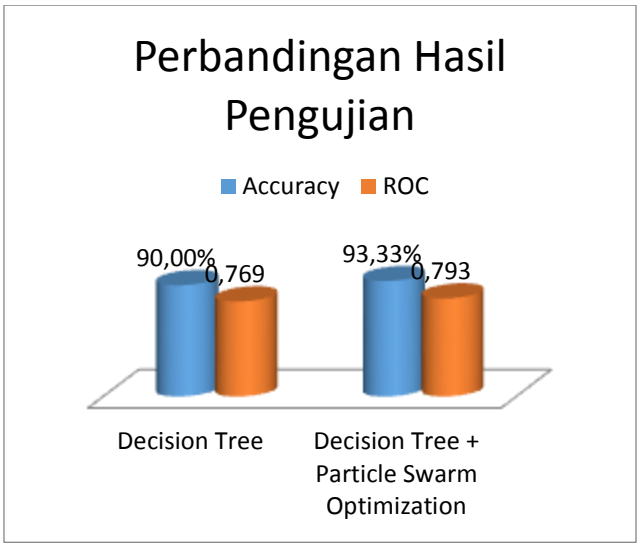
4.3. Pembahasan

Dari hasil penelitian yang telah dilakukan, maka pada tahapan ini akan dilakukan evaluasi sehingga diketahui peningkatan nilai akurasi dari klasifikasi metode *Decision Tree* sebelum dilakukan optimasi bobot atribut dan *Decision Tree* setelah dilakukan optimasi bobot atribut menggunakan *Particle Swarm Optimization* dengan cara menghitung *data testing* dalam bentuk nilai akurasi dan grafik ROC. Berikut perbandingan nilai akurasi *Decision Tree* sebelum dilakukan optimasi bobot atribut dan *Decision Tree* setelah dilakukan optimasi bobot atribut menggunakan *Particle Swarm Optimization*.

Tabel 4. Perbandingan Hasil Pengujian

No.	Metode	*Accuracy*	ROC
1	*Decision Tree*	90,00%	0.769
2	*Decision Tree + Particle Swarm Optimization*	93,33%	0.793

 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 6. Grafik Perbandingan Hasil Pengujian

Pada Gambar 6 dapat dilihat bahwa hasil klasifikasi metode *Decision Tree* mengalami peningkatan akurasi sebersar 3.33% setelah dilakukan optimasi bobot menggunakan *Particle Swarm Optimization*. Sehingga terbukti bahwa optimasi *Particle Swarm Optimization* dapat meningkatkan performa atau kinerja algoritma *Decision Tree* pada klasifikasi kesuburan pria.

Pada perbandingan hasil penelitian sebelumnya juga dilakukan sehingga dapat mengevaluasi hasil keseluruhan penelitian. Penelitian ini merupakan penelitian lanjutan dari penelitian-penelitian sebelumnya dengan menggunakan data yang sama yaitu *Fertility Dataset*.

5 Kesimpulan

Telah diterapkan algoritma klasifikasi *Decision Tree* berbasis *Particle Swarm Optimization* pada *Fertility Dataset*. Setelah dilakukan optimasi pada *Fertility Dataset* menggunakan *Decision Tree* berbasis *Particle Swarm Optimization* menghasilkan peningkatan akurasi dan performa. Hasil akurasi dengan algoritma *Decision Tree* pada *Fertility Dataset* menghasilkan nilai akurasi *confusion matrix* sebesar 90.00% dan AUC sebesar 0.769 setelah dilakukan optimasi menggunakan *Particle Swarm Optimization* untuk pembobotan atribut pada *Fertility Dataset* serta hasil akurasi dari *Decision Tree* meningkat menjadi sebesar 93.33% dan AUC sebesar 0.793. sehingga dapat disimpulkan bahwa *Particle Swarm Optimization* dapat meningkatkan akurasi dari algoritma *Decision Tree* pada *Fertility Dataset*.

Referensi

[1] A. Amrulloh and A. P. Wibowo, "Implementasi Algoritma Decission Tree Untuk Mengklasifikasi Kondisi Kesuburan Pria," *JASIEK (Jurnal Aplikasi Sains, Informasi, Elektron. dan Komputer)*, vol. 1, no. 1, pp. 7–11, 2019, doi: 10.26905/jasiek.v1i1.3096.

[2] G. A. Irawan and A. Muliantara, "Prediksi Kesuburan (Fertility) dengan menggunakan Principal Component Analysis dan Klasifikasi Naive Bayes, " *Jurnal Ilmiah Komputer Universitas Udayana*, vol. X no. 2, pp. 10–14, 2017.

[3] H. Harafani and A. Maulana, "Penerapan Algoritma Genetika pada Support Vector Machine Sebagai Pengoptimasi Parameter untuk Memprediksi Kesuburan," *Jurnal Teknik Informatika. STMIK Antar Bangsa*, vol. V, no. 1, pp. 51–59, 2019.

[4] K. Rusman, "Pengaruh Aktivitas Merokok Terhadap Hasil Analisa Sperma Pada Kasus Infertilitas Pria di Makassar" *UMI Medical Journal*, vol. 4, no. 1, pp. 117–129, 2019.

[5] Y. J. S. Ningsih and A. Farich, "Determinan Kejadian Infertilitas Pria di Kabupaten Tulang Bawang," *J. Kesehat.*, vol. VII, no. 2, pp. 242–249, 2016.

[6] D. C. P. Buani, "Optimasi Algoritma Naïve Bayes dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (Fertility)," *EVOLUSI: Jurnal Sains dan Teknologi*, " vol. 4, no. 1, pp. 1–10, 2016.
[7] S. Hosana and D. S. D. Putra, "Penerapan Data Mining Untuk Prediksi Kualitas Sperma (Fertility) Berbasis Android," *Jurnal Algor*, vol. 1, no. 1, 2019.
[8] R. T. Vulandari, *Data Mining Teori dan Aplikasi Rapidminer*, Cetakan 1. Yogyakarta: Penerbit Gava Media, 2017.
[9] H. Sulastri and A. I. Gufroni, "Penerapan Data Mining dalam Pengelompokan Penderita Thalassaemia," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 299–305, 2017, doi: 10.25077/teknosi.v3i2.2017.299-305.
[10] Y. Mardi, "Data Mining: Klasifikasi menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
[11] R. Sari, "Komparasi Algoritma Support Vector Machine, Naïve Bayes dan C4.5 untuk Klasifikasi SMS," *IJCIT (Indonesia J. Comput. Infomation Technol.*, vol. 2, no. 2, pp. 7–13, 2017.
[12] D. A. Ningtyas, M. Wahyudi, and N. Nurajijah, "Klasifikasi Siswa SMK Berpotensi Putus Sekolah Menggunakan Algoritma Decision Tree , Support Vector Machine dan Naive Bayes," *J. Khatulistiwa Inform.*, vol. VII, no. 2, pp. 85–90, 2019.
[13] H. Noor, "Optimasi Model Klasifikasi C4.5 dan Particle Swarm Optimization untuk Prediksi Siswa Bermasalah," *Technol. J. Ilm.*, vol. 9, no. 4, p. 228, 2018, doi: 10.31602/tji.v9i4.1537.
[14] A. Sudibyo, T. Asra, and B. Rifai, "Klasifikasi Seleksi Atribut pada Serangan Spam," *J. PILAR Nusa Mandiri*, vol. 14, no. 2, pp. 145–150, 2018, doi: https://doi.org/10.33480/pilar.v14i2.31.
[15] I. A. Marie, L. Hakim, D. Sugiarto, and W. Septiani, "Perbandingan Performansi Teknik Klasifikasi Breakdown Mesin pada Proses Produksi Pembuatan Battery Mobil," *J. Ilm. Tek. Ind.*, vol. 18, no. 1, pp. 33–41, 2019, doi: 10.23917/jiti.v18i1.7232.
[16] Sulistiyanto, "Penerapan C4.5 Berbasis Particle Swarm Optimization (PSO) dalam Memprediksi Siswa Lolos Seleksi Perguruan Tinggi," *Semin. Nas. Teknol. dan Bisnis*, pp. 162–170, 2018.
[17] T. Arifin and A. Herliana, "Optimasi Metode Klasifikasi dengan Menggunakan Particle Swarm Optimization untuk Identifikasi Penyakit Diabetes Retinopathy," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 4, no. 2, p. 77, 2018, doi: 10.23917/khif.v4i2.6825.
[18] T. Arifin, "Implementasi Algoritma PSO dan Teknik Bagging untuk Klasifikasi Sel Pap Smear," *J. Inform.*, vol. 4, no. 2, pp. 155–162, 2017, doi: 10.31311/JI.V4I2.2129.
[19] A. Rakhman, "Prediksi Ketepatan Kelulusan Mahasiswa menggunakan Metode Decision Tree Berbasis Particle Swarm Optimation (PSO)," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 6, no. 1, pp. 193–197, 2017, doi: 10.30591/SMARTCOMP.V6I1.466.
[20] S. Rustam, H. A. Santoso, and C. Supriyanto, "Optimasi K-Means Clustering untuk Identifikasi Daerah Endemik Penyakit Menular dengan Algoritma Particle Swarm Optimization di Kota Semarang," *Ilkom. Jurnal Ilmiah*, vol. 10, no. 3, pp. 251–259, 2018, doi: 10.33096/ilkom.v10i3.342.251-259.
[21] E. Budianita, F. R. Hustianto, F. Syafria, and M. Nasir, "Implementasi Algoritma Jaringan Syaraf Tiruan (JST) Hopfield untuk Klasifikasi Kualitas Kesuburan Pria," *Seminar Nasional Teknologi Informasi, Komununikasi dan Industri*, 2018.
[22] E. Nurelasari, "Komparasi Algoritma Naive Bayes dengan Support Vector Machine Berbasis Particle Swarm Optimization untuk Prediksi Kesuburan," *Bina Insani ICT Journal*, vol. 5, no. 1, pp. 61–70, 2018.
[23] T. A. Y. Siswa and Prihandoko, "Analisis Penerapan Optimasi Perbandingan Kinerja Algoritma C4.5 dan Naïve Bayes Berbasis Particle Swarm Optimization (PSO) Untuk Mendeteksi Kanker Payudara," *Bangkit Indonesia*, vol. 2, no. VII, pp. 1–9, 2018.
[24] A. Rifai and R. Aulianita, "Komparasi Algoritma Klasifikasi C4.5 dan Naïve Bayes Berbasis Particle Swarm Optimization Untuk Penentuan Resiko Kredit," *J. speed- sentra Penelit. enginering dan edukasi*, vol. 10, no. 2, 2018.
[25] C. W. Dawson, *Projects in Computing and Information Systems : A Student's Guide*, Second.,

vol. 2. England: Addison Wesley, 2009.
[26] T. Arifin and D. Ariesta, "Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization," *J. Tekno Insentif*, vol. 13, no. 1, pp. 26–30, 2019, doi: 10.36787/jti.v13i1.97.
[27] H. Amalia, "Penerapan Metode Neural Network Berbasis Particle Swam," *Paradigma*, vol. XVII, no. 1, pp. 1–8, 2015.
[28] S. H. Waluyo and Prihandoko, "Klasifikasi Pemanfaat Program Beras Sejahtera (RASTRA) Berdasarkan Tingkat Kemiskinan Dengan Menggunakan Algoritma Decision Tree C4.5 Berbasis Particle Swarm Optimization," vol. 7, no. 2, pp. 19–24, 2017.
[29] A. P. Windarto, D. R. S. P, D. Hartama, and I. S. Damanik, "Penerapan Klasifikasi C4.5 Dalam Meningkatkan Sistem Pembelajaran Mahasiswa," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 593–597, 2019, doi: 10.30865/komik.v3i1.1665.
[30] Y. S. Luvia, D. Hartama, A. P. Windarto, and Solikhun, "Penerapan Algoritma C4.5 untuk Klasifikasi Predikat Keberhasilan Mahasiswa dI AMIK Tunas Bangsa," *JURASIK (Jurnal Ris. Sist. Inf. Tek. Inform.*, vol. 1, no. 1, 2016.