

Optimasi Naïve Bayes Classifier Menggunakan Particle Swarm Optimization Pada Klasifikasi Data Penerima Bantuan Sosial

Ari Kurniawan Saputra^{1*}, Wiwin Susanty^{2*}, Ahmad Shaleh Kurniawan³

¹ Informatika, Fakultas Ilmu Komputer, Universitas Bandar Lampung, Bandar Lampung

² Sistem Informasi, Fakultas Ilmu Komputer, Universitas Bandar Lampung, Bandar Lampung

³ Informatika, Fakultas Ilmu Komputer, Universitas Bandar Lampung, Bandar Lampung

* ari.kurniawan@ubl.ac.id, ² wiwin.susanty@ubl.ac.id, ³ ahmad.22421044@student.ubl.ac.id

ABSTRACT – Data classification in the case study of social assistance recipient data serves to classify based on labels or classes that indicate recommendations for social assistance recipients. Naive Bayes Classifier (NBC) is one of the classification methods or algorithms in data mining that has high speed. However, this assumption is not always correct so that it can affect the level of accuracy. This is due to the long time and prediction accuracy that is still not optimal. To overcome this problem, it is necessary to optimize NBC so that it can work more optimally in various dataset conditions. This research aims to optimize the Naive Bayes Classifier algorithm using Particle Swarm Optimization to improve accuracy in the classification process of social assistance recipients. Based on Naive Bayes optimization testing using PSO produces optimal values of model parameters that improve classification accuracy. In testing with 10 data, PSO successfully optimized Naive Bayes parameters so that 9 out of 10 data were correctly classified as beneficiaries, and 1 data was classified as non-recipients. The fitness value of the optimized model is close to 0.9 (90% accuracy), indicating that PSO is effective in adjusting parameters to improve Naive Bayes performance. It can be concluded that using the PSO algorithm, the Naive Bayes model becomes more accurate and reliable in data classification, especially in the context of decision making to determine social assistance recipients.

Keywords: Optimization; Particle Swarm Optimization; Naïve Bayes; Classification

ABSTRAK – Klasifikasi data pada studi kasus data penerima bantuan sosial berfungsi untuk mengklasifikasi berdasarkan label atau kelas yang menjadi indikasi rekomendasi penerima bantuan sosial. *Naive Bayes Classifier* (NBC) adalah salah satu metode atau algoritma klasifikasi dalam data mining yang memiliki kecepatan tinggi. Namun, asumsi tersebut tidak selalu tepat sehingga dapat mempengaruhi tingkat keakuratan. Hal ini disebabkan oleh waktu yang lama dan akurasi prediksi yang masih belum optimal. Untuk mengatasi masalah tersebut, perlu dilakukan optimasi terhadap NBC agar dapat bekerja lebih optimal dalam berbagai kondisi dataset. Penelitian ini bertujuan untuk mengoptimalkan algoritma *Naive Bayes Classifier* menggunakan *Particle Swarm Optimization* guna meningkatkan akurasi dalam proses klasifikasi penerima bantuan sosial. Berdasarkan pengujian optimasi *Naive Bayes* menggunakan PSO menghasilkan nilai optimal dari parameter model yang meningkatkan akurasi klasifikasi. Dalam pengujian dengan 10 data, PSO berhasil mengoptimalkan parameter *Naive Bayes* sehingga 9 dari 10 data diklasifikasikan dengan benar sebagai penerima bantuan, dan 1 data diklasifikasikan sebagai bukan penerima bantuan. Nilai *fitness* dari model yang dioptimalkan mendekati 0.9 (90% akurasi), menunjukkan bahwa PSO efektif dalam menyesuaikan parameter untuk meningkatkan performa *Naive Bayes*. Maka hal ini dapat disimpulkan bahwa penggunaan algoritma PSO, model *Naive Bayes* menjadi lebih akurat dan andal dalam klasifikasi data, khususnya dalam konteks pengambilan keputusan untuk menentukan penerima bantuan sosial.

Kata Kunci: Optimasi; *Particle Swarm Optimization*; *Naive Bayes*; Klasifikasi

1. PENDAHULUAN

Klasifikasi merupakan metode untuk membentuk model dari data yang belum memiliki klasifikasi, yang nantinya digunakan untuk mengklasifikasikan data baru [1]. Klasifikasi data pada studi kasus data penerima bantuan sosial berfungsi untuk mengklasifikasi berdasarkan label atau kelas yang menjadi indikasi rekomendasi penerima bantuan sosial. Salah satu metode

yang digunakan untuk klasifikasi adalah *Naive Bayes Classifier* (NBC).

Naive Bayes Classifier (NBC) adalah salah satu metode atau algoritma klasifikasi dalam data mining yang memiliki kecepatan tinggi, mudah diterapkan, serta memiliki struktur sederhana dan efisien dalam melakukan klasifikasi [2]. *Naive Bayes* didasari pada asumsi sederhana bahwa setiap nilai atribut secara kondisional tidak saling bergantung jika nilai output telah diketahui. Dengan kata



lain, probabilitas pengamatan bersama adalah hasil kali dari probabilitas masing-masing atribut, asalkan nilai output sudah diketahui [3][4]. Namun, asumsi tersebut tidak selalu tepat sehingga dapat mempengaruhi tingkat keakuratan [5]. Hal ini disebabkan oleh waktu yang lama dan akurasi prediksi yang masih belum optimal [6]. Untuk mengatasi masalah tersebut, perlu dilakukan optimasi terhadap NBC agar dapat bekerja lebih optimal dalam berbagai kondisi *dataset*. Salah satu metode optimasi yang populer digunakan adalah *Particle Swarm Optimization* (PSO).

Particle Swarm Optimization (PSO) merupakan metode optimasi berbasis stokastik yang melibatkan populasi, terinspirasi oleh kecerdasan kolektif dari kelompok hewan seperti sekumpulan burung atau ikan [7]. Metode *Particle Swarm Optimization* menawarkan keuntungan berupa konsep yang sederhana, kemudahan dalam penerapan, serta efisiensi komputasi yang lebih tinggi jika dibandingkan dengan algoritma matematis dan metode optimasi lainnya [8].

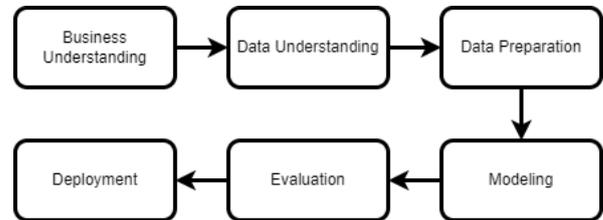
Berdasarkan penelitian yang dilakukan M. Rangga Ramadhan Saelan, *et al* tahun 2020 menyatakan bahwa *Naïve Bayes* hanya mencapai akurasi sebesar 82,60%. Berdasarkan analisis ROC, dapat disimpulkan bahwa kinerja *Naïve Bayes* tergolong kurang memuaskan, dengan tingkat *Equal Error Rate* (EER) mencapai 65% [9]. Berdasarkan penelitian yang dilakukan oleh Fitriana Sholehah, *et al* tahun 2024 menyatakan bahwa algoritma *Naïve Bayes* mencapai akurasi 79%, sementara algoritma *K-Nearest Neighbors* (KNN) mencatat akurasi tertinggi sebesar 82%. Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma K-NN dengan pembagian data 50:50 lebih efisien dalam melakukan prediksi dan klasifikasi data [10]. Berdasarkan penelitian yang dilakukan oleh Anggi Tri Dewi Septiani, *et al* tahun 2024 menyatakan bahwa berdasarkan analisis, metode *K-Nearest Neighbor* terbukti lebih efektif dalam menganalisis sentimen, dengan hasil *confusion matrix* yang menunjukkan akurasi *K-Nearest Neighbor* mencapai 84,06% dan 83,06%. Sementara itu, *Naïve Bayes Classifier* mendapatkan akurasi sebesar 83,06% dan 82,56% [11].

Berdasarkan uraian penelitian-penelitian sebelumnya, maka penelitian ini bertujuan untuk mengoptimalkan algoritma *Naïve Bayes Classifier* menggunakan *Particle Swarm Optimization* guna meningkatkan akurasi dalam proses klasifikasi penerima bantuan sosial. Optimasi ini dilakukan karena NBC sering kali mengalami penurunan performa terutama saat menghadapi dataset yang tidak independen atau memiliki ketidakseimbangan kelas. Dengan menerapkan PSO, diharapkan dapat dicapai parameter optimal yang mampu meningkatkan akurasi, presisi, dan *recall*, sehingga dapat mengurangi kesalahan klasifikasi, memastikan bantuan yang didistribusikan tepat sasaran.

2. METODOLOGI

Metode penelitian yang digunakan dalam penelitian ini adalah metode CRISP-DM. CRISP-DM atau *Cross-*

industry Standard Process for Data Mining, adalah sebuah metode yang memberikan panduan standar dalam proses data mining. Metode ini dapat digunakan dalam strategi pemecahan masalah umum di dunia bisnis maupun penelitian, yang terdiri dari beberapa tahapan [12]. Tahapan metode CRISP-DM diilustrasikan pada Gambar 1 sebagai berikut [13].



Gambar 1. Metode Penelitian CRISP-DM

1. *Business Understanding*

Pada tahap ini, beberapa langkah dilakukan seperti memahami kebutuhan dan tujuan dari perspektif bisnis, kemudian menerjemahkan pengetahuan tersebut ke dalam bentuk perumusan masalah dalam *data mining*, serta merancang rencana dan strategi untuk mencapai tujuan *data mining* [14]. Tahapan ini bertujuan untuk memahami tujuan dan kebutuhan bisnis, serta untuk menilai kelayakan proyek tersebut apakah dapat dilaksanakan atau tidak [15].

2. *Data Understanding*

Data understanding merupakan pemahaman kebutuhan data untuk mencapai tujuan dalam merumuskan strategi penelitian yang efektif dan efisien [16]. Tahap ini bertujuan untuk memahami pola dan struktur data, mengingat data teks adalah data yang tidak terstruktur. Oleh karena itu, akan dilakukan beberapa proses identifikasi, seperti mengidentifikasi data yang hilang dan duplikasi, menghitung jumlah data untuk setiap kelas, serta menganalisis frekuensi kata dalam dataset [17].

3. *Data Preparation*

Tahap *data preparation* mencakup tahap *preprocessing* yang harus dilakukan. Salah satu tahapan *preprocessing* dalam persiapan data adalah *data cleaning*, yang bertujuan untuk menghapus *missing value* pada satu atau beberapa atribut [18].

4. *Modeling*

Tahap *modeling* merupakan langkah krusial dalam penerapan pendekatan *data mining*. Proses ini melibatkan pembuatan model dengan menggunakan algoritma dari teknik tertentu, seperti aturan asosiasi, untuk diterapkan pada *dataset* yang telah disiapkan [19].

5. *Evaluation*

Pada tahap *evaluation*, dilakukan analisis terhadap hasil dari pengolahan data yang telah dilaksanakan. Interpretasi hasil juga merupakan bagian dari tahap evaluasi ini. *Evaluation* bertujuan untuk menilai kecocokan model yang digunakan selama proses pembelajaran data [20].

6. *Deployment*

Tahap *deployment*, atau penyebaran, merujuk pada proses pembuatan laporan atau presentasi berdasarkan pengetahuan yang diperoleh selama evaluasi dalam data mining [21].

Penerapan metode penelitian mengacu pada Gambar 1, metode CRISP-DM di terapkan dalam penelitian ini sebagai berikut.

1. *Business Understanding*

Pada tahap ini, dilakukan dengan menguraikan permasalahan dalam penelitian, metode pemecahan masalah penelitian, dan penentuan kebutuhan data. Tahap ini di deskripsikan pada Tabel 1 sebagai berikut

3. HASIL DAN PEMBAHASAN

Tabel 1. Tahap *Business Understanding*

Business Understanding	Keterangan
Permasalahan dalam Penelitian	<i>Naive Bayes</i> didasari pada asumsi sederhana bahwa setiap nilai atribut secara kondisional tidak saling bergantung jika nilai output telah diketahui. Namun, asumsi tersebut tidak selalu tepat sehingga dapat mempengaruhi tingkat keakuratan. Hal ini disebabkan oleh waktu yang lama dan akurasi prediksi yang masih belum optimal
Metode Pemecahan Masalah Penelitian	Metode yang digunakan adalah metode optimasi <i>Particle Swarm Optimization</i> yang di implementasikan untuk optimasi klasifikasi <i>Naive Bayes Classifier</i> . Berikut formula atau rumus PSO dan NBC yang diterapkan dalam penelitian ini [22]
	a. Rumus <i>Naive Bayes Classifier</i>

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad \text{(Rumus 1)}$$

Keterangan:

X = Data dengan kelas yang belum diketahui

H = Hipotesis data X merupakan suatu kelas spesifik

$P(H|X)$ = Probabilitas hipotesis H berdasar kondisi X (*posterior probability*)

$P(H)$ = Probabilitas hipotesis H (*prior probability*)

$P(X|H)$ = Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

b. Rumus *Particle Swarm Optimization*

$$V_{i,m}^{new} = W \cdot V_{i,m}^{old} + C_1 x (P_{i,m}^{local\ best} - X_{i,m}^{old}) + C_2 R x (P_{i,m}^{global\ best} - X_{i,m}^{old}) \quad \text{(Rumus 2)}$$

Mengukur kecepatan terkini masing-masing partikel:

$$X_{i,m}^{new} = X_{i,m}^{old} + V_{i,m}^{new} \quad \text{(Rumus 3)}$$

Keterangan:

n = Jumlah partikel dalam kelompok

d = Dimensi

V_i = Kecepatan partikel ke- i pada iterasi ke- i

C_1, C_2 = Konstanta akselerasi (*learning rate*)

R = Bilangan random (0-1)



Penentuan Kebutuhan Data	<p>X, = Posisi saat ini dari partikel ke-i</p> <p>P_{best_i} = Posisi terbaik sebelumnya dari partikel ke-i</p> <p>P_{gbest} = Partikel terbaik dari semua partikel dalam satu kelompok atau populasi</p> <p>Data yang dibutuhkan dalam penelitian ini adalah data calon penerima bantuan sosial</p>
--------------------------	---

2. Data Understanding

Pada tahap ini, data yang digunakan adalah data Penyasaran Percepatan Penghapusan Kemiskinan Ekstrem (P3KE). Data tersebut merupakan data pengajuan calon penerima bantuan sosial dari salah satu kabupaten di Provinsi Lampung tahun 2024. Parameter yang digunakan seperti Jenis Kelamin, Status Marital, Pekerjaan, Pendidikan, dan Usia. Verifikasi terhadap klasifikasi rekomendasi penerima bantuan sosial dikategorikan menjadi Penerima Bantuan dan Bukan Penerima Bantuan. *Data understanding* pada tahap ini di deskripsikan pada Tabel 2. Pada Tabel 3 merupakan dataset yang digunakan dalam tahap *Data Understanding*.

Tabel 2. Kategori Penerima Bantuan Sosial

Kategori Penerima	Kelayakan	Kategori Kemiskinan
Penerima Bantuan		1 = Sangat Miskin 2 = Miskin 3 = Rentan Miskin
Bukan Penerima Bantuan		4 = Menengah Kebawah 5 Penerima Tidak Layak

Jumlah dataset yang digunakan dalam artikel penelitian ini merupakan jumlah data yang tidak sebenarnya. Jumlah data yang sebenarnya digunakan dalam penelitian ini adalah 10004 *record*

Tabel 3. Dataset Calon Penerima Bantuan

ID Individu	Jenis Kelamin	Status Marital	Pekerjaan	Pendidikan	Usia 2024
1	Laki-laki	Belum kawin	Pekerja Lepas	Tamat SLTA/ sederajat	30
2	Laki-laki	Kawin	Pekerja Lepas	Tamat SLTA/ sederajat	53
3	Perempuan	Kawin	Pekerja Lepas	Tamat SD/ sederajat	51
4	Perempuan	Belum kawin	Pekerja Lepas	Tamat SLTA/ sederajat	24
5	Laki-laki	Belum kawin	Belum/ Tidak Bekerja	Masih SLTA/ sederajat	19
6	Laki-laki	Kawin	Pensiunan	Tamat SLTA/ sederajat	67
7	Perempuan	Kawin	Belum/ Tidak Bekerja	Tamat SLTP/ sederajat	64
8	Laki-laki	Belum kawin	Pegawai Negeri	Masih PT/ akademi	24
9	Laki-laki	Kawin	Pekerja Lepas	Tamat SLTA/ sederajat	37
10	Perempuan	Kawin	Belum/ Tidak Bekerja	Tamat SLTA/ sederajat	36

3. Data Preparation

Pada tahap *data preparation*, dilakukan pemilihan data dan transformasi data. Data yang telah digabungkan akan diproses lebih lanjut pada tahap *data preparation*. Pada *data preparation* memiliki proses yang dilakukan, yaitu sebagai berikut:

a. Data Selection

Proses *data selection* dilakukan karena tidak semua data akan relevan untuk digunakan, sehingga perlu dilakukan seleksi data. Data yang dipilih adalah yang sesuai dengan analisis

yang akan diterapkan. Proses yang terjadi pada *data selection* akan menghilangkan kolom ID

Individu, yang ditunjukkan pada Tabel 4 berikut ini

Tabel 4. Proses *Data Selection*

Jenis Kelamin	Status Marital	Pekerjaan	Pendidikan	Usia 2024
Laki-laki	Belum kawin	Pekerja Lepas	Tamat SLTA/ sederajat	30
Laki-laki	Kawin	Pekerja Lepas	Tamat SLTA/ sederajat	53
Perempuan	Kawin	Pekerja Lepas	Tamat SD/ sederajat	51
Perempuan	Belum kawin	Pekerja Lepas	Tamat SLTA/ sederajat	24
Laki-laki	Belum kawin	Belum/ Tidak Bekerja	Masih SLTA/ sederajat	19
Laki-laki	Kawin	Pensiunan	Tamat SLTA/ sederajat	67
Perempuan	Kawin	Belum/ Tidak Bekerja	Tamat SLTP/ sederajat	64
Laki-laki	Belum kawin	Pegawai Negeri	Masih PT/ akademi	24
Laki-laki	Kawin	Pekerja Lepas	Tamat SLTA/ sederajat	37
Perempuan	Kawin	Belum/ Tidak Bekerja	Tamat SLTA/ sederajat	36

b. Data Transformation

Data yang telah dipilih akan diproses lebih lanjut pada tahap *data transformation*. Proses *data transformation* ini melibatkan klasifikasi data ke dalam beberapa kategori yang telah ditentukan. Berikut adalah tabel pengklasifikasian data:

1. Kriteria Jenis Kelamin

Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 2 nilai kriteria yaitu 1 dan 2. Pada Tabel 5 merupakan Kriteria Jenis Kelamin yang telah di klasifikasikan

Tabel 5. Klasifikasi Kriteria Jenis Kelamin

Kriteria	Kategori Kriteria	Nilai Kriteria
Jenis Kelamin	Laki-laki	1
	Perempuan	2

2. Kriteria Hubungan Keluarga

Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 4 nilai kriteria yaitu 1,2,3 dan 4. Pada Tabel 6 merupakan Kriteria Hubungan Keluarga yang telah di klasifikasikan

Tabel 6. Klasifikasi Kriteria Hubungan Keluarga

Kriteria	Kategori Kriteria	Nilai Kriteria
Hubungan keluarga	Kepala	1
	Keluarga Suami/Istri	2
	Anak	3
	Anggota	4
	Keluarga Lainnya	

3. Kriteria Status Marital

Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 3 nilai kriteria yaitu 1,2 dan 3. Pada Tabel 7 merupakan Kriteria Status Marital yang telah di klasifikasikan

Tabel 7. Klasifikasi Kriteria Status Marital

Kriteria	Kategori Kriteria	Nilai Kriteria
Status Marital	Belum Kawin	1
	Kawin	2
	Cerai	3



4. Kriteria Pekerjaan
Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 4 nilai kriteria yaitu 1,2,3, dan 4. Pada Tabel 8 merupakan Kriteria Pekerjaan yang telah di klasifikasikan

Tabel 8. Klasifikasi Kriteria Pekerjaan

Kriteria	Kategori Kriteria	Nilai Kriteria
Pekerjaan	Belum/Tidak Bekerja	1
	Pekerja Lepas	2
	Pegawai	3
	Pensiunan	4

5. Kriteria Pendidikan
Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 5 nilai kriteria yaitu 1,2,3,4, dan 5. Pada Tabel 9 merupakan Kriteria Pendidikan yang telah di klasifikasikan

Tabel 9. Klasifikasi Kriteria Pendidikan

Kriteria	Kategori Kriteria	Nilai Kriteria
Pendidikan	Tidak Sekolah	1
	SD	2
	SMP	3
	SMA	4
	Perguruan Tinggi	5

6. Kriteria Usia
Variabel ini berisi range nilai yang telah diklasifikasikan menjadi 4 nilai kriteria yaitu 1,2,3 dan 4. Pada Tabel 10 merupakan Kriteria Usia yang telah di klasifikasikan

Tabel 10. Klasifikasi Kriteria Usia

Kriteria	Kategori Kriteria	Nilai kriteria
Usia	> 60 Tahun	1
	41-60 Tahun	2
	18-40 Tahun	3
	< 18 Tahun	4

4. Modeling

Pada tahap ini, data yang telah diklasifikasikan akan diproses lebih lanjut melalui beberapa pemodelan. Proses pemodelan ini bertujuan

untuk mengoptimalkan hasil yang diperoleh. Berikut merupakan tahapan optimasi *Naive Bayes* menggunakan *Particle Swarm Optimization*. Pada Langkah ini, algoritma PSO digunakan untuk mengoptimasi parameter pada *Naive Bayes*, dengan tujuan untuk mengoptimalkan hasil klasifikasi data. Berikut merupakan langkah-langkah penerapan algoritma PSO dalam optimasi parameter *Naive Bayes*:

1. Inisiasi Partikel dan Kecepatan
 - a. Inisialisasi Posisi Partikel (x_i):
Dalam penelitian ini variable partikel ditentukan sebagai berikut

$$x_{min} = 10$$

$$x_{max} = 50$$

Maka nilai acak (random) yang dihasilkan = 0.25

Selanjutnya hitung posisi awal $x_i(0)$ sebagai berikut:

$$x_i(0) = 0.25 \times (50 - 10) + 10$$

$$x_i(0) = 0.25 \times 40 + 10$$

$$x_i(0) = 10 + 10$$

$$x_i(0) = 20$$

Jadi, posisi awal partikel $x_i(0)$ adalah 20.

- b. Inisialisasi Kecepatan Partikel (v_i):
Dalam penelitian ini variabel kecepatan partikel ditentukan sebagai berikut

$$v_{min} = -5$$

$$v_{max} = 10$$

Maka nilai acak (random) yang dihasilkan = 0.6

Selanjutnya hitung kecepatan awal $v_i(0)$ sebagai berikut

$$v_i = 0.6 \times (10 - (-5)) + (-5)$$

$$v_i = 0.6 \times 15 + (-5)$$

$$v_i = 9 - 5$$

$$v_i = 4$$

Jadi, kecepatan awal partikel $v_i(0) = 4$

2. Evaluasi *Fitness* dengan *Naive Bayes*
Evaluasi setiap partikel dengan menghitung akurasi *Naive Bayes* berdasarkan parameter yang diwakili oleh posisi partikel. Dalam penelitian ini berdasarkan *dataset* pada Tabel 4, hasil klasifikasi *Naive Bayes* adalah 9 data yang dinyatakan sebagai Penerima Bantuan

dan 1 data dinyatakan Bukan Penerima. Berikut adalah tahapan Evaluasi *fitness* dengan *Naive Bayes*

a. *Probabilitas Prior*

$$P(\text{Penerima Bantuan}) = \frac{9}{10} = 0.9$$

$$P(\text{Bukan Penerima Bantuan}) = \frac{1}{10} = 1.0$$

b. *Likelihood*

Likelihood Jenis Kelamin

$$P(\text{Laki - laki} | \text{Penerima Bantuan}) = \frac{5}{9} = 0.56$$

$$P(\text{Laki - laki} | \text{Bukan Penerima Bantuan}) = \frac{1}{1} = 1.0$$

Likelihood Status Marital

$$P(\text{Belum Kawin} | \text{Penerima Bantuan}) = \frac{3}{9} = 0.33$$

$$P(\text{Belum Kawin} | \text{Bukan Penerima Bantuan}) = \frac{1}{1} = 1.0$$

Likelihood Pekerjaan

$$P(\text{Pegawai Negeri} | \text{Penerima Bantuan}) = \frac{0}{9} = 0$$

$$P(\text{Pegawai Negeri} | \text{Bukan Penerima Bantuan}) = \frac{1}{1} = 1.0$$

Likelihood Pendidikan

$$P(\text{Masih PT} / \text{Akademi} | \text{Penerima Bantuan}) = \frac{0}{9} = 0$$

$$P(\text{Masih PT} / \text{Akademi} | \text{Bukan Penerima Bantuan}) = 1.0$$

Tabel 11. Perhitungan Iterasi dan Optimalisasi dalam Proses PSO

Iterasi	v_i (Kecepatan)	Inersia ($w \cdot v$)	Komponen Kognitif ($c1 \cdot r1 \cdot (P_{best} - x)$)	Komponen Sosial ($c2 \cdot r2 \cdot (G_{best} - x)$)	x_i Posisi	Posisi Baru ($x + v$)
1	2	0.52=1.0	1.5·0.8·(6-5)=1.2	1.5·0.8·(7-5)=2.4	5	5+4.6=9.6
2	4.6	Dihitung ulang	Dihitung ulang	Dihitung ulang	9.6	Dihitung ulang

$$P(\text{Bantuan}) = \frac{1}{1} = 1.0$$

c. *Probabilitas Posterior*

Untuk penerima:

$$P(\text{Penerima Bantuan} | \text{Fitur}) = 0.9 \times 0.56 \times 0.33 \times 0 \times 0 = 0$$

Untuk Bukan Penerima

$$P(\text{Bukan Penerima} | \text{Fitur}) = 0.1 \times 1.0 \times 1.0 \times 1.0 \times 1.0 = 1$$

Hasil = Karena probabilitas *posterior* untuk Bukan Penerima (0.1) lebih tinggi dibandingkan dengan Penerima (0), individu ini diklasifikasikan sebagai Bukan Penerima Bantuan.

3. *Update Local Best* dan *Global Best*

Asumsi awal:

Fitness Partikel Saat Ini=85

Local Best (p_i) sebelumnya=80

Global Best (g) sebelumnya=82

a. *Update Local Best* (p_i):

Karena *fitness* partikel saat ini (85) lebih baik dari nilai *local best* sebelumnya (80):

$$p_i = 85$$

b. *Update Global Best* (g)

Local Best yang baru = 85

Global Best sebelumnya = 82

Karena *local best* (85) lebih baik dari *global best* (82), maka:

$$g = 85$$

Maka:

Local Best p_i baru = 85

Global Best g = baru = 85

4. Iterasi Hingga Konvergensi atau Kondisi Berhenti

4. KESIMPULAN

Berdasarkan pengujian optimasi *Naive Bayes* menggunakan PSO menghasilkan nilai optimal dari parameter model yang meningkatkan akurasi klasifikasi. Dalam pengujian dengan 10 data, PSO berhasil mengoptimalkan parameter *Naive Bayes* sehingga 9 dari 10 data diklasifikasikan dengan benar sebagai penerima bantuan, dan 1 data diklasifikasikan sebagai bukan penerima bantuan. Nilai *fitness* dari model yang dioptimalkan mendekati 0.9 (90% akurasi), menunjukkan bahwa PSO efektif dalam menyesuaikan parameter untuk meningkatkan performa *Naive Bayes*. Maka hal ini dapat disimpulkan bahwa penggunaan algoritma PSO, model *Naive Bayes* menjadi lebih akurat dan andal dalam klasifikasi data, khususnya dalam konteks pengambilan keputusan untuk menentukan penerima bantuan sosial.

DAFTAR PUSTAKA

- [1] D. Dalbergio, M. N. Hayati, and Y. N. Nasution, "Klasifikasi Lama Studi Mahasiswa Menggunakan Metode C5.0 pada Studi Kasus Data Kelulusan Mahasiswa Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Mulawarman Tahun 2017," *Pros. Semin. Nas. Mat. Stat. dan Apl.* 2019, vol. 1, no. 1, pp. 36–42, 2019.
- [2] I. Purnamasari, "Analisa Klasifikasi Loyalitas Siswa Lembaga Pendidikan Tari dengan Metode Naive Bayes," *J. Ilm. Komputasi*, vol. 19, no. 1, pp. 59–68, Mar. 2020, doi: 10.32409/jikstik.19.1.157.
- [3] S. Iskandar, N. R. Refisis, and B. A. Ginting, "Metode Naive Bayes Classifier Dalam Penentuan Penerima Beasiswa Bidikmisi Di Universitas Negeri Medan," *Karismatika*, vol. 7, no. 1, pp. 10–23, 2021.
- [4] A. J. Susilo, K. K. Kustanto, and D. Remawati, "Implementasi Naive Bayes Dalam Pemilihan Jenis Bahan Pembuatan Meja," *J. Ilm. SINUS*, vol. 21, no. 1, p. 39, Jan. 2023, doi: 10.30646/sinus.v21i1.674.
- [5] E. Feronica, Y. N. Nasution, and I. Purnamasari, "Optimasi Algoritma Naive Bayes Menggunakan Algoritma Genetika Untuk Memprediksi Kelulusan," *EKSPONENSIAL*, vol. 13, no. 2, p. 147, Nov. 2022, doi: 10.30872/eksponeisial.v13i2.1057.
- [6] E.- Mutiara, "Algoritma Klasifikasi Naive Bayes Berbasis Particle Swarm Optimization Untuk Prediksi Penyakit Tuberculosis (TB)," *Swabumi*, vol. 8, no. 1, pp. 46–58, Mar. 2020, doi: 10.31294/swabumi.v8i1.7668.
- [7] H. B. Jatmiko, N. Tedi Kurniadi, and D. Maulana, "Optimasi Naive Bayes Dengan Particle Swarm Optimization Untuk Analisis Sentimen Formula E-Jakarta," *J. Autom. Comput. Inf. Syst.*, vol. 2, no. 1, pp. 22–30, Jun. 2022, doi: 10.47134/jacis.v2i1.35.
- [8] H. Setiani, A. Sunyoto, and A. Nasiri, "Metode Naive Bayes dan Particle Swarm Optimization untuk Klasifikasi Penyakit Jantung," *EXPLORE*, vol. 12, no. 2, p. 6, Jul. 2022, doi: 10.35200/explore.v12i2.566.
- [9] M. R. R. Saelan, D. A. Sahputra, W. Widiastuti, and W. Gata, "Komparasi Algoritma Klasifikasi untuk Prediksi Minat Sekolah Tinggi Pelajar pada Students Alcohol Consumption," *J. Sains dan Inform.*, vol. 6, no. 2, pp. 120–129, Nov. 2020, doi: 10.34128/jsi.v6i2.236.
- [10] F. Sholekhah, A. D. Putri, R. Rahmadden, and L. Efrizoni, "Perbandingan Algoritma Naive Bayes dan K-Nearest Neighbors untuk Klasifikasi Metabolik Sindrom," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 507–514, Feb. 2024, doi: 10.57152/malcom.v4i2.1249.
- [11] A. T. Dewi Septiani, A. P. Kuncoro, P. Subarkah, and R. Riyanto, "Perbandingan Kinerja Metode Naive Bayes Classifier dan K-Nearest Neighbor pada Analisis Sentimen Ulasan Mobile Banking Jenius," *J. Krisnadana*, vol. 3, no. 2, pp. 67–77, Jan. 2024, doi: 10.58982/krisnadana.v3i2.516.
- [12] S. Shedriko and M. Firdaus, "Penentuan Klasifikasi Dengan Crisp-Dm Dalam Memprediksi Kelulusan Mahasiswa Pada Suatu Mata Kuliah," *Semnas Ristek (Seminar Nas. Ris. dan Inov. Teknol.)*, vol. 6, no. 1, pp. 826–831, Jan. 2022, doi: 10.30998/semnasristek.v6i1.5814.
- [13] M. M. Fajar, A. R. Putri, and K. F. H. Holle, "Perbandingan Metode Klasifikasi Data Mining Untuk Deteksi Keaslian Lowongan Pekerjaan di Medsos," *J. Ilm. Inform.*, vol. 7, no. 1, pp. 41–48, Jun. 2022, doi: 10.35316/jimi.v7i1.41-48.
- [14] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 103–108, Oct. 2021, doi: 10.30871/jaic.v5i2.3200.
- [15] M. Yamin Nurzaman and B. Nurina Sari, "Implementasi K-Means Clustering Dalam Pengelompokan Banyaknya Jumlah Petani Berdasarkan Kecamatan Di Provinsi Jawa Barat," *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 3, pp. 131–144, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [16] M. Rizki and M. Mulyawan, "Penerapan Metode

- K-Means Pada Data Penjualan Untuk Mendapatkan Produk Terlaris Di PT. Titian Nusantara Boga,” *JATI (Jurnal Mbs. Tek. Inform.,* vol. 7, no. 2, pp. 1303–1307, Sep. 2023, doi: 10.36040/jati.v7i2.6562.
- [17] A. F. Riany and G. Testiana, “Penerapan Data Mining untuk Klasifikasi Penyakit Stroke Menggunakan Algoritma Naïve Bayes,” *J. SAINTEKOM*, vol. 13, no. 1, pp. 42–54, Mar. 2023, doi: 10.33020/saintekom.v13i1.352.
- [18] A. Fathiarahma, A. Voutama, T. Ridwan, and N. Heryana, “Analisis Text Mining Klasifikasi Kegiatan Keluarga menggunakan Orange dengan Metode Naive Bayes,” *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 35–41, Jul. 2023, doi: 10.54914/jtt.v9i1.606.
- [19] D. Kurniawan, M. Sahata Sipayung, R. Ismayanti, M. Rivani Ibrahim, Y. Bintan, and S. Aulia Miranda, “Optimalisasi Strategi Pemenuhan Persediaan Stok Barang Menggunakan Algoritma Frequent Pattern Growth,” *METIKJ.*, vol. 6, no. 2, pp. 104–114, Dec. 2022, doi: 10.47002/metik.v6i2.387.
- [20] N. Ajjah and A. Kurniawan, “Klasifikasi Teks Mining Terhadap Analisa Isu Kegiatan Tenaga Lapangan Menggunakan Algoritma K-Nearest Neighbor (KNN),” *J-SAKTI (Jurnal Sains Komput. Inform.,* vol. 7, no. 1, pp. 254–262, 2023.
- [21] Muhammad Nur Akbar, Hariani, And Asep Indra Syahyadi, “Analisis Prediksi Ketepatan Masa Studi Mahasiswa Menggunakan Algoritma Naive Bayes Classifier Dan Feature Selection,” *J. Instek (Informatika Sains dan Teknol.,* vol. 7, no. 2, pp. 306–315, Oct. 2022, doi: 10.24252/instek.v7i2.32576.
- [22] T. Wuriyanto, H. B. Setiawan, and A. B. Tjandrarini, “Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit yang Berisiko Menggunakan Algoritma Support Vector Machine,” *J. Ilm. Scroll (Jendela Teknol. Informasi),* vol. 10, no. 1, pp. 1–6, 2022, [Online]. Available: <https://univ45sby.ac.id/ejournal/index.php/informatika/article/view/291>