

ALGORITMA KLASIFIKASI C4.5 BERBASIS *PARTICLE SWARM OPTIMIZATION* UNTUK PREDIKSI PENERIMA BANTUAN PANGAN NON TUNAI

Erni Ermawati

Program Studi Sistem Informasi Kampus Kabupaten Karawang, Fakultas Teknologi Informasi,
Universitas Bina Sarana Informatika
Jl. Banten No. 1 Karangpawitan Karawang, Telp (0267) 8454893
Email: erni.ert@bsi.ac.id

(Diterima: 29 Agustus 2019, direvisi: 30 Agustus 2019, disetujui: 09 September 2019)

ABSTRAK

Program Bantuan Pangan Non Tunai (BPNT) merupakan peralihan dari program RASTRA atau RASKIN. BPNT adalah bantuan sosial pangan dalam bentuk non tunai dari pemerintah yang diberikan kepada Keluarga Penerima Manfaat (KPM) setiap bulannya melalui mekanisme akun elektronik yang digunakan hanya untuk membeli bahan pangan di pedagang bahan pangan atau e-warong. Bahan pangan dalam program BPNT ini adalah beras dan atau telur. Salah satu kesulitan yang terkadang dihadapi oleh pemerintah dalam penyaluran BPNT adalah proses pembagian bantuan sosial yang tidak merata dan tidak tepat sasaran. Munculnya masalah tersebut diakibatkan kurangnya pertimbangan atau kematangan analisis penerima manfaat BPNT dalam menentukan kelayakan pada saat mengajukan pendataan. Oleh karena itu perlu dilakukan analisis lanjutan sehingga dapat mengetahui kelayakan dari suatu permasalahan prediksi penerima manfaat BPNT. Melalui hasil analisis pendataannya, dapat diketahui apakah warga layak atau tidak untuk menerima BPNT. Dari permasalahan yang ada digunakan metode klasifikasi untuk memprediksi kelayakan penerima manfaat BPNT yaitu dengan menggunakan dua model, model algoritma klasifikasi C4.5 dan model algoritma klasifikasi C4.5 berbasis *Particle Swarm Optimization* (PSO) tanpa *pruning*. Karena proses pengujian yang dilakukan tanpa *pruning* menghasilkan akurasi yang lebih tinggi dibandingkan dengan yang menggunakan *pruning*. Setelah dilakukan pengujian dengan dua model tersebut didapatkan hasil yaitu algoritma klasifikasi C4.5 menghasilkan nilai akurasi sebesar 98,56% dan nilai AUC sebesar 0,979 dengan tingkat diagnosa *Excellent Classification*, namun setelah dilakukan optimasi dengan algoritma klasifikasi C4.5 berbasis *Particle Swarm Optimization* nilai akurasi sebesar 98,92% dan nilai AUC sebesar 0,988 dengan tingkat diagnosa *Excellent Classification*. Sehingga kedua metode tersebut memiliki perbedaan tingkat akurasi sebesar 0,36%.

Keywords: Algoritma C4.5, Klasifikasi, Kelayakan BPNT, *Particle Swarm Optimization*.

1 PENDAHULUAN

Kemiskinan merupakan salah satu masalah yang dialami oleh beberapa negara berkembang, termasuk Indonesia. Banyak cara yang dilakukan untuk menanggulangi kemiskinan, diantaranya dengan program bantuan sosial untuk rakyat miskin [1]. Masalah kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun [2]. Pangan adalah kebutuhan pokok sekaligus menjadi aspek yang esensial bagi kehidupan manusia [3]. Ketahanan pangan merupakan salah satu isu paling strategis dalam pembangunan nasional, terlebih bagi negara berkembang seperti Indonesia [4].

Program Bantuan Pangan Non Tunai (BPNT) merupakan peralihan dari program RASTRA atau RASKIN. Program ini dimulai pada tahun 2017 di beberapa kota-kota di Indonesia dengan cara memberikan subsidi non tunai melalui sebuah kartu elektronik [5]. BPNT merupakan bantuan sosial pangan dalam bentuk non tunai dari pemerintah yang diberikan kepada Keluarga Penerima Manfaat (KPM) setiap bulannya melalui mekanisme akun elektronik yang digunakan hanya untuk membeli bahan pangan di pedagang bahan pangan atau *e-warong* [6]. Bahan pangan dalam program BPNT ini adalah beras dan atau telur [7].

Program BPNT memiliki kelebihan seperti kualitas sembako terutama beras jauh lebih bagus, ketersediaan jenis sembako lebih banyak dan masyarakat tidak perlu mengeluarkan uang untuk menebus sembako tersebut [5]. Calon penerima BPNT dipilih berdasarkan kriteria yang telah ditentukan oleh Kementerian Sosial Republik Indonesia. BPNT dilakukan oleh pemerintah untuk membantu seseorang yang kurang mampu dalam memenuhi kebutuhan sehari-hari. Untuk membantu menentukan siapa yang menerima BPNT diperlukan suatu metode yang dapat memberikan rekomendasi BPNT yang valid [8]. Salah satu kesulitan yang terkadang dihadapi oleh pemerintah adalah proses pembagian bantuan sosial yang tidak merata dan tepat sasaran [1]. Program E-warong ini secara khusus diluncurkan untuk mencegah distribusi bantuan pemerintah yang tidak tepat kualitas, kuantitas dan sasaran [9].

14 kriteria dari BPS untuk menentukan penerima BPNT yakni (1) luas lantai kurang dari 8 meter persegi, (2) jenis lantai, (3) jenis dinding, (4) fasilitas jamban, (5) sumber air untuk minum/memasak, (6) sumber penerangan, (7) bahan bakar, (8) dalam seminggu tidak pernah mengonsumsi daging, susu, atau hanya sekali dalam seminggu, (9) kemampuan membeli pakaian baru/tahun, (10) frekuensi makan/hari, (11) kemampuan berobat ke puskesmas/poliklinik, (12) pekerjaan (13) pendidikan tertinggi kepala keluarga, (14) tidak memiliki aset [8].

Beras adalah bahan makanan utama yang dikonsumsi masyarakat Indonesia sebagai sumber karbohidrat [10]. Beras adalah pangan pokok bangsa Indonesia yang dikonsumsi dan setiap tahun selalu mengalami peningkatan seiring dengan laju penambahan penduduk [11]. Beras sebagai bahan makanan pokok menyebabkan komoditas ini menjadi salah satu indikator pertumbuhan ekonomi maupun tingkat kemakmuran masyarakat [12]. Dengan demikian, beras merupakan bahan makanan pokok masyarakat Indonesia dan selalu mengalami peningkatan seiring dengan laju penambahan penduduk serta menjadi salah satu indikator pertumbuhan ekonomi maupun tingkat kemakmuran masyarakat.

Telur adalah makanan sumber protein hewani yang murah dan mudah untuk didapatkan oleh masyarakat Indonesia. Telur memiliki kandungan gizi yang lengkap, namun demikian telur merupakan bahan organik yang kualitasnya sangat terpengaruh pada kondisi lingkungan [13]. Telur salah satu bahan makanan hewani yang dikonsumsi selain daging, ikan dan susu. Umumnya telur yang dikonsumsi berasal dari jenis-jenis unggas, seperti ayam, bebek, dan angsa. Telur merupakan bahan makanan yang sangat akrab dengan kehidupan kita sehari-hari [14]. Oleh sebab itu, telur merupakan salah satu bahan makanan yang dikonsumsi sehari-hari yang memiliki kandungan gizi yang lengkap serta mengandung sumber protein hewani.

Penelitian sebelumnya untuk penentuan penerima bantuan pangan non tunai dilakukan dengan menggunakan salah satu metode dari *Fuzzy MADM* yaitu *Simple Additive Weighting* (SAW). Metode ini dipilih karena mampu menyeleksi alternatif terbaik dari sejumlah alternatif yang ada. Didapatkan hasil dari simulasi percobaan untuk metode ini cukup optimal untuk mendukung pengambilan keputusan dibuat orang B lebih miskin dibanding orang A dan C dan hasilnya orang B diprioritaskan untuk mendapatkan bantuan terlebih dahulu [8].

Dalam penelitian ini, dilakukan analisis algoritma klasifikasi data mining yaitu C4.5 dan C4.5 berbasis PSO yang dapat digunakan untuk membangun model klasifikasi. Algoritma tersebut menghasilkan akurasi model. Semakin tinggi nilai akurasi yang dihasilkan maka semakin akurat pula algoritma tersebut digunakan untuk menentukan kelayakan penerima bantuan pangan non tunai.

Algoritma C4.5 adalah algoritma terbaik [15]. Algoritma C4.5 adalah salah satu algoritma *decision tree* yang paling efektif untuk melakukan klasifikasi [16]. *Particle Swarm Optimization* (PSO) adalah teknik optimasi stokastik berbasis populasi yang dalam prosesnya dipengaruhi oleh sifat individu dan kelompok dalam solusi optimal [17]. Selain itu, PSO dikenal dapat meningkatkan akurasi berbagai macam klasifikasi data mining [18]. Berdasarkan penjelasan tersebut Algoritma C4.5 merupakan algoritma *decision tree* terbaik dan yang paling efektif untuk melakukan klasifikasi. Selain itu, PSO merupakan teknik optimasi stokastik berbasis populasi yang dapat meningkatkan akurasi berbagai macam klasifikasi data mining.

Tujuan dari penelitian ini adalah untuk membandingkan dan menentukan algoritma klasifikasi yang menghasilkan akurasi paling akurat yang nantinya akan digunakan untuk membuat sebuah sistem yang dapat membantu dalam memprediksi dan menentukan kelayakan penerima bantuan pangan non tunai (BPNT). Harapan dari penelitian ini adalah untuk dapat membantu pemerintah daerah dalam menentukan penerima BPNT.

2 TINJAUAN PUSTAKA

2.1. Kemiskinan

Kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun [2][19]. Kemiskinan merupakan salah satu masalah yang dialami oleh beberapa negara berkembang, termasuk Indonesia. Banyak cara yang dilakukan untuk menanggulangi kemiskinan, diantaranya dengan program bantuan sosial untuk rakyat miskin [1]. Berdasarkan penjelasan tersebut dapat ditarik kesimpulan bahwa kemiskinan adalah masalah atau persoalan yang dihadapi oleh banyak negara berkembang dan menjadi pusat perhatian di negara manapun.

2.2. Bantuan Pangan Non Tunai

Program Bantuan Pangan Non Tunai (BPNT) merupakan peralihan dari program RASTRA atau RASKIN. Program ini dimulai pada tahun 2017 di beberapa kota-kota di Indonesia dengan cara memberikan subsidi non tunai melalui sebuah kartu elektronik [5]. BPNT merupakan bantuan sosial pangan dalam bentuk non tunai dari pemerintah yang diberikan kepada Keluarga Penerima Manfaat (KPM) setiap bulannya melalui mekanisme akun elektronik yang digunakan hanya untuk membeli bahan pangan di pedagang bahan pangan atau e-warong [6]. Bahan pangan dalam program BPNT ini adalah beras dan atau telur [7]. Program BPNT memiliki kelebihan seperti kualitas sembako terutama beras jauh lebih bagus, ketersediaan jenis sembako lebih banyak dan masyarakat tidak perlu mengeluarkan uang untuk menebus sembako tersebut [5]. Calon penerima BPNT dipilih berdasarkan kriteria yang telah ditentukan oleh Kementerian Sosial Republik Indonesia. Untuk membantu menentukan siapa yang menerima BPNT diperlukan suatu metode yang dapat memberikan rekomendasi BPNT yang valid [8]. Salah satu kesulitan yang terkadang dihadapi oleh pemerintah adalah proses pembagian bantuan sosial yang tidak merata dan tepat sasaran [1]. Program E-warong ini secara khusus diluncurkan untuk mencegah distribusi bantuan pemerintah yang tidak tepat kualitas, kuantitas dan sasaran [9]. Kesimpulannya BPNT merupakan bantuan sosial pangan dalam bentuk non tunai seperti beras dan atau telur sebagai bentuk kepedulian dari pemerintah diberikan kepada calon penerima sesuai dengan kriteria yang telah ditentukan oleh Kementerian Sosial Republik Indonesia.

2.3. Ketahanan Pangan

Ketahanan pangan merupakan bagian terpenting dalam pemenuhan hak atas pangan sekaligus merupakan salah satu pilar utama hak azasi manusia. Ketahanan pangan juga merupakan bagian sangat penting dari ketahanan nasional [12]. Ketahanan pangan merupakan salah satu isu paling strategis dalam pembangunan nasional, terlebih bagi negara berkembang seperti Indonesia [4]. Kesimpulannya ketahanan pangan adalah bagian pemenuhan hak atas pangan dan bagian sangat penting dari ketahanan nasional serta pembangunan nasional suatu bangsa.

2.4. Beras

Beras adalah makanan pokok masyarakat mayoritas di Indonesia. Beras adalah komoditas strategis dan merupakan pangan pokok bangsa Indonesia. Konsumsi beras setiap tahun selalu meningkat seiring dengan laju penambahan penduduk [11]. Beras sebagai bahan makanan pokok menyebabkan komoditas ini menjadi salah satu indikator pertumbuhan ekonomi maupun tingkat kemakmuran masyarakat [12]. Dengan demikian, beras merupakan bahan makanan pokok masyarakat Indonesia dan selalu mengalami peningkatan seiring dengan laju penambahan penduduk serta menjadi salah satu indikator pertumbuhan ekonomi maupun tingkat kemakmuran masyarakat.

2.5. Telur

Telur adalah makanan sumber protein hewani yang murah dan mudah untuk didapatkan oleh masyarakat Indonesia. Telur memiliki kandungan gizi yang lengkap, namun demikian telur merupakan bahan organik yang kualitasnya sangat terpengaruh pada kondisi lingkungan [13]. Telur salah satu bahan makanan hewani yang dikonsumsi selain daging, ikan dan susu. Umumnya telur yang dikonsumsi berasal dari jenis-jenis unggas, seperti ayam, bebek, dan angsa. Telur merupakan bahan makanan yang sangat akrab dengan kehidupan kita sehari-hari [14]. Oleh sebab itu, telur merupakan salah satu bahan makanan yang dikonsumsi sehari-hari yang memiliki kandungan gizi yang lengkap serta mengandung sumber protein hewani.

Ermawati, Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai

2.6. Klasifikasi

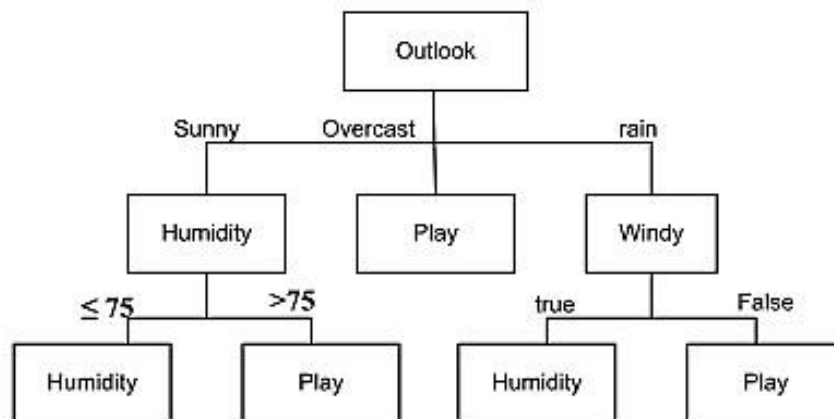
Klasifikasi adalah suatu teknik pembentukan model dari data yang belum terklasifikasi, untuk digunakan mengklasifikasi data baru. Klasifikasi termasuk ke dalam tipe supervised learning, artinya dibutuhkan data pelatihan untuk membangun suatu model klasifikasinya [14]. Klasifikasi merupakan topik yang 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 [14].

2.7. Data Mining

Data mining telah menarik banyak perhatian dalam dunia sistem informasi dan dalam masyarakat secara keseluruhan dalam beberapa tahun terakhir, karena ketersediaan luas dalam jumlah besar data dan kebutuhan segera untuk mengubah data tersebut menjadi informasi yang berguna dan pengetahuan [20]. Data mining bisa diartikan serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data. Tujuan data mining adalah untuk melakukan klasifikasi, klusterisasi, menemukan pola asosiasi hingga melakukan peramalan (*predicting*).

2.8. Algoritma C4.5

Algoritma C4.5 dan pohon keputusan merupakan dua model yang tak terpisahkan, karena untuk membangun sebuah pohon keputusan dibutuhkan algoritma C4.5 [21]. Pada dasarnya konsep dari algoritma C4.5 adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan (rule). C4.5 adalah algoritma yang cocok untuk masalah klasifikasi dan data mining, C4.5 memetakan nilai atribut menjadi class yang dapat diterapkan untuk klasifikasi baru. Seperti Gambar 1 Dibawah ini:



Gambar 1 Pohon keputusan

Pemilihan atribut sebagai simpul, baik simpul akar (root) atau simpul internal didasarkan pada nilai Gain tertinggi dari atribut-atribut yang ada. Penghitungan nilai Gain digunakan rumus seperti dalam Persamaan 1.

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

- S : Himpunan kasus
- A : Atribut
- n : Jumlah partisi himpunan atribut A
- |S_(i)| : Jumlah kasus pada partisi ke- i
- |S| : Jumlah kasus dalam S

Untuk menghitung nilai Entropy dapat dilihat pada Persamaan 2.

$$\text{Entropy} (S) = \sum_{i=1}^n - p_i * \log_2(p_i) \dots\dots\dots (2)$$

n : Jumlah partisi S

pi : Proporsi dari S_(i) terhadap S

2.9. Particle Swarm Optimization

Particle Swarm Optimization (PSO) merupakan metode optimisasi yang terinspirasi dari perilaku populasi ikan atau unggas dalam mencari sumber makanan [22]. *Particle Swarm Optimization* (PSO) memiliki proses algoritma yang lebih sederhana dibanding algoritma genetika [23]. Jadi PSO adalah metode optimisasi sederhana yang terinspirasi dari perilaku populasi ikan atau unggas dalam mencari sumber makanan.

2.10. Pengujian K-Fold Cross Validation

Cross Validation adalah teknik untuk mengevaluasi model dengan cara mempartisi sampel asli ke dalam training set untuk melatih model, dan test set untuk mengevaluasi model [24]. Pengujian ini dilakukan sebanyak k kali dengan mengganti-ganti partisi yang berfungsi sebagai data training dan data testing [25]. Secara umum pengujian nilai k dilakukan sebanyak 10 kali untuk memperkirakan akurasi estimasi. Dalam penelitian ini nilai k yang digunakan berjumlah 10 atau 10-fold *Cross Validation*.

2.11. Evaluasi dan Validasi Metode

Evaluasi yang dilakukan pada algoritma C4.5 dan algoritma C4.5 berbasis *particle swarm optimization* menggunakan model *confusion matrix*, dan *ROC curve* (*Receiver Operating Characteristic*).

a. Confusion Matrix

Confusion Matrix adalah *tools* yang digunakan untuk mengevaluasi model klasifikasi yang digunakan untuk memperkirakan objek yang benar dan yang salah. Hasil prediksi akan dibandingkan dengan kelas asli dari data tersebut. *Confusion Matrix* mengevaluasi kinerja model berdasarkan pada kemampuan akurasi prediktif suatu model [26]. *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 berikut [18]:

$$\text{Precision} = \text{TP}/(\text{TP}+\text{Fp})$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN})$$

$$\text{Accuracy} = (\text{TP}+\text{TN})/(\text{TP}+\text{TN}+\text{FP}+\text{FN})$$

Dimana :

TP : Jumlah kasus positif yang diklasifikasikan sebagai positif

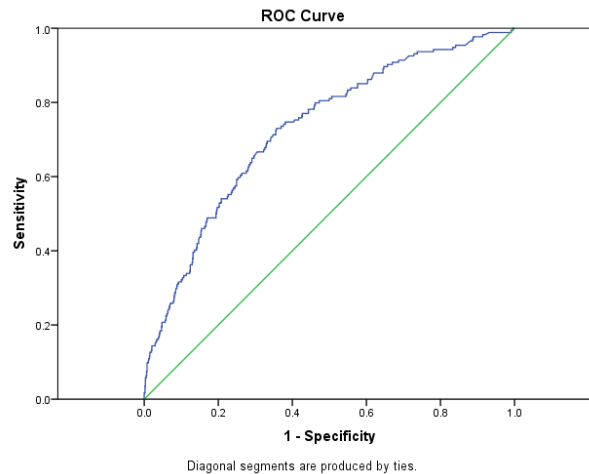
FP : Jumlah kasus negatif yang diklasifikasikan sebagai positif

TN : Jumlah kasus negatif yang diklasifikasikan sebagai negatif

FN : Jumlah kasus positif yang diklasifikasikan sebagai negatif

b. ROC Curve

Kurva ROC adalah perangkat dua dimensi yang digunakan untuk menilai kinerja klasifikasi yang menggunakan dua class keputusan, masing-masing objek dipetakan ke salah satu elemen dari himpunan pasangan, positif atau negatif [18].



Gambar 2 Grafik ROC

Pada Gambar 6 Garis diagonal membagi ruang ROC, yaitu:

1. Poin diatas garis diagonal merupakan hasil klasifikasi yang baik.
2. Poin dibawah garis diagonal merupakan hasil klasifikasi yang buruk.

Dapat disimpulkan bahwa, satu point pada kurva ROC adalah lebih baik dari pada yang lainnya jika arah garis melintang dari kiri bawah ke kanan atas didalam grafik. Menurut Gournescu (2011) dalam [16] Tingkat akurasi dapat didiagnosa sebagai berikut:

- a. Akurasi 0.90 – 1.00 = *Excellent classification*
- b. Akurasi 0.80 – 0.90 = *Good classification*
- c. Akurasi 0.70 – 0.80 = *Fair classification*
- d. Akurasi 0.60 – 0.70 = *Poor classification*
- e. Akurasi 0.50 – 0.60 = *Failure*

3 METODE PENELITIAN

1. Jenis penelitian

Dua pendekatan utama dalam penelitian yaitu pendekatan kualitatif dan pendekatan kuantitatif. Penelitian kualitatif adalah penelitian yang menampilkan beberapa fakta yang terjadi dari beberapa sumber dan menggunakan tipe penelitian deskriptif eksplanatif [11]. Sedangkan penelitian kuantitatif yaitu data yang dikumpulkan berupa angka-angka. Penelitian kuantitatif adalah penelitian ilmiah yang sistematis terhadap bagian-bagian dan fenomena serta hubungan-hubungannya [27]. Untuk itu dalam penelitian ini digunakan metode penelitian kuantitatif, dan metode ini dibagi menjadi tiga sub kategori yaitu inferentia, experimental, dan simulasi. Desain riset yang digunakan dalam penelitian ini adalah eksperimen. Desain eksperimen dibagi dua yaitu eksperimen absolut dan eksperimen komparatif [15]. Pada penelitian ini, jenis penelitian yang diambil adalah eksperimen komparatif. Dalam penelitian ini melakukan penerapan metode C4.5 dengan metode C4.5 yang berbasis PSO dalam hal pengujian metode akan dipilih salah satu metode yang paling baik tingkat akurasinya.

2. Metode Pemilihan Populasi dan Sampel

1. Populasi

populasi adalah keseluruhan obyek penelitian dengan batas-batas persoalan yang sudah cukup jelas [28]. Menurut sugiyono dalam [29] populasi adalah wilayah generalisasi yang terdiri atas obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya. Jadi populasi adalah keseluruhan objek penelitian dengan batas persoalan yang telah ditentukan sebelumnya untuk dipelajari dan kemudian menghasilkan suatu kesimpulan. Populasi dalam penelitian ini merupakan warga yang dinyatakan layak mendapatkan

BPNT dan warga yang tidak layak mendapatkan BPNT tahun 2016-2017 yang berasal dari data internal di salah satu kecamatan yang ada di Karawang Kota.

2. Sampel

Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi tersebut [30]. Penelitian menggunakan sampel yang diambil dari populasi. Apa yang dipelajari dari sampel itu, kesimpulannya akan dapat diberlakukan untuk populasi. Untuk itu sampel yang diambil dari populasi harus betul-betul representatif (mewakili). Sampel dari penelitian ini adalah data warga yang layak mendapatkan BPNT dan warga yang tidak mendapatkan BPNT. Data tersebut bersifat intern yang belum dipublikasikan oleh pihak kecamatan dapat dilihat pada Table 1 jumlah sampel yang diambil:

Tabel 1 Sampel Dataset

| Layak | Tidak Layak | Total Sampel |
|-------|-------------|--------------|
| 292 | 817 | 1109 |

3. Metode Pengumpulan Data

Metode pengumpulan data dibagi menjadi dua sumber data yaitu data primer dan data sekunder. Data Primer diperoleh melalui pertanyaan lisan dengan menggunakan metode wawancara, observasi, kuesioner dll [31]. Data sekunder adalah data dari sumber data lain atau data yang diperoleh dari obyek penelitian yaitu berupa dokumentasi [31] Data sekunder diperoleh dari instansi dan bacaan terkait penelitian [10]. Jadi data primer adalah data yang langsung diperoleh dari sumber dan data sekunder adalah data yang diperoleh dari sumber lain baik dokumentasi, literatur, buku, jurnal, dan informasi lainnya yang ada hubungannya dengan masalah yang diteliti. Dalam penelitian ini metode pengumpulan data untuk mendapatkan sumber data yang digunakan adalah metode pengumpulan data sekunder. Data utama diperoleh dari data warga yang layak mendapatkan BPNT dan warga yang tidak layak mendapatkan BPNT sedangkan data pendukung lainnya didapat dari buku, jurnal dan publikasi lainnya.

4. Instrumen Penelitian

1. Penelitian ini menggunakan data sekunder berupa data warga yang mendapatkan BPNT dan warga yang tidak layak mendapatkan BPNT yang akan digunakan sebagai instrumen guna memperoleh data dalam proses penentuan penerima BPNT.
2. Data disajikan dalam bentuk tabulasi model dan variabel masing- masing sebanyak 1109 warga terdiri dari warga layak sebanyak 292 dan warga tidak layak sebanyak 817 yang diambil dari tahun 2016/2017.
3. Perangkat lunak yang digunakan untuk menganalisis adalah Rapidminer dan Graphical User Interface (GUI) untuk menguji rule algoritma terpilih adalah PHP.

5. Metode Analisis dan Pengujian Data

Teknik Analisis data menggunakan data kuantitatif berupa matematika terhadap angka atau numerik dan nominal. Pada penelitian ini, analisis data dilakukan melalui data warga salah satu kecamatan yang ada di kota Karawang dengan nilai rata-rata warga yang mendapatkan BPNT dan warga yang tidak layak mendapatkan BPNT. Data diolah dan di uji dalam pengujian pada algoritma C4.5 dan algoritma C4.5 berbasis PSO. Kemudian pengujian Rule yang diperoleh algoritma C4.5 dan algoritma C4.5 berbasis PSO tersebut kemudian diuji dengan Confusion Matrix dan Kurva Receiver Operating Characteristic (ROC) untuk mengukur tingkat akurasi yang akan dihasilkan dari metode tersebut.

Dengan pengujian diatas dapat diperoleh rule algoritma yang mempunyai akurasi yang lebih baik. Sehingga dapat diterapkan pada Graphical User Interface (GUI) dengan baik. Selanjutnya aplikasi yang dibuat akan dievaluasi untuk menghasilkan pengetahuan (*knowledge*) baru.

Metode penelitian yang digunakan pada eksperimen ini adalah model Cross-Standard Industry for Data Mining (CRISP-DM) yang terdiri dari 6 fase (Brown, 2014), yaitu *Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, Deployment (using models in everyday business)*.

a. *Business Understanding*

Pada tahap business understanding bisa disebut juga tahap pemahaman penelitian, menentukan tujuan proyek penelitian dalam perumusan mendefinisikan masalah data mining. Berdasarkan kondisi warga periode Desember 2016, menunjukkan bahwa penerima manfaat bantuan pangan non tunai masih belum tepat sasaran. Banyaknya data warga yang seharusnya berhak mendapat bantuan akan tetapi mereka tidak mendapatkan bantuan pangan non tunai itu menunjukkan kurangnya analisis yang akurat dalam menentukan kelayakan pemberian penerima manfaat bantuan pangan non tunai. Teknik klasifikasi algoritma C4.5 digunakan dalam prediksi penerima bantuan pangan non tunai dengan tujuan analisis yang lebih akurat. Pada penelitian ini dilakukan pengembangan dengan algoritma klasifikasi C4.5 berbasis PSO (*Particle Swarm Optimization*) dengan tujuan meningkatkan akurasi dari perhitungan algoritma klasifikasi C4.5

b. *Data Understanding*

Pada tahap Data Understanding, dilakukan pengumpulan data, melakukan analisis penyelidikan data (data warga) untuk mengenali lebih lanjut data dan pencarian pengetahuan awal kemudian mengevaluasi kualitas dari data tersebut. Sehingga dapat diketahui persentase dari kelompok data yang akan dianalisis misalnya jumlah data yang akan diambil, jumlah data dengan keterangan layak atau tidak.

c. *Data Preparation*

Jumlah data yang diperoleh untuk penelitian ini sebanyak 1109 record warga yang layak menerima BPNT maupun yang tidak layak menerima. Untuk mendapatkan data yang berkualitas, beberapa teknik preprocessing digunakan (Vecellis, 2009), yaitu:

- 1) data validation, untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*), data yang tidak konsisten, dan data yang tidak lengkap (*missing value*).
- 2) data integration and transformation, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal dan kontinyu.
- 3) data size reduction and discretization, untuk memperoleh data set dengan jumlah atribut dan record yang lebih sedikit tetapi bersifat informatif. Di dalam data training yang digunakan dalam penelitian ini, dilakukan seleksi atribut dan penghapusan data duplikasi menggunakan software RapidMiner

d. *Modelling*

Pada tahap ini dilakukan pemrosesan data training sehingga akan menghasilkan beberapa aturan dan akan membentuk sebuah pohon keputusan. Model yang akan digunakan ada dua yaitu algoritma klasifikasi C4.5 dan algoritma klasifikasi C4.5 berbasis Particle Swarm Optimization (PSO).

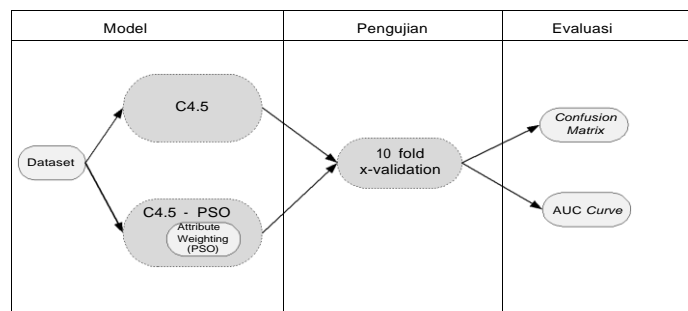
e. *Evaluation*

Pada tahap evaluasi, disebut tahap klasifikasi karena pada tahap ini akan ditentukan pengujian untuk akurasi. Tahap pengujiannya adalah melihat hasil akurasi pada proses klasifikasi Algoritma C4.5 dan klasifikasi Algoritma C4.5 berbasis Particle Swarm Optimization serta evaluasi dengan ROC Curve.

f. *Deployment*

Pada tahapan deployment, dilakukan penerapan model algoritma klasifikasi C4.5 berbasis Particle Swarm Optimization untuk menentukan penerima bantuan pangan non tunai (BPNT).

6. Model yang diusulkan



Gambar 3 Model algoritma yang diusulkan

Berdasarkan model eksperimen yang dijelaskan diatas, model algoritma yang dipakai dalam penelitian ini adalah:

1. Algoritma C4.5 yaitu model untuk mengubah data menjadi pohon keputusan dengan aturan-aturannya (rules).
2. *Particle Swarm Optimization* yaitu pencarian solusi yang terbaik dengan meningkatkan atribut (*attribute weight*), berupa partikel yang bergerak kearah posisi yang terbaik dari posisi sebelumnya.

4 HASIL DAN PEMBAHASAN

4.1. Pohon Keputusan C4.5

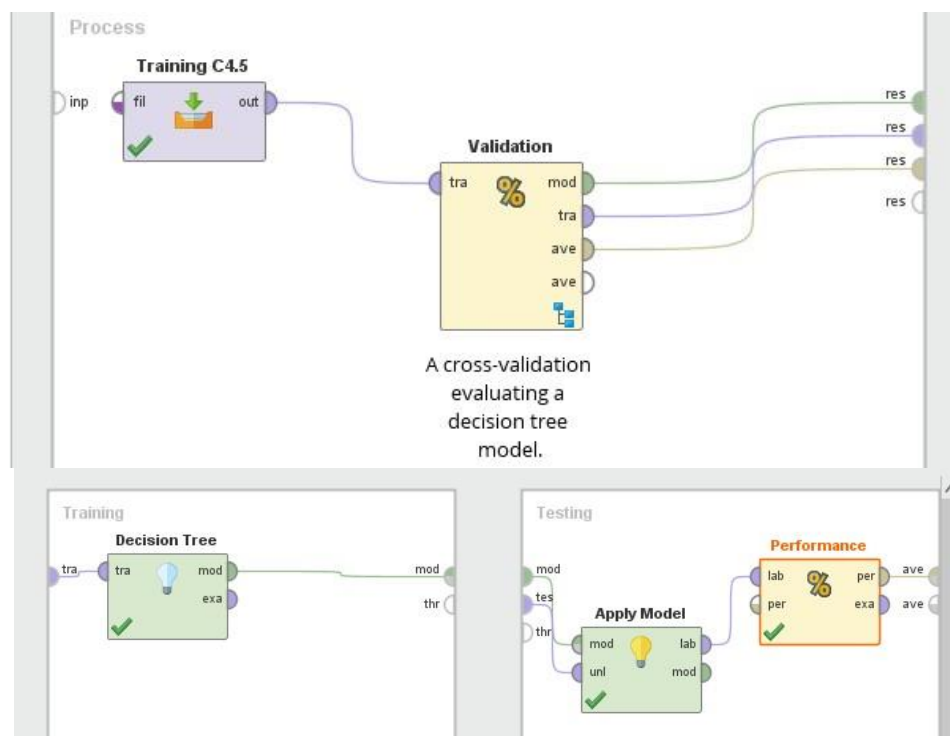
Adapun *rule* yang didapat dari penelitian adalah sebagai berikut:

1. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat PT/Akademik *THEN* Hasil = Tidak
2. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
3. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
4. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
5. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
6. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
7. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
8. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
9. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
10. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tidak/Belum Sekolah *THEN* Hasil = Layak
11. *IF* Sumber Air Minum = Sumur Terlindung/Pompa *THEN* Hasil = Layak
12. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Tidak *THEN* Hasil = Layak
13. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Fasilitas BAB = Jamban Bersama *THEN* Hasil = Layak
14. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Fasilitas BAB = Jamban Sendiri *THEN* Hasil = Tidak
15. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *THEN* Hasil = Layak
16. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Bambu *THEN* Hasil = Layak
17. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SD/MI *AND* Sumber Air Minum = Ledeng/Kemasan *THEN* Hasil = Tidak
18. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SD/MI *AND* Sumber Air Minum = Sumur Terlindung/Pompa *THEN* Hasil = Tidak
19. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
20. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
21. *IF* Pekerjaan = Pegawai Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Jenis Dinding = Tembok *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Tidak
22. *IF* Pekerjaan = Pensiunan *THEN* Hasil = Tidak
23. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak

Ermawati, Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai

24. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
25. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
26. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
27. *IF* Pekerjaan = PNS/TNI/POLRI *THEN* Hasil = Tidak
28. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *THEN* Hasil = Layak
29. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
30. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
31. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *AND* Aset = Tidak *THEN* Hasil = Layak
32. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *AND* Aset = Ya *THEN* Hasil = Layak
33. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak/Belum Sekolah *THEN* Hasil = Layak
34. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat PT/Akademik *THEN* Hasil = Tidak
35. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Tidak
36. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
37. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *THEN* Hasil = Tidak

Berikut adalah gambar pengujian menggunakan metode *K-Fold Cross Validation*:



Gambar 4 Pengujian *K-Fold cross validation* algoritma C4.5

4.2. Evaluasi Confusion Matrix C4.5

Tabel 2. Konversi confusion matrix algoritma klasifikasi C4.5

accuracy: 98.56% +/- 1.29% (mikro: 98.56%)

| | true LAYAK | true TIDAK | class precision |
|--------------|------------|------------|-----------------|
| pred. LAYAK | 278 | 2 | 99.29% |
| pred. TIDAK | 14 | 815 | 98.31% |
| class recall | 95.21% | 99.76% | |

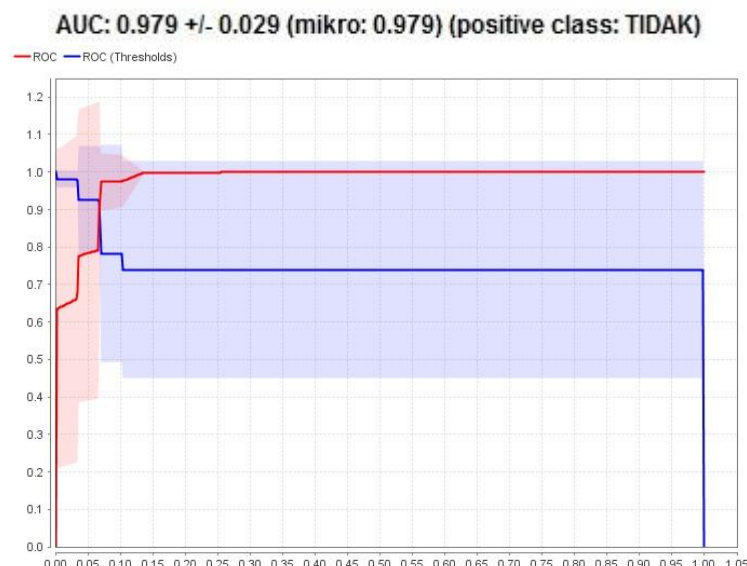
Jumlah *True Positive* (TP) adalah 278 *record* diklasifikasikan sebagai LAYAK dan *False Negative* (FN) sebanyak 14 *record* diklasifikasikan sebagai LAYAK tetapi TIDAK. Berikutnya 815 *record* untuk *True Negative* (TN) diklasifikasikan sebagai TIDAK, dan 2 *record False Positive* (FP) diklasifikasikan sebagai TIDAK ternyata LAYAK. Berdasarkan Tabel 3 tersebut menunjukkan bahwa, tingkat akurasi dengan menggunakan algoritma C4.5 adalah sebesar 98,56%.

Tabel 3 Hasil Perhitungan Algoritma C4.5

| | Nilai (%) |
|--------------------|-----------|
| <i>Accuracy</i> | 98,56 |
| <i>Sensitivity</i> | 95,21 |
| <i>Specitivity</i> | 99,76 |
| <i>PPV</i> | 99,29 |
| <i>NPV</i> | 98,31 |

Berdasarkan Tabel 3 menunjukkan bahwa, tingkat akurasi menggunakan algoritma klasifikasi C4.5 adalah sebesar **98,56%**.

4.3. Evaluasi dengan ROC C4.5



Gambar 5 Nilai AUC dalam grafik ROC algoritma C4.5

Pada Gambar 5 menunjukkan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,979 dengan tingkat diagnosa *Excellent Classification*.

4.4. Pohon Keputusan C4.5 berbasis PSO

Adapun *rule* yang didapat dari penggunaan metode C4.5 berbasis PSO adalah berikut:

1. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat PT/Akademik *THEN* Hasil = Tidak
2. *IF* Pekerjaan =Lainnya *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
3. *IF* Pekerjaan = Lainnya *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
4. *IF* Pekerjaan =Lainnya *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
5. *IF* Pekerjaan =Lainnya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
6. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
7. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
8. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
9. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
10. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum =Ledeng/Kemasan *AND* Pendidikan Terakhir KK = Tidak/Belum Sekolah *THEN* Hasil = Layak
11. *IF* Pekerjaan = Pedagang *AND* Sumber Air Minum = Sumur Terlindung/Pompa *THEN* Hasil = Layak
12. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Tidak *THEN* Hasil = Layak
13. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Masih PT/Akademik *THEN* Hasil = Tidak
14. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Masih SLTA/MA *AND* Sumber Air Minum = Ledeng/Kemasan *THEN* Hasil = Layak
15. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Masih SLTA/MA *AND* Sumber Air Minum = Sumur Terlindung/Pompa *THEN* Hasil = Tidak
16. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Tamat PT/Akademik *THEN* Hasil = Tidak
17. *IF* Pekerjaan = Pegawai Swasta *AND* Aset =Ya *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Tidak
18. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
19. *IF* Pekerjaan = Pegawai Swasta *AND* Aset = Ya *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
20. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *THEN* Hasil = Layak
21. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Tidak *THEN* Hasil = Layak
22. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Dinding = Bambu *THEN* Hasil = Layak
23. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SD/MI *AND* Sumber Air Minum = Ledeng/Kemasan *THEN* Hasil = Tidak
24. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset = Ya *AND* Dinding =Tembok *AND* Pendidikan Terakhir KK = Tamat SD/MI *AND* Sumber Air Minum =Sumur Terlindung/Pompa *THEN* Hasil = Tidak
25. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset =Ya *AND* Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
26. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset =Ya *AND* Dinding = Tembok *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak

27. *IF* Pekerjaan = Pekerja Lepas *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Aset =Ya *AND* Dinding = Tembok *AND* Pendidikan Terakhir KK = Tidak/Belum Sekolah *THEN* Hasil = Tidak
28. *IF* Pekerjaan = Pensiunan *THEN* Hasil = Tidak
29. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
30. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
31. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tamat SLTP/MTSN *THEN* Hasil = Tidak
32. *IF* Pekerjaan = Petani *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *THEN* Hasil = Layak
33. *IF* Pekerjaan = PNS/TNI/POLRI *THEN* Hasil = Tidak
34. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *THEN* Hasil = Layak
35. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Layak
36. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
37. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *AND* Aset = Tidak *THEN* Hasil = Layak
38. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak Tamat SD/MI *AND* Aset = Ya *THEN* Hasil = Layak
39. *IF* Pekerjaan = Tidak/Belum Bekerja *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *AND* Pendidikan Terakhir KK = Tidak/Belum Sekolah *THEN* Hasil = Layak
40. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat PT/Akademik *THEN* Hasil = Tidak
41. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat SD/MI *THEN* Hasil = Tidak
42. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Tidak *AND* Pendidikan Terakhir KK = Tamat SLTA/MA *THEN* Hasil = Tidak
43. *IF* Pekerjaan = Wiraswasta *AND* Membeli Daging,Ayam, Susu/Minggu = Ya *THEN* Hasil = Tidak

4.5. Evaluasi Confusion Matrix C4.5 Berbasis PSO

Tabel 4 Konversi Confusion Matrix Algoritma Klasifikasi C4.5 Berbasis PSO

accuracy: 98.92% +/- 1.13% (mikro: 98.92%)

| | true LAYAK | true TIDAK | class precision |
|--------------|------------|------------|-----------------|
| pred. LAYAK | 283 | 3 | 98.95% |
| pred. TIDAK | 9 | 814 | 98.91% |
| class recall | 96.92% | 99.63% | |

Jumlah *True Positive* (TP) adalah 283 *record* diklasifikasikan sebagai LAYAK dan *False Negative* (FN) sebanyak 9 *record* diklasifikasikan sebagai LAYAK tetapi TIDAK. Berikutnya 814 *record* untuk *True Negative* (TN) diklasifikasikan sebagai TIDAK, dan 3 *record* *False Positive* (FP) diklasifikasi sebagai TIDAK ternyata LAYAK. Berdasarkan Tabel 5 tersebut menunjukkan bahwa, tingkat akurasi dengan menggunakan algoritma C4.5 adalah sebesar 98,92.

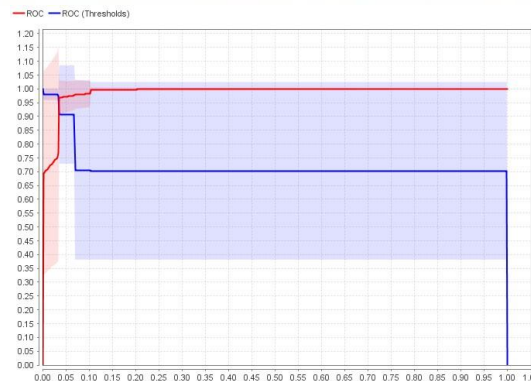
Tabel 5 Hasil Perhitungan Algoritma C4.5 Berbasis PSO

| | Nilai (%) |
|--------------------|-----------|
| <i>Accuracy</i> | 98,92 |
| <i>Sensitivity</i> | 96,92 |
| <i>Specitivity</i> | 99,63 |
| <i>PPV</i> | 98,95 |
| <i>NPV</i> | 98,91 |

Berdasarkan Tabel 5 menunjukkan bahwa, tingkat akurasi menggunakan algoritma klasifikasi C4.5 berbasis *particle swarm optimization* adalah sebesar **98,92%**.

4.6. Evaluasi dengan ROC C4.5 Berbasis PSO

AUC: 0.988 +/- 0.016 (mikro: 0.988) (positive class: TIDAK)



Gambar 6 Nilai AUC dalam grafik ROC algoritma C4.5 berbasis (PSO)

Pada Gambar 6 menunjukkan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,988 dengan tingkat diagnosa *Excellent Classification*.

5 KESIMPULAN

Dari hasil analisis optimasi algoritma C4.5 dengan *particle swarm optimization* (PSO) tanpa *pruning* dapat disimpulkan bahwa nilai akurasi yang didapat model algoritma C4.5 berbasis PSO adalah 98,92%. Lebih baik jika dibandingkan dengan model algoritma C4.5 menghasilkan nilai akurasi sebesar 98,56%. Dari hasil tersebut didapatkan selisih antara kedua model tersebut sebesar 0,36%. Sedangkan untuk evaluasi menggunakan ROC curve untuk kedua model tersebut memiliki selisih 0,009 dengan rincian model algoritma C4.5 menghasilkan nilai AUC sebesar 0,979 dan model algoritma C4.5 berbasis PSO menghasilkan nilai AUC sebesar 0,988 dengan tingkat diagnosa *Excellent Classification*. Sehingga dapat disimpulkan bahwa pengujian data tanpa *pruning* menghasilkan akurasi yang lebih besar dibandingkan pengujian yang menggunakan *pruning* dan penerapan teknik optimasi PSO dapat meningkatkan nilai akurasi pada algoritma klasifikasi C4.5 sehingga hasilnya dapat diterapkan menjadi sebuah sistem yang dapat membantu memprediksi penerima bantuan pangan non tunai (BPNT).

REFERENSI

- [1] D. Iskandar and Y. K. Suprpto, "PERBANDINGAN AKURASI KLASIFIKASI TINGKAT KEMISKINAN ANTARA ALGORITMA C 4.5 DAN NAÏVE BAYES," *Ilm. NERO*, vol. 2, no. 1, pp. 37–43, 2015.
- [2] I. Kurniawan and R. A. Saputra, "Penerapan Algoritma C5 . 0 Pada Sistem Pendukung Keputusan Kelayakan Penerimaan Beras Masyarakat Miskin," *J. Inform.*, vol. 4, no. 2, pp. 236–240, 2017.
- [3] E. K. Kiha and W. Rindayati, "KONVERGENSI HARGA PANGAN POKOK ANTAR WILAYAH DI INDONESIA," *J. Ekon. dan Kebijak. Pembang.*, vol. 2, no. 1, pp. 30–46, 2013.
- [4] S. Mariyani, F. E. Prasmatiwi, and R. Adawiyah, "KETERSEDIAAN PANGAN DAN FAKTOR-FAKTOR YANG MEMPENGARUHI KETERSEDIAAN PANGAN RUMAH TANGGA PETANI PADI ANGGOTA LUMBUNG PANGAN DI KECAMATAN AMBARAWA KABUPATEN PRINGSEWU," *JIIA*, vol. 5, no. 3, pp. 304–311, 2017.
- [5] M. S. Junaidi, B. M. Setiawan, and W. D. Prastiwi, "THE SATISFACTION COMPARISON OF BANTUAN PANGAN NON TUNAI RECIPIENTS AND RASTRA RECIPIENTS IN CAKUNG DISTRICT, EAST JAKARTA," *J. Ilm. Econosains*, vol. 15, no. 2, pp. 273–288, 2017.

- [6] T. A. Choesni *et al.*, *Pedoman Pelaksanaan Bantuan Pangan Non Tunai*. 2017.
- [7] P. Maharani, *Pedoman umum bantuan pangan non tunai*, 1st ed. Jakarta Pusat, 2017.
- [8] R. L. Fuady and A. M. Abadi, "Penentuan Penerimaan Bantuan Pangan Non Tunai (BPNT) Dengan Menggunakan Fuzzy Multiple Attribute Descission Making," in *SEMINAR MATEMATIKA DAN PENDIDIKAN MATEMATIKA*, 2017, pp. 203–210.
- [9] M. Yusnita, D. Agustina, and W. Anggita, "E-Warong sebagai Salah Satu Solusi Mengentaskan Kemiskinan Masyarakat," in *Proceeding of Community Development*, 2018, vol. 1, no. 2017, pp. 146–150.
- [10] F. I. Bazai, W. D. Sayekti, and D. A. H. Lestari, "PENERAPAN STRATEGI PEMASARAN DAN AKSESIBILITAS RUMAH TANGGA TERHADAP BIHUN TAPIOKA DI KOTA METRO," *JIIA*, vol. 5, no. 4, pp. 399–405, 2017.
- [11] C. S. Dewi and I. Pahlawan, "LANDASAN FAKTOR KEBIJAKAN IMPOR BERAS INDONESIA ASAL THAILAND PADA TAHUN 2009-2011," *JOM FISIP*, vol. 5, no. 1, pp. 1–11, 2018.
- [12] O. H. Putri and S. I. Gunawan, "Pengaruh produksi dan harga terhadap konsumsi beras di kabupaten kerinci," *J. Akrab Juara*, vol. 3, no. 1, pp. 59–69, 2018.
- [13] M. A. Djaelani, "Pengaruh Pencelupan pada Air Mendidih dan Air Kapur Sebelum Penyimpanan Terhadap Kualitas Telur Ayam Ras (Gallus L.)," *Bul. Anat. dan Fisiol.*, vol. XXIII, no. 1, pp. 24–30, 2015.
- [14] A. F. Basuki, D. Dr.Ir. Bambang Hidayat, and S. . Prof.Dr.Ir. Sjafril Darana, "DETEKSI KUALITAS DAN KESEGERAN TELUR BERDASARKAN SEGMENTASI WARNA DENGAN METODE FUZZY COLOR HISTOGRAM DAN WAVELET DENGAN KLASIFIKASI KNN QUALITY AND FRESHNESS DETECTION OF CHICKEN EGG BASED ON COLOR SEGMENTATION USING FUZZY COLOR HISTOGRAM (FCH) AND WAV," *e-proceeding Eng.*, vol. 3, no. 3, pp. 4404–4411, 2016.
- [15] A. Faisal, "KAJIAN PENERAPAN ALGORITMA C4.5, NAÏVE BAYE DAN NEURAL NETWORK UNTUK MEMENUHI PENILAIAN DATA KARYAWAN SERVICE LEVEL AGREEMENT DI BANK," *J. Fakt. Exacta*, vol. 10, no. 4, pp. 350–361, 2017.
- [16] A. Puspita and M. Wahyudi, "Algoritma C4.5 Berbasis Decision Tree untuk Prediksi Kelahiran Bayi Prematur," pp. 97–102, 2015.
- [17] R. Wajhillah, "Optimasi Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penyakit Jantung," *Swabumi*, vol. I, no. 1, pp. 26–36, 2014.
- [18] M. Y. Kurniawan and M. E. Rosadi, "Optimasi Decision Tree Menggunakan Particle Swarm Optimization Pada Data Siswa Putus Sekolah," *JTIULM*, vol. 2, no. 1, pp. 15–22, 2017.
- [19] S. Firdyana, D. Cahyadi, and I. F. Astuti, "Penerapan Metode Weighted Product Untuk Menentukan Penerima Bantuan Beras Masyarakat Miskin (Raskin)," *Pros. Semin. Ilmu Komput. Dan Teknol. Inf.*, vol. 2, no. 1, pp. 1–7, 2017.
- [20] W. D. Septiani, "Komparasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis," *Pilar Nusa Mandiri*, vol. 13, no. 1, pp. 76–84, 2017.
- [21] A. Purwanto and E. A. Darmadi, "PERBANDINGAN MINAT SISWA SMU PADA METODE KLASIFIKASI MENGGUNAKAN 5 ALGORITMA," *J. IKRAITH-INFORMATIKA*, vol. 2, no. 1, pp. 43–47, 2018.
- [22] D. Rahmalia, "Teknik penalti pada optimisasi berkendala menggunakan particle swarm optimization (," *J. Mat. dan Pendidik. Mat.*, vol. 3, no. 1, pp. 44–52, 18AD.
- [23] A. Setiawan, L. W. Santoso, and R. Adipranata, "Penerapan Algoritma Particle Swarm Optimization (PSO) untuk Optimisasi Pembangunan Negara dalam Turn Based Strategy Game," in *Penerapan Algoritma Particle Swarm Optimization (PSO) untuk Optimisasi Pembangunan Negara dalam Turn Based Strategy Game*, 2019, pp. 249–255.
- [24] G. A. Sandag, J. Leopold, and V. F. Ong, "Klasifikasi Malicious Websites Menggunakan Algoritma K-NN Berdasarkan Application Layers dan Network Characteristics Malicious Websites Classification Using K-NN Algorithm Based on Application Layers and Network Characteristics," *Cogito Smart J.*, vol. 4, no. 1, pp. 37–45, 2018.
- [25] R. A. Nugroho, Tarno, and A. Prahutama, "1 , 2 , 3 1," *Gaussian*, vol. 6, pp. 439–448, 2017.
- [26] A. Khoerunnisa, B. Irawan, and M. R. Rumani, "Analisis dan implementasi perbandingan

- algoritma c.45 dengan naïve bayes untuk prediksi penawaran produk,” in *e-proceeding of engineering*, 2016, vol. 3, no. 3, pp. 5029–5035.
- [27] A. Rahman and R. Y. Dewantara, “‘ PENGARUH KEMUDAHAN PENGGUNAAN DAN KEMANFAATAN TEKNOLOGI INFORMASI TERHADAP MINAT MENGGUNAKAN SITUS JUAL BELI ONLINE ’ (Studi Kasus Pada Pengguna Situs Jual Beli ‘ Z ’),” *J. Adm. Bisnis*, vol. 52, no. 1, p. 52, 2017.
- [28] E. Cahyani and D. Mulyadi, “PENGARUH KEPEMIMPINAN DAN MOTIVASI TERHADAP KINERJA PEGAWAI KANTOR CAMAT LUBUKLINGGAU SELATAN II KOTA LUBUKLINGGAU,” *INTERPROF (Jurnal Manajemen) Progr. Stud. Magister Manaj. STIE MURA*, vol. 3, no. 1, pp. 60–73, 2017.
- [29] A. Supriyatna, “Analisis Dan Evaluasi Kepuasan Pengguna Sistem Informasi Perpustakaan Dengan Menggunakan Pieces Framework,” *J. Pilar Nusa Mandiri*, vol. XI, no. 1, pp. 43–52, 2015.
- [30] M. Pransisya and B. Sudaryanto, “Analisis kualitas pelayanan dan harga produk yang berpengaruh terhadap kepuasan untuk mengurangi perpindahan merek pengguna xl axiata pada mahasiswa universitas diponegoro,” vol. 6, pp. 1–15, 2017.
- [31] M. Hot *et al.*, “PELAYANAN KEPEGAWAIAN (SAPK) BADAN KEPEGAWAIAN NEGARA SURABAYA Ayu Meirani Kusumawati,” pp. 1–10.