

PENERAPAN ALGORITMA C4.5 BERBASIS PARTICLE SWARM OPTMIZATION UNTUK PREDIKSI HASIL LAYANAN KEMUDAHAN DONASI ZAKAT DAN PROGRAM

Nur Lutfiyana

Sistem Informasi
STMIK Nusa Mandiri
<http://www.nusamandiri.ac.id>
nur.lutfiyana@nusamandiri.ac.id

Abstract— Cashless Service is a donation facility donated to donors using non-cash donation services through banking or other channeling. Convenience is part of the service commitment of RZ institutions as a means for donors or partners who wish to donate. Efforts to develop the organization to a better direction in improving the satisfaction of the donors to the service process provided is a necessity, therefore it is needed an introduction in the prediction of loyal and disloyal donor transactions, in order to know the easy access donations that have been applied. This, the researchers used two comparative models: algorithm model C4.5 and Particle Swarm Optimization (C4.5) based C4.5 algorithm model. After testing the accuracy value obtained for model classification algorithm C4.5 of 83.91% and PSO (Particle Swarm Optimization) C4.5 algorithm resulted in AUC 0.889 value of accuracy of PSO-based C4.5 algorithm based on diagnostic value Good Classification.

Keywords: Ease Donation, C4.5, Particle Swarm Opimization (PSO)

Intisari— Layanan *Cashless* adalah Layanan kemudahan donasi yang diberikan kepada donatur dengan menggunakan layanan donasi non tunai melalui perbankan atau channeling lainnya. Kemudahan adalah bagian dari komitmen layanan dari lembaga RZ sebagai sarana untuk donatur atau mitra yang ingin berdonasi. Upaya untuk mengembangkan organisasi ke arah yang lebih baik dalam meningkatkan kepuasan para donatur terhadap proses pelayanan yang diberikan merupakan suatu keharusan, oleh karena itu dibutuhkan pengenalan dalam prediksi terhadap transaksi donatur yang loyal dan tidak loyal, agar bisa diketahui akses kemudahan donasi yang telah di terapkan. Dengan demikian, peneliti menggunakan dua model perbandingan yaitu: model algoritma C4.5 dan model algoritma C4.5 berbasis *Particle Swarm Optimization* (PSO). Setelah dilakukan pengujian nilai akurasi yang di dapatkan untuk model algoritma klasifikasi C4.5 sebesar 83,91 % dan algoritma C4.5 berbasis PSO

(*Particle Swarm Optimization*) menghasilkan nilai AUC 0.889 dari nilai akurasi algoritma klasifikasi C4.5 berbasis PSO, dengan nilai diagnosa *Good Classification*.

Kata kunci: Kemudahan Donasi, C4.5, *Particle Swarm Opimization* (PSO)

PENDAHULUAN

Kemudahan adalah bagian dari komitmen layanan dari lembaga RZ sebagai sarana untuk donatur atau mitra yang ingin berdonasi. Layanan *Cashless* adalah Layanan kemudahan donasi yang diberikan kepada donatur dengan menggunakan layanan donasi non tunai melalui perbankan atau *channeling* lainnya. Lembaga Zakat adalah Lembaga yang memfasilitasi untuk mengamanahkan zakat, infak/shadaqah serta program-program lainnya dikemas dalam bentuk yang kreatif dan inovatif, sehingga donatur akan terus bertambah. Tetapi dalam kenyataannya donatur suka berpindah Lembaga zakat satu ke lembaga yang lainnya. Sehingga Lembaga Zakat yang memiliki jumlah donatur yang banyak memiliki kecenderungan kompetisi yang ketat dan mempertahankan donatur merupakan tantangan besar pada suatu lembaga.

Sepanjang tahun 2001 sampai sekarang merupakan era kebangkitan baru dalam perekonomian Indonesia yaitu dengan kehadiran Badan Amil Zakat (BAZ) serta Lembaga-Lembaga Amil Zakat (LAZ) dalam meningkatkan pertumbuhan ekonomi dhuafa sangat signifikan, terbukti dengan tumbuhnya pengelolaan zakat dengan pola distribusi pada arah pengembangan produktivitas mustahik.

Menurut BAZNAS: Potenzi zakat di Indonesia mencapai Rp 217 triliun, namun baru terserap dan dikelola oleh lembaga amil zakat sebesar Rp 2,73 trilun (sekitar 1%). Dengan pengelolaan yang baik, zakat merupakan sumber dana potensial yang dapat dimanfaatkan untuk memajukan kesejahteraan umum bagi seluruh

masyarakat. Optimalisasi dana zakat terhadap upaya penanggulangan kemiskinan telah terbukti dapat menekan angka kemiskinan sebesar 10,79 persen.

Pengentasan kemiskinan memerlukan suatu upaya pemberdayaan yang bersifat bottomup dengan melibatkan seluruh elemen masyarakat melalui budaya tolong-menolong (Munir, 2005). Dengan hadirnya peran serta dari pihak instansi, tataran pemerintahan dan *stakeholder* lainnya agar lebih peduli untuk kolaborasi terhadap optimalisasi dalam penanggulangan kemiskinan, baik dari segi potensi dana zakat atau program pemberdayaan lainnya yang ada sampai pengelolaan dari dana zakat tersebut. Jika tidak adanya saling sinergi akan semakin meningkatnya tingkat kemiskinan. Untuk itu layanan kemudahan donasi di lembaga RZ ini membuat layanan yang membuat para donatur semakin mudah untuk bisa berdonasi dimanapun dan kapanpun.

Dalam proses perencanaan kerja lembaga zakat, bagi sebuah lembaga zakat untuk dapat memproyeksikan jumlah pendapatan lembaga zakat dari tahun-tahun mendatang. Agar pengentasan kemiskinan semakin berkurang, dengan terbentuknya lembaga zakat bergerak di bidang sosial dapat menentukan besarnya pencapaian dari jumlah donasi donatur aktif yang dimiliki. Untuk dapat mengetahui jumlah donasi donatur aktif, maka perlu sebuah model yang dapat melihat hasil donasi donatur sehingga lembaga zakat dapat mengenali prediksi donasi transaksi donatur loyal dan tidak loyal dan mampu memprediksi seberapa banyak donatur yang akan meninggalkan lembaga zakat dan berpindah kepada lembaga lain. Dengan mengetahui prediksi layanan kemudahan donasi donatur berbasis program, lembaga zakat juga dapat meningkatkan hubungan dengan para donaturnya.

Maka untuk bisa mempertahankan donatur kita harus mengetahui pola donatur agar bisa memprediksi donasi zakat dan berdonasi berbasis program berdasarkan transaksi data yang ada. Klasifikasi merupakan topik yang penting dalam penelitian data mining. Dimana terdiri dari sekumpulan data yang masing-masing sudah dikelompokkan kedalam kelas tertentu, masalah klasifikasi yang diperhatikan yaitu dengan penentuan aturan yang memungkinkan adanya klasifikasi walaupun data yang ada belum menunjukkan hal tersebut.

Data mining merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santosa & Willy, 2011). Peneliti Hermansyah

(Hermansyah, 2011) melakukan penelitian strategi pengembangan donatur pada lembaga amal zakat dengan pendekatan data mining, metode yang digunakan *clustering* dan hasil pengelompokan tersebut di evaluasi dengan *artificial neural network* dan menggunakan metode *association rules* untuk menganalisa peta pemasaran serta dirumuskan strategi pengembangan donatur berdasarkan data transaksi dan visi misi. Hasil terdapat dua kelompok donatur lembaga amal zakat yaitu kelompok *low* 25% dan *high* 75% beserta karakteristik program dan potensi pengembangan donatur. Akurasi *decision tree* lebih baik dibandingkan dengan *neural network* (Tsai & Chen, 2010)

BAHAN DAN METODE

Peneliti Hermansyah (Hermansyah, 2011) melakukan penelitian strategi pengembangan donatur pada lembaga amal zakat dengan pendekatan data mining, metode yang digunakan *clustering* dan hasil pengelompokan tersebut dievaluasi dengan *artificial neural network* dan menggunakan metode *association rules* untuk menganalisa peta pemasaran serta dirumuskan strategi pengembangan donatur berdasarkan data transaksi dan visi misi. Hasil terdapat dua kelompok donatur lembaga amal zakat yaitu kelompok *low* 25% dan *high* 75% beserta karakteristik program dan potensi pengembangan donatur.

Peneliti Ragita (Aprillisa & Wirawan, 2013) analisis manajemen hubungan terhadap donatur untuk meningkatkan loyalitas dan jumlah donatur vihara. Peneliti menggunakan pendekatan kualitatif dengan metode wawancara semi terstruktur, observasi serta analisis dokumen. Berdasarkan time horizon, data yang diambil merupakan data one shot dan cross-sectional di dapatkan hasil yang dapat mengacu loyalitas dari tiap donatur.

Peneliti Aryani dan Rosinta (Aryani & Rosinta, 2011) pengaruh kualitas layanan terhadap kepuasan pelanggan dalam membentuk loyalitas pelanggan. Peneliti menggunakan pendekatan kuantitatif berdasarkan teknik pengumpulan data dan informasi ada dua macam data yaitu data primer dan data sekunder. Model dilakukan dalam peneliti ini menggunakan metode estimasi Maximum Likelihood dan menggunakan metode structural equation modeling (SEM) yang dinilai lebih efisien dan unbiased jika asumsi multivariate terpenuhi. Dimensi terkuat dalam menjelaskan kualitas layanan berturut-turut adalah reliability, responsiveness, assurance, empathy dan tangibility. Hal ini terbukti dari hasil penelitian yang menunjukkan bahwa sebesar

72,9% variabel kepuasan pelanggan dapat dijelaskan kualitas layanan, sedangkan sisanya sebesar 27,1% dipengaruhi oleh faktor lain di luar variabel kualitas layanan.

Layanan Kemudahan Donasi

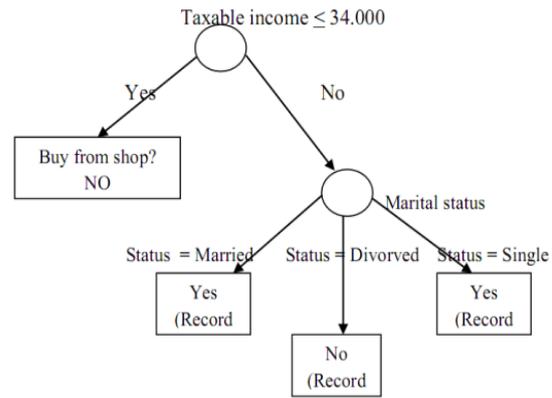
Layanan *Cashless* adalah Layanan kemudahan donasi yang diberikan kepada donatur dengan menggunakan layanan donasi non tunai melalui perbankan atau channeling lainnya. Kemudahan adalah bagian dari komitmen layanan dari lembaga RZ sebagai sarana untuk donatur atau mitra yang ingin berdonasi.

Tuntutan akan kinerja untuk menjadi organisasi yang profesional dan memiliki program kerja yang jelas dan terencana bagi organisasi merupakan suatu kebutuhan untuk menjawab tantangan zaman. Upaya untuk mengembangkan organisasi ke arah yang lebih baik dalam meningkatkan kepuasan para donatur terhadap proses pelayanan yang diberikan merupakan suatu keharusan, oleh karena itu dibutuhkan strategi layanan donasi agar lebih memudahkan para donatur.

Algoritma C4.5

Algoritma C4.5 adalah Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun juga mudah untuk dipahami, sehingga *Decision Tree* ini merupakan metode klasifikasi yang paling populer digunakan. *Decision Tree* adalah *flow-chart* seperti struktur *tree*, dimana tiap internal *node* menunjukkan sebuah *test* pada sebuah atribut, tiap cabang menunjukkan hasil dari test dan *leaf node* menunjukkan class-class atau *class distribution* (Sunjana, 2010). Algoritma ini digunakan untuk membangun sebuah pohon keputusan yang mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar (Gorunescu, 2011). Sebelumnya diakhir tahun 1970 sampai awal tahun 1980 J. Ross Quinlan, seorang peneliti dibidang *machine learning*, membuat sebuah algoritma *decision tree* yang dikenal dengan ID3 (*Iterative Dichotomiser*).

Algoritma C4.5 atau pohon keputusan mirip sebuah pohon dimana terdapat node internal (bukan daun) yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas. Pohon keputusan dengan mudah dapat dikonversi ke aturan klasifikasi. Secara umum keputusan pengklasifikasi pohon memiliki akurasi yang baik, namun keberhasilan penggunaan tergantung pada data yang akan diolah.



Sumber: (Gorunescu, 2011).

Gambar 1. Contoh konsep pohon keputusan sederhana

Pada Gambar 1 variabel target untuk pohon keputusan adalah membeli pada toko, dengan pengklasifikasian ya atau tidak. Variabel *predictor* adalah taxable income (<34.000 atau > 34.000), marital status (married, divorced, atau single). Simpul akar merupakan simpul keputusan, pengujiannya apakah taxable income < 34.000 atau > 34.000.

Algoritma C4.5 menggunakan konsep *information gain* atau *entropy reduction* untuk memilih pembagian yang optimal (Larose, 2005). Tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 (Gorunescu, 2011). yaitu:

1. Mempersiapkan data *training*, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masing atribut atau berdasarkan nilai *index entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai *index entropy*, dengan rumus:

$$Entropy(i) = \sum_{j=1}^m f(i,j) \cdot \log 2f[(i,j)] \dots\dots\dots (1)$$

Keterangan:

- i = himpunan kasus
- m = jumlah partisi i
- f(i,j) = proposi j terhadap i

3. Hitung nilai *gain* dengan rumus:

$$Entropy\ split = - \sum_{i=1}^p \binom{n_1}{n} \times IE(i) \dots\dots\dots (2)$$

- p = jumlah partisi atribut
- ni = proporsi ni terhadap i
- n = jumlah kasus dalam n

4. Ulangi langkah ke-2 hingga semua *record* terpartisi. Proses partisi pohon keputusan akan

berhenti disaat:

- a) Semua tupel dalam *record* dalam simpul m mendapat kelas yang sama
- b) Tidak ada atribut dalam *record* yang dipartisi lagi
- c) Tidak ada *record* didalam cabang yang kosong.

Particle Swarm Optimization

Particle Swarm Optimimization (PSO) adalah didasarkan pada perilaku sekawanan burung atau ikan. Algoritma PSO meniru perilaku sosial organisme ini. Perilaku sosial terdiri dari tindakan individu dan pengaruh dari individu-individu lain dalam suatu kelompok (Santosa & Willy, 2011). PSO merupakan metode pencarian penduduk yang berasal dari penelitian untuk gerakan sekelompok burung atau ikan. Serupa dengan algoritma genetik (GA), PSO melakukan pencarian menggunakan populasi (*swarm*) dari individu (partikel) yang akan diperbaharui dari iterasi ke iterasi.

Untuk PSO dapat diasumsikan sebagai Kata partikel menunjukkan, misalnya, seekor burung dalam kawanan burung. Setiap individu atau partikel berperilaku 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 (Santosa & Willy, 2011). PSO digunakan untuk memecahkan masalah optimasi.

Menemukan solusi yang optimal, masing-masing partikel bergerak kearah posisi yang terbaik sebelumnya dan posisi terbaik secara global. Sebagai contoh, partikel ke-i dinyatakan sebagai: $xi = (xi1, xi2, \dots, xid)$ dalam ruang d-dimensi. Posisi terbaik sebelumnya dari partikel ke-i disimpan dan dinyatakan sebagai $pbesti = (pbesti,1, pbesti,2, \dots, pbesti,d)$. Indeks partikel terbaik diantara semua partikel dalam kawanan group dinyatakan sebagai $gbestd$. Kecepatan partikel ke-i dinyatakan sebagai: $vi = (vi,1, vi,2, \dots, vi,d)$. Modifikasi kecepatan dan posisi tiap partikel dapat dihitung menggunakan kecepatan saat ini dan jarak $pbest i,d$ ke $gbestd$ seperti ditunjukkan persamaan berikut:

$$vi,d = w * vi,d + c1 * R * (pbesti,d - xi,d) + c2 * R * (gbestd - xi,d) \dots\dots\dots(3)$$

$$xid = xi,d + vi,d \dots\dots\dots(4)$$

Dimana:

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 diantara semua partikel dalam satu kelompok atau populasi
- n = Jumlah partikel dalam kelompok
- d = Dimensi

Metode penelitian eksperimen, terdapat beberapa tahapan penelitian yang dilakukan seperti berikut:

1. Pengumpulan data
Pada tahap ini ditentukan data yang akan diproses. Mencari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data kedalam data set, termasuk variabel yang diperlukan dalam proses.
2. Pengolahan awal data
Ditahap ini dilakukan penyeleksian data, data dibersihkan dan ditransformasikan kebentuk yang diinginkan sehingga dapat dilakukan persiapan dalam pembuatan model.
3. Model yang diusulkan
Pada tahap ini data dianalisis, dikelompokan variabel mana yang berhubungan dengan satu sama lainnya. Setelah data dianalisis lalu diterapkan model-model yang sesuai dengan jenis data. Pembagian data kedalam data latihan (training data) dan data uji (testing data) juga diperlukan untuk pembuatan model.
4. Eksperimen dan pengujian model
Pada tahap ini model yang diusulkan akan diuji untuk melihat hasil berupa rule yang akan dimanfaatkan dalam pengambilan keputusan.
5. Evaluasi dan validasi hasil
Pada tahap ini dilakukan evaluasi terhadap model yang ditetapkan untuk mengetahui tingkat keakurasian model.

HASIL DAN PEMBAHASAN

Model Algoritma C4.5

Algoritma C4.5 untuk model yang pertama dilakukan. Berikut langkah-langkah yang akan dilakukan sebagai berikut:

1. Menghitung jumlah kasus class YES dan class NO serta nilai Entropy dari semua kasus. Kasus dibagi berdasarkan atribut pada Tabel 1. dengan jumlah kasus 3356 record,

kelas YES ada 1739 record dan kelas NO sebanyak 1617 record sehingga didapat entropy:

$$Entropy(i) = \sum_{j=1}^M (i,j) \cdot \log 2 f[(i,j)] \dots\dots (5)$$

$$= (-1739/3356 \cdot \text{Log}_2(1739/3356)) + (-1617/3356 \cdot \text{Log}_2(1617/3356)) = 0,999$$

2. Hitung nilai Gain dari masing-masing atribut pada Tabel 3.3, sebagai contoh untuk jumlah donasi:

$$\leq 1745000 = 94/3356$$

$$> 1745000 = 125/3356$$

Atribut jumlah donasi ≤ 1745000 terdiri dari 19 class YES, 75 class NO dan untuk Atribut jumlah donasi > 1745000 terdiri dari 120 class YES, 5 class NO. Nilai entropynya dapat dihitung sebagai berikut:

$$Entropy.Split = \sum_{i=1}^p \frac{n_i}{n} \dots\dots\dots (6)$$

$$\text{Jumlah donasi } \leq 1745000 = (-19/94 \cdot \text{Log}_2(19/94)) + (-75/94 \cdot \text{Log}_2(75/94)) = 0,726$$

$$\text{Jumlah Donasi } > 1745000 = (-120/125 \cdot \text{Log}_2(120/125)) + (-5/125 \cdot \text{Log}_2(5/125)) = 0,242$$

$$E_{split} \text{ jumlah donasi} = ((94/3356 \cdot (0,726)) + ((125/3356 \cdot (0,242))) = 0,029$$

$$\text{Gain jumlah donasi} = 0,999 - 0,029 = 0,969$$

Tabel 1. Informasi Gain untuk Algoritma C4.5

| Candidate Split | Child Nodes | Entropy | Gain |
|-----------------|-----------------|---------|-------|
| 1 | Jumlah Mustahik | | |
| | ≤ 0.500 | 0,725 | 0,273 |
| | > 0.500 | | |
| 2 | Bulan Donasi | | |
| | Januari | 0,009 | 0,361 |
| | Februari | 0,020 | |
| | Maret | 0,029 | |
| | April | 0,056 | |
| | Mei | 0,043 | |
| | Juni | 0,038 | |
| | Juli | 0,136 | |
| | Agustus | 0,117 | |
| | September | 0,029 | |
| | Oktober | 0,062 | |
| | November | 0,060 | |
| Desember | 0,039 | | |
| 3 | Metode Bayar | | |
| | ATM | 0,064 | 0,883 |
| | Autodebet | 0,010 | |

| | | | |
|----------------|--------------------|-------|-------|
| | Ebanking | 0,008 | |
| | EDC | 0,001 | |
| | Setoran tunai | 0,023 | |
| | Setoran via bank | 0,003 | |
| | Zisco setor teller | 0,006 | |
| 4 | Jml bulan | | |
| | ≤ 0.500 | 0,05 | 0,949 |
| | > 0.500 | | |
| | ≤ 1.500 | 0 | 0 |
| | > 1.500 | | |
| | ≤ 2 | 0 | 0 |
| | > 2 | | |
| | ≤ 6.500 | 0,121 | 0,878 |
| | > 6.500 | | |
| | ≤ 7.500 | 0,008 | 0,990 |
| > 7.500 | | | |
| 5 | Jml Donasi | | |
| | ≤ 7500 | 0,009 | 0,989 |
| | > 7500 | | |
| | ≤ 53250 | 0,004 | 0,995 |
| | > 53250 | | |
| | ≤ 66000 | 0 | 0 |
| | > 66000 | | |
| | ≤ 705000 | 0,004 | 0,995 |
| | > 705000 | | |
| | ≤ 186430 | 0 | 0 |
| | > 186430 | | |
| | ≤ 199500 | 0,006 | 0,993 |
| | > 199500 | | |
| | ≤ 197000 | 0,013 | 0,986 |
| | > 197000 | | |
| | ≤ 450000 | 0,008 | 0,091 |
| > 450000 | | | |
| ≤ 1400000 | 0,014 | 0,985 | |
| > 1400000 | | | |
| ≤ 1745000 | 0,209 | 0,969 | |
| > 1745000 | | | |
| ≤ 2050000 | 0,006 | 0,993 | |
| > 2050000 | | | |

Sumber: (Lutfiyana, 2017)

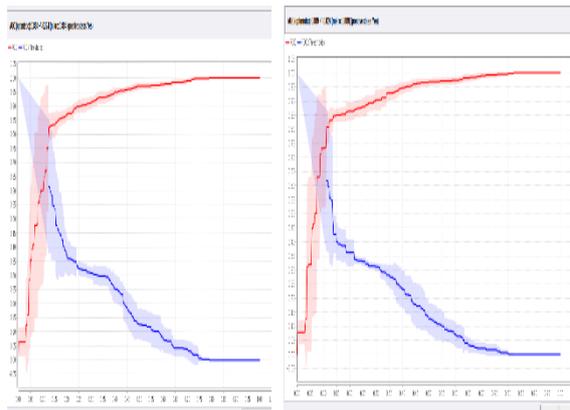
Analisis Evaluasi dan Validasi Hasil Dari hasil pengujian, dengan dilakukan evaluasi baik secara *confusion matrix* maupun *ROC Curve* maka terbukti bahwa pengujian yang dilakukan algoritma klasifikasi C4.5 memiliki nilai akurasi yang lebih tinggi dibandingkan dengan menggunakan algoritma klasifikasi C4.5. Nilai akurasi untuk model algoritma klasifikasi C4.5 sebesar 83,91 % dan nilai akurasi algoritma klasifikasi C4.5 berbasis PSO sebesar 83,49 % dengan selisih akurasi sebesar 0,42 %, dapat dilihat pada Tabel 2 di bawah ini:

Tabel 2. Pengujian Algoritma C4.5 dan Algoritma C4.5 berbasis PSO

| | Accuracy | AUC |
|-------------------|----------|-------|
| C4.5 | 83,91 % | 0,884 |
| C4.5 Berbasis PSO | 83,49 % | 0,889 |

Sumber: Sumber: (Lutfiyana, 2017)

Untuk evaluasi menggunakan ROC curve sehingga menghasilkan nilai AUC (*Area Under Curve*) untuk model algoritma C4.5 menghasilkan nilai 0.884 dengan nilai diagnosa *Good Classification*, sedangkan untuk algoritma C4.5 berbasis PSO (*Particle Swarm Optimization*) menghasilkan nilai 0.889 dengan nilai diagnosa *Good Classification*, dan selisih nilai keduanya sebesar 0.005. Dapat dilihat pada Gambar 2 dibawah ini.



Sumber: (Lutfiyana, 2017)

Gambar 2. Hasil ROC curve (Algoritma C4.5 dan Algoritma C4.5 berbasis PSO)

Pembahasan

Berdasarkan hasil eksperimen yang dilakukan untuk memecahkan masalah prediksi hasil layanan kemudahan donasi zakat dan program, dapat disimpulkan bahwa eksperimen menggunakan metode C4.5 dilakukan mulai dari eksperimen untuk menghitung dan mendapatkan rule-rule model algoritma, eksperimen langkah awal menghitung jumlah kasus, 3356 record, menghitung entropy dan nilai gain untuk semua atribut dilakukan, untuk mendapatkan nilai gain tertinggi. Gain jumlah mustahik: 0,273, gain bulan donasi 0,361, gain metode bayar 0,833, gain jumlah bulan entropy 0,05 yaitu 0,949, gain jumlah donasi dengan entropy 0,004 yaitu gain 0,995.

Eksperimen selanjutnya menggunakan metode C4.5 berbasis PSO dilakukan mulai dari eksperimen untuk menghitung dan mendapatkan rule-rule model algoritma, eksperimen langkah awal menghitung jumlah kasus, 3356 record, menghitung entropy dan nilai gain untuk semua atribut dilakukan, untuk mendapatkan nilai gain tertinggi. Gain jumlah mustahik, dengan entropy 0,940 yaitu gain 0,588, gain bulan donasi 0,202, gain metode bayar 0,882, gain jumlah bulan entropy 0,001 yaitu 0,998, gain jumlah donasi dengan entropy 0,001 yaitu gain 0,998.

Berdasarkan hasil eksperimen, dengan melakukan pengujian model untuk mengukur tingkat akurasi dan AUC. Model *Confusion Matrix* untuk Algoritma C4.5, bahwa untuk jumlah *True Positive* (TP) adalah 1507 untuk *False Negative* (FN) adalah 430, untuk *False Positive* (FP) adalah 110, dan untuk *True Negative* (TN) adalah 1309. Tingkat akurasi dengan menggunakan metode algoritma C4.5 adalah sebesar 83.91% nilai AUC sebesar 0.884 dengan nilai akurasi *Good Classification*. Model *Confusion Matrix* untuk Algoritma C4.5 berbasis PSO, bahwa untuk jumlah *True Positive* (TP) adalah 1538 untuk *False Negative* (FN) adalah 475, untuk *False Positive* (FP) adalah 79, dan untuk *True Negative* (TN) adalah 1264. Tingkat akurasi dengan menggunakan metode algoritma C4.5 berbasis PSO adalah sebesar 83.49% dan nilai AUC sebesar 0.889 dengan nilai akurasi *Good Classification*.

KESIMPULAN

Sebuah lembaga untuk memprediksi hasil layanan kemudahan donasi, salah satu algoritma data mining yang dapat digunakan adalah algoritma C4.5. Selain menghasilkan model sederhana yang mudah dimengerti, algoritma ini juga memiliki tingkat akurasi yang tinggi dalam proses klasifikasi donasi/transaksi. Algoritma C4.5 juga cukup handal dalam mengolah nilai numerik dan nominal.

Untuk mendapatkan model yang lebih akurat, algoritma PSO dapat diterapkan untuk proses pembobotan atribut. Pembobotan atribut ini diharapkan dapat memperbaiki atribut atribut numerik sehingga dapat memiliki nilai informasi yang lebih tinggi saat dihitung dengan algoritma C4.5. Dengan adanya penyempitan skala pada nilai numerik, atribut atribut numerik akan lebih mudah dipecah oleh algoritma C4.5. Model yang diperoleh pun menjadi lebih akurat dalam mengklasifikasikan loyalitas pelanggan.

Hasil penelitian untuk nilai akurasi yang didapat pada model algoritma C4.5 adalah 83,91% lebih baik jika dibandingkan dengan model algoritma C4.5 berbasis PSO adalah 83,49%. Dari hasil tersebut didapatkan selisih antara kedua model yaitu sebesar 0.42%. Sementara untuk evaluasi menggunakan ROC curve untuk kedua model yaitu, untuk model algoritma C4.5 nilai AUC adalah 0,884 dengan tingkat diagnosa *Good Classification*, dan untuk model algoritma C4.5 berbasis PSO nilai AUC adalah 0,889 dengan tingkat diagnosa *Good Classification*, selisih nilai AUC sebesar 0,005. Sehingga dapat disimpulkan bahwa hanya

penerapan algoritma C4.5 dapat meningkatkan nilai akurasi.

Dengan model prediksi kemudahan layanan donasi yang baik, lembaga zakat dapat meningkatkan layanan kepada donatur, serta dapat menggunakan model sebagai acuan dalam meningkatkan layanan berdasarkan zakat dan program.

1857

Tsai, C.-F., & Chen, M.-Y. (2010). Variable selection by association rules for customer churn prediction of multimedia on demand. *Expert Systems with Applications*, 37(3), 2006–2015.
<https://doi.org/10.1016/J.ESWA.2009.06.076>

REFERENSI

Aprillisa, R., & Wirawan, A. R. (2013). ANALISIS DAN PERANCANGAN SISTEM INFORMASI MANAJEMEN HUBUNGAN TERHADAP DONATUR UNTUK MENINGKATKAN LOYALITAS DAN JUMLAH DONATUR VIHARA DHAMMADIPA SURABAYA. *CALYPTRA*, 2(1), 1–14. Retrieved from <http://journal.ubaya.ac.id/index.php/jimus/article/view/120>

Aryani, D., & Rosinta, F. (2011). Pengaruh Kualitas Layanan Terhadap Kepuasan Pelanggan Dalam Membentuk Loyalitas Pelanggan. *BISNIS & BIROKRASI: Jurnal Ilmu Administrasi Dan Organisasi*, 7(2).

Gorunescu, F. (2011). *Data Mining Techniques and Models*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-19721-5_5

Hermansyah, A. (2011). *Strategi Pengembangan Donatur Pada Lembaga Amil Zakat Dengan Pendekatan Data Mining*. Universitas Indonesia. Retrieved from [http://lib.ui.ac.id/file?file=digital/20172095-T28731-Strategi pengembangan.pdf](http://lib.ui.ac.id/file?file=digital/20172095-T28731-Strategi%20pengembangan.pdf)

Lutfiyana, N. (2017). *Laproan Akhir Penelitian Mandiri*. Jakarta.

Munir, M. (2005). *Matra Dakwah Pengembangan Masyarakat, Pengembangan Sumber Daya Manusia* (1st ed.). Yogyakarta: Pustaka Pesantren.

Santosa, B., & Willy, P. (2011). *Metoda Metaheuristik, Konsep dan Implementasinya* (1st ed.). Surabaya: Guna Widya. Retrieved from <http://gunawidya.com/metode-metaheuristik-konsep-dan-implementasi/>

Sunjana. (2010). Seminar Nasional Aplikasi Teknologi Informasi (SNATI). In *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)* (pp. 24–29). Yogyakarta: UII. Retrieved from <http://journal.uui.ac.id/Snati/article/view/>

