

OPTIMASI PSO PADA NAÏVE BAYES UNTUK PREDIKSI CALON PENDONOR DARAH TETAP (STUDI KASUS PMI KABUPATEN KUDUS)

Gentur Wahyu Nyipto Wibowo¹, Adi Sucipto², Hery Kristanto³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Nahdlatul UlamaJI. Taman Siswa(Pekeng) Tahunan, Jepara, Jawa Tengah

e-mail: ¹gentur23@yahoo.com, ²adisucipto@gmail.com, ³191240000985@unisnu.ac.id,

Abstrak

Donor darah merupakan salah satu kegiatan yang krusial dan salah satu aspek terpenting guna menunjang keselamatan nyawa bagi pasien rumah sakit yang sedang membutuhkan transfusi darah. Akan tetapi dalam pelaksanaannya, permintaan kebutuhan kantong darah selalu lebih tinggi daripada jumlah kantong darah yang dihasilkan sehingga mengakibatkan ketersediaan kantong darah yang selalu mengalami kekosongan dan meningkatkan resiko pasien yang tidak tertolong. Berangkat dari masalah itu diperlukan suatu teknik klasifikasi data mining untuk memprediksi calon pendonor darah tetap. Dataset awal yang diperoleh sebanyak 911 record pada periode Januari sampai Maret 2023 dengan 18 (delapan belas) atribut. Kemudian dilakukan preprocessing data sehingga dataset yang akan digunakan 7 (tujuh) atribut dengan rincian 1 (satu) atribut ID yaitu nama, 5 (lima) atribut regular yaitu umur, jenis kelamin, golongan darah, hemoglobin, berat badan dan 1 (satu) atribut kelas yaitu status pendonor rutin. Penelitian menggunakan metode Naïve Bayes yang dioptimasi menggunakan Particle Swarm Optimization dengan pengujian menggunakan confusion matrix untuk menghasilkan nilai akurasi dan AUC dengan analisa software Rapidminer. Dari hasil penelitian, akurasi yang dihasilkan Naïve Bayes adalah 75,10 % sedangkan setelah dioptimasi menggunakan Particle Swarm Optimization akurasinya mengalami peningkatan sebesar 5,31 % sehingga naik menjadi 80,41 %. Diperoleh nilai AUC (Area Under Curve) sebesar 0,716 yang dikategorikan sebagai fair classification. Dengan dihasilkan sebuah pengetahuan dalam proses penelitian tersebut, kedepannya mampu mengatasi kesenjangan persediaan stok kantong darah di PMI Kabupaten Kudus.

Kata Kunci: Prediksi, Donor Darah, Naïve Bayes, Particle Swarm Optimization

Abstract

Blood donation is one of the crucial activities and one of the most important aspects to support the life safety of hospital patients who are in need of blood transfusions. However, in practice, the demand for blood bags is always higher than the number of blood bags produced, resulting in the availability of blood bags which are always empty and increasing the risk of patients who are not helped. Departing from this problem, a data mining classification technique is needed to predict prospective blood donors. The initial dataset obtained was 911 records in the period January to March 2023 with 18 (eighteen) attributes. Then data preprocessing is carried out so that the dataset will use 7 (seven) attributes with details of 1 (one) ID attribute, namely name, 5 (five) regular attributes, namely age, sex, blood group, hemoglobin, weight and 1 (one) attribute class i.e. routine donor status. The research uses the Naïve Bayes method which is optimized using Particle Swarm Optimization with testing using a confusion matrix to produce accuracy and AUC values with rapidminer software analysis. an increase of 5.31% so that it rose to 80.41%. An AUC (Area Under Curve) value of 0.716 was obtained which was categorized as a fair classification. By generating knowledge in the research process, in the future it will be able to overcome the gap in supply of blood bags at PMI Kudus Regency.

Keywords: Prediction, Blood Donation, Naïve Bayes, Particle Swarm Optimization

1. Pendahuluan

Dalam dunia medis, persediaan kantong darah di rumah sakit merupakan aspek penting dan krusial guna menunjang keselamatan nyawa bagi pasien yang sedang membutuhkan transfusi darah. Kegiatan mendonorkan darah secara rutin di Unit Donor Darah Palang Merah Indonesia (UDD PMI) yang sudah tersebar pada tiap Kabupaten di Indonesia merupakan hal fundamental dalam menuntaskan permasalahan yang dihadapi. Akan tetapi hal tersebut sulit untuk terjamin ketersediaannya akan kantong darah karena beberapa kendala. Di masa sekarang, Indonesia tiap tahunnya kekurangan setidaknya 500 ribu kantong darah, data tersebut merujuk dari *World Health Organization* (WHO) yang menunjukkan angka minimal yang harus didapatkan untuk kebutuhan kantong darah sebanyak 2,5% dari keseluruhan jumlah penduduk di Indonesia. Dibutuhkan setidaknya 5,1 juta kantong darah per tahun yang mana berbanding terbalik dengan ketersediaannya per tahun hanya berkisar 4,9 juta kantong darah dari total 3,05 juta pendonor darah. [1]

Berangkat dari permasalahan yang ada, peneliti memilih menggunakan *Naïve Bayes* yang memiliki karakteristik dalam mengolah banyak jumlah data dan dapat menangani atribut yang tidak sama, gangguan yang terdapat pada data serta kemampuannya menangani data yang kosong (*missing value*) sehingga sesuai untuk diterapkan dengan tipikal data dalam ruang lingkup objek penelitian. Akan tetapi, *Naïve Bayes* merupakan metode klasifikasi yang memiliki kekurangan dimana probabilitasnya tidak akan bisa menghitung seberapa besar akurasi sebuah klasifikasi. Metode klasifikasi *Naïve Bayes* dalam ruang lingkup penelitian ini perlu dioptimalkan dengan memberikan nilai bobot untuk setiap atribut dengan penggunaan algoritma *Particle Swarm Optimization* (PSO) yang dapat meningkatkan akurasi sebagai tujuannya [2].

Beberapa penelitian terdahulu yang diajukan sebagai acuan yaitu [3] yang berjudul "Komparasi Algoritma C4.5 dan *Naïve Bayes* dalam Penentuan Status Kelayakan Donor Darah" yang bertujuan untuk memprediksi kelayakan donor darah. Hasil penelitian menunjukkan bahwa algoritma *Naïve Bayes* memiliki akurasi tinggi dibanding algoritma lain sebesar 93,26% dan termasuk dalam kategori baik.

Selanjutnya [4] yang berjudul "Penerapan Data Mining Untuk Penentuan Kelayakan Calon Pendonor Darah Menggunakan Metode *Naïve Bayes* (Studi Kasus PMI Kabupaten Kotawaringin Timur)" yang bertujuan untuk menentukan kriteria indikator dalam uji kelayakan donor darah. Hasil penelitian didapatkan akurasi yang tinggi sebesar 77,00 % dan termasuk dalam kategori baik.

Selanjutnya [5] yang berjudul "Penerapan Algoritma *Naïve Bayes* Untuk Prediksi Heregistrasi Calon Mahasiswa Baru" yang membahas kemungkinan calon mahasiswa baru pada penerimaan calon mahasiswa baru di UNISNU Jepara yang akan heregistrasi atau tidaknya. Hasil penelitian tersebut menunjukkan akurasi dari *Naïve Bayes* diklasifikasikan baik dengan nilai AUC 0,841 dan akurasi sebesar 92,67%.

Selanjutnya [6] yang berjudul "Penerapan *Particle Swarm Optimization* (PSO) Dalam Pemilihan Atribut Untuk Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis Dengan Metode *Naïve Bayes*" yang bertujuan menentukan prediksi pasien dengan penyakit hepatitis. Hasil yang diperoleh dari penelitian tersebut menunjukkan bahwa algoritma *Naïve Bayes* yang dioptimasi dengan PSO menghasilkan akurasi lebih tinggi 7,65% dengan tingkat akurasi sebesar 92,50% dibandingkan dengan algoritma *Naïve Bayes*.

Berdasarkan hal tersebut, maka peneliti memilih menggunakan algoritma *Naïve Bayes* yang dioptimasi dengan *Particle Swarm Optimization* untuk memprediksi calon pendonor darah di PMI Kabupaten Kudus. Dataset yang digunakan adalah data pendonor yang berhasil mendonorkan darahnya pada rentang periode Januari-Maret 2023 dengan 7 (tujuh) atribut dengan rincian 1 (satu) atribut ID yaitu nama, 5 (lima) atribut regular yaitu umur, jenis kelamin, golongan darah, hemoglobin, berat badan, dan 1 (satu) atribut kelas yaitu status pendonor rutin. Melalui metode algoritma *Naïve Bayes* dengan optimasi *Particle Swarm Optimization* diharapkan mampu memprediksi calon pendonor darah tetap dengan lebih akurat.

2. Metode Penelitian

Tahapan yang dilalui dalam penelitian ini terdiri dari :

1. Pengumpulan Data
Data yang dikumpulkan pada penelitian ini adalah pendonor darah yang berhasil mendonor darahnya yang tercatat dalam database PMI Kabupaten Kudus rentang periode Januari – Maret 2023 yang digunakan sebagai dasar untuk memprediksi calon pendonor darah tetap.
2. *Pre-processing* Data
Tahapan ini meliputi penyusunan dan pengkategorian data, seleksi atribut, dan pembersihan data agar memberikan hasil yang optimal.
3. Penerapan Algoritma *Naïve Bayes* + *PSO*
Setelah data di *pre-processing* maka data akan diolah dengan algoritma *Naïve Bayes* yang dioptimasi *Particle Swarm Optimization* menggunakan tools RapidMiner
4. Evaluasi dan Validasi
Setelah dilakukan penerapan algoritma *Naïve Bayes* + *PSO* maka akan dilakukan pengujian menggunakan confusion matrix dan kurva ROC (*Receiver Operating Characteristic*)

3. Hasil dan Pembahasan

3.1 Pengumpulan Data

Tahapan pengumpulan data pada penelitian ini dilakukan di PMI Kabupaten Kudus di bagian IT. Tahapan ini dilakukan dengan cara observasi dan wawancara perihal Pendonor Darah yang telah berhasil mendonorkan darahnya dan tercatat dalam database PMI Kabupaten Kudus. Data yang diperoleh sebanyak 911 record pada periode Januari – Maret 2023 dengan 7 (tujuh) atribut dengan rincian 1 (satu) atribut ID yaitu nama, 5 (lima) atribut regular yaitu umur, jenis kelamin, golongan darah, hemoglobin, berat badan, dan 1 (satu) atribut kelas yaitu status pendonor rutin dalam bentuk file Microsoft Excel sehingga dapat mempermudah pengolahan data.

3.2 *Pre-processing* Data

a. Penyusunan dan Pengkategorian Data

Setelah data didapatkan data kemudian dilakukan penyusunan dan pengkategorian data. Data dikategorikan agar mempermudah dan memperoleh hasil maksimal dalam penerapan algoritma. Atribut dikategorikan berdasarkan kriteria atribut yang sudah

ditentukan menggunakan software Microsoft Excel dapat dilihat pada Tabel 1.

Tabel 1. Konversi Kriteria Atribut

No.	Atribut	Kategori	Nilai Atribut
1.	Umur	17-30 tahun	1
		31-50 tahun	2
		51-65 tahun	3
2.	Hemoglobin	12.5 – 13.0	1
		13.1 – 14.0	2
		14.1 – 15.0	3
		15.1 – 16.0	4
		16.1 – 17.0	5
3.	Berat Badan	45-65 Kg	1
		66-85 Kg	2
		86 – 105 Kg	3
		>= 106 Kg	4

b. Seleksi Atribut

Tahapan ini dilakukan untuk menyeleksi atribut yang akan digunakan dari semua atribut yang ada pada data awal. Pada penelitian ini atribut yang digunakan yaitu Nama, Umur, Golongan Darah, Jenis Kelamin, Hemoglobin, Berat Badan, dan Status Pendonor Rutin. Tipe data masing-masing atribut dapat dilihat pada tabel berikut:

Tabel 2 Tipe Data Atribut

No	Nama Atribut	Tipe Data	Role
1.	Nama	Polynomial	ID
2.	Umur	Integer	
3.	Golongan Darah	Polynomial	
4.	Jenis Kelamin	Binomial	
5.	Hemoglobin	Integer	
6.	Berat Badan	Integer	
7.	Status Pendonor Darah	Polynomial	Label

c. Pembersihan Data

Tahapan ini dilakukan untuk mengidentifikasi dan menghilangkan data yang kosong (*missing*)

value) menggunakan operator *replace missing values*.

Gambar 3. Hasil Pembersihan Data

Dari gambar diatas dapat dilihat bahwa pada keterangan filter yang dihasilkan yaitu sebanyak 911 data, hal ini berarti dari proses pembersihan data tidak terdeteksi adanya data yang kosong.

3.3 Metode *Naive Bayes*

Setelah melalui tahapan pre-processing data, langkah selanjutnya adalah mengolah data menggunakan algoritma *Naive Bayes*. Pada penelitian ini, data dibagi menjadi 2 (dua) bagian, yaitu 73% digunakan sebagai data training dan 27% digunakan untuk data testing. Pembagian data dilakukan menggunakan operator split data.

Berikut ini adalah data yang akan diimplementasikan terhadap algoritma *Naive Bayes*:

N	Nama	U	GD	JK	HB	:	Sta
o							tus
1	Brahma Ari	I	O	Pria	III	:	Ya
2	Julianawati	II	O	Wanita	II	:	Ya
3	M. Amiruddin	I	A	Pria	IV	:	Ya
:	:	:	:	:	:	:	:
9	M..Andi N	II	A	Pria	IV	:	Tidak
9	Khoirun Nuha	I	O	Pria	II	:	Tidak
9	Irfan Agus	I	O	Pria	IV	:	Tidak
9	M. Khoirul Riza	I	A	Pria	IV	:	Ya

Tabel 3. Data Penelitian

a. Menghitung Probabilitas Kelas
Sebelum menentukan prediksi, langkah pertama yang dilakukan adalah menghitung probabilitas masing-masing kategori kelas pada atribut status pendonor rutin dengan cara menjumlahkan masing-masing kategori kelas dibagi jumlah semua kategori kelas pada atribut status pendonor rutin.

Maka :

$$1. P(C \text{ Pertama}) = P(\text{Class Pertama})$$

$$P(Ya) = 724/911$$

$$P(Ya) = 0,795$$

$$2. P(C \text{ Kedua}) = P(\text{Class Kedua})$$

$$P(\text{Tidak}) = 187/911$$

$$P(\text{Tidak}) = 0,205$$

Sehingga diperoleh hasil sebagai berikut:

Tabel 4. Hasil Probabilitas Kelas

No	Status Pendonor Rutin	Probabilitas
1	Ya	0,795
2	Tidak	0,205
Jumlah Keseluruhan		1

b. Menghitung Probabilitas Atribut Target Prediksi/Reguler

Setelah menghitung probabilitas atribut kelas, langkah selanjutnya adalah menghitung probabilitas masing-masing kategori atribut target prediksi/reguler yaitu umur, golongan darah, jenis kelamin, hemoglobin dan berat badan. Sampel yang digunakan adalah atribut jenis kelamin.

Maka :

$$1. P(X \mid \text{Jenis Kelamin} = \text{"Pria"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) / 724$$

$$= 614 / 724$$

$$= 0.848$$

$$2. P(X \mid \text{Jenis Kelamin} = \text{"Wanita"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) / 724$$

$$= 110 / 724$$

$$= 0.152$$

Sehingga diperoleh hasil sebagai berikut:

Tabel 5. Hasil Probabilitas Atribut Jenis Kelamin

Probabilitas	Status Pendonor Rutin	
	Ya	Tidak
Pria	0.848	0.717
Wanita	0.152	0.283
Jumlah	1	1

c. Menentukan Hasil Prediksi

Setelah menghitung probabilitas pada masing-masing kategori atribut dan kelas,

langkah selanjutnya adalah menentukan hasil prediksi terhadap dataset berdasarkan hasil perhitungan probabilitas kategori atribut dan kelas.

Maka :

1. Brahma Ari Prabowo =

$$\begin{aligned} & \bullet P(\text{Status Pendonor Rutin} = \text{"Ya"}) \\ &= P(\text{Status Pendonor Rutin} = \text{"Ya"}) * P(\text{Umur} = \text{"2"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) * P(\text{Golongan Darah} = \text{"O Positif"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) * P(\text{Jenis Kelamin} = \text{"Pria"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) * P(\text{Hemoglobin} = \text{"3"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) * P(\text{Berat Badan} = \text{"3"} \mid \text{Status Pendonor Rutin} = \text{"Ya"}) / (P(X = \text{"Ya"}) + P(X = \text{"Tidak"})) \\ &= 0.79 * 0.283 * 0.431 * 0.848 * 0.267 * 0.130 / 0.00341 \\ &= 0.8344 \end{aligned}$$

$$\begin{aligned} & \bullet P(\text{Status Pendonor Rutin} = \text{"Tidak"}) \\ &= P(\text{Status Pendonor Rutin} = \text{"Tidak"}) * P(\text{Umur} = \text{"2"} \mid \text{Status Pendonor Rutin} = \text{"Tidak"}) * P(\text{Golongan Darah} = \text{"O Positif"} \mid \text{Status Pendonor Rutin} = \text{"Tidak"}) * P(\text{Jenis Kelamin} = \text{"Pria"} \mid \text{Status Pendonor Rutin} = \text{"Tidak"}) * P(\text{Hemoglobin} = \text{"3"} \mid \text{Status Pendonor Rutin} = \text{"Tidak"}) * P(\text{Berat Badan} = \text{"3"} \mid \text{Status Pendonor Rutin} = \text{"Tidak"}) / (P(X = \text{"Ya"}) + P(X = \text{"Tidak"})) \\ &= 0.21 * 0.214 * 0.385 * 0.283 * 0.225 * 0.428 / 0.00341 \\ &= 0.1655 \end{aligned}$$

Perhitungan diatas juga berlaku untuk record selanjutnya sampai record terakhir pada data. Sehingga diperoleh hasil prediksi sebagai berikut :

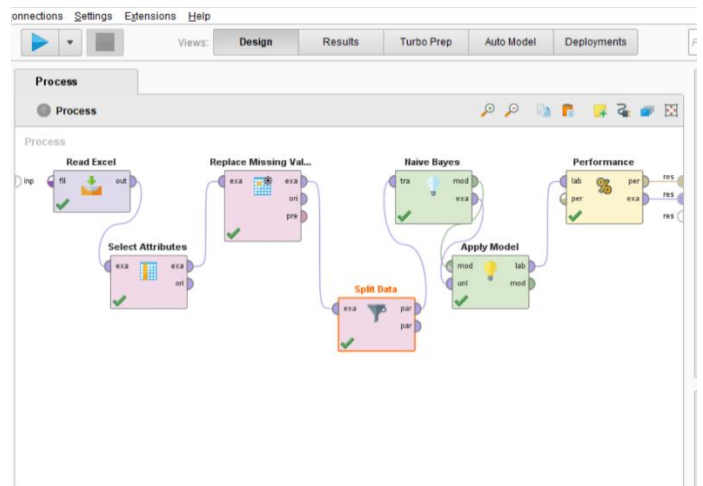
Tabel 6. Hasil Perhitungan Prediksi

No	Nama	Class Predicted	
		Ya	Tidak
1	Brahma Ari Prabowo	0.8344	0.1655
2	Julianawati	0.8160	0.1839
3	Muchamad Amiruddin	0.9591	0.0408
4	Muh Yusuf Hamzah	0.8968	0.1031

5	Athur Galih Prakoso	0.8968	0.1031
6	Shofiyan Nor Arifin	0.8446	0.1553
7	Martina Elfiana	0.7577	0.2422
8	Arris Kristianto	0.6842	0.3157
9	Sugeng Triyono	0.8255	0.1744
10	Agus Triadhi	0.8920	0.1079
11	FX Didiek Setyawan	0.8830	0.1169
12	Vita Aprillia Sari	0.7620	0.2379
13	Bambang Suwono	0.9468	0.0531
:	:	:	:
907	Siswo Suryo K	0.8165	0.1834
908	Muhamad Andi N	0.8361	0.1638
909	Khoirun Nuha Akmal	0.5961	0.4038
910	Irfan Agus Saputro	0.7768	0.2231
911	Muhammad Khoirul Riza	0.5547	0.4452

1.	Brahma Ari Prabowo	Ya	0.8344	Ya
2	Julianawati	Ya	0.8160	Ya
3	Muchamad Amiruddin	Ya	0.9591	Ya
4	Muh Yusuf Hamzah	Ya	0.8968	Ya
5	Athur Galih Prakoso	Ya	0.8968	Ya
6	Shofiyan Nor Arifin	Ya	0.8446	Ya
7	Martina Elfiana	Ya	0.7577	Ya
8	Arris Kristianto	Tidak	0.6842	Ya
9	Sugeng Triyono	Tidak	0.8255	Ya
10	Agus Triadhi	Ya	0.8920	Ya
11	FX Didiek Setyawan	Ya	0.8830	Ya
12	Vita Aprillia Sari	Tidak	0.7620	Ya
13	Bambang Suwono	Ya	0.9468	Ya
:	:	:	:	:
907	Siswo Suryo K	Tidak	0.8165	Ya
908	Muhamad Andi N	Tidak	0.8361	Ya
909	Khoirun Nuha Akmal	Tidak	0.5961	Ya
910	Irfan Agus Saputro	Tidak	0.7768	Ya
911	Muhammad Khoirul Riza	Ya	0.5547	Ya

Berikut ini pemodelan *Naïve Bayes* pada *tools* RapidMinner



Sehingga diperoleh hasil prediksi sebagai berikut:

Tabel 7. Perhitungan Probabilitas Akhir

No	Nama	Konsentrasi Status Pendoron	Nilai Max	Hasil Klasifikasi
----	------	-----------------------------	-----------	-------------------

Gambar 4. Pemodelan Algoritma *Naïve Bayes*

Hasil prediksi dari *tools* RapidMiner ditampilkan sebagai berikut :

Gambar 5. Hasil Klasifikasi

3.4 Metode *Naïve Bayes* + *PSO*

Setelah melalui tahapan pemrosesan algoritma *Naïve Bayes*, langkah selanjutnya adalah mengoptimasi data menggunakan metode *Particle Swarm Optimization*. Pada penelitian ini, data dibagi menjadi 2 (dua) bagian, yaitu 73% digunakan sebagai data training dan 27% digunakan untuk data testing.

a. Menghitung Nilai Probabilitas

Untuk menghitung mencari nilai baru harus menentukan parameter terlebih dahulu dengan ketentuan r (vector) dicari secara acak dengan rentang 0-1 dan C , didapatkan nilai $C_1 = 1$, $C_2 = 1$, $r_1 = 0.4$, dan $r_2 = 0.5$. Setelah didapatkan nilai parameter maka selanjutnya mencari nilai baru sesuai rumus *Particle Swarm Optimization*, sampel yang digunakan adalah atribut umur yang dapat dilihat perhitungan berikut ini :

$$X_1 = 0 + (1 * 0.4 * (2 - 2)) + (1 * 0.5 * (5 - 2)) = 1.5$$

$$X_2 = 0 + (1 * 0.4 * (3 - 3)) + (1 * 0.5 * (5 - 2)) = 1$$

$$X_3 = 0 + (1 * 0.4 * (2 - 2)) + (1 * 0.5 * (5 - 2)) = 1.5$$

$$X_4 = 0 + (1 * 0.4 * (2 - 2)) + (1 * 0.5 * (5 - 2)) = 1.5$$

$$X_5 = 0 + (1 * 0.4 * (2 - 2)) + (1 * 0.5 * (5 - 2)) = 1.5$$

Setelah itu hasil dari perhitungan ditambahkan dengan nilai masing-masing populasi, maka dapat dilihat pada tabel berikut ini:

Tabel 8. Perhitungan Nilai Umur

	Umur	f(x)	Pbest	Gbest	V	V(j)	P Umur
x1	2	9604	2	5	v1	0	1.5
x2	3	9409	3		v2	0	1
x3	2	9604	2		v3	0	1.5
x4	2	9604	2		v4	0	1.5
x5	2	9604	2		v5	0	1.5

b. Menentukan Hasil Prediksi

Setelah didapatkan nilai baru pada masing-masing atribut, selanjutnya menghitung $P(X|H)$ *Naïve Bayes* dengan penambahan bobot atribut dari hasil G_{best} . Kemudian dilanjutkan langkah *Naïve Bayes* dari perhitungan probabilitas akhir untuk setiap kelas sampai akhir. Selanjutnya dari masing-masing atribut dan nilai probabilitas kelas dikalikan.

Dari kedua hasil yang sudah ditentukan pada tiap kelas, bandingkan nilai yang paling tinggi. Jika kelas "Ya" bernilai paling tinggi, maka hasilnya "Ya". Begitu sebaliknya jika nilai "Tidak" bernilai tinggi maka hasilnya "Tidak". Perhitungan nilainya berada pada tabel berikut :

Tabel 9. Hasil Perhitungan Klasifikasi

No	Nama	Class Predicted	
		Ya	Tidak
1	Brahma Ari	0.518	0.482
	Prabowo		
2	Julianawati	0.527	0.473
3	Muchamad	0.468	0.532
	Amiruddin		
4	Muh Yusuf	0.533	0.467
	Hamzah		
5	Athur Galih	0.533	0.467
	Prakoso		

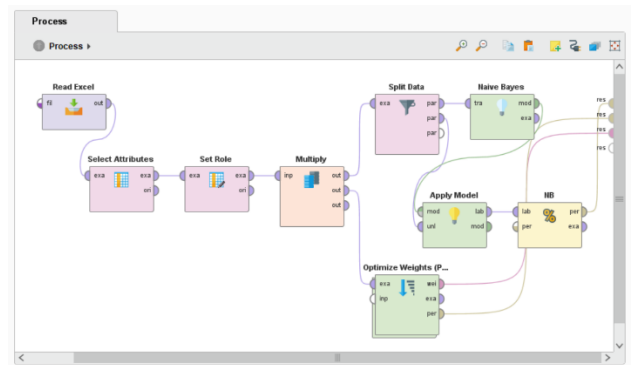
6	Shofiyen Nor Arifin	0.516	0.484
7	Martina Elfiana	0.531	0.469
8	Arris Kristianto	0.520	0.480
:	:	:	:
2	Annas	0.533	0.467
4	Wibowo		
4	Masvan		
5	Yulianto	0.523	0.477

Sehingga diperoleh hasil prediksi sebagai berikut

Tabel 10. Perhitungan Probabilitas NB+PSO

No	Nama	Konsentrasi Status Pendoron	Nilai Max	Hasil Klasifikasi
1.	Brahma Ari Prabowo	Ya	0.518	Ya
2	Julianawati	Ya	0.527	Ya
3	Muchamad Amiruddin	Tidak	0.532	Tidak
4	Muh Yusuf Hamzah	Ya	0.533	Ya
5	Athur Galih Prakoso	Ya	0.533	Ya
6	Shofiyen Nor Arifin	Ya	0.516	Ya
7	Martina Elfiana	Ya	0.531	Ya
8	Arris Kristianto	Tidak	0.520	Ya
:	:	:	:	:
244	Annas Wibowo	Ya	0.533	Ya
245	Masvan Yulianto	Ya	0.523	Ya

Berikut ini pemodelan Naïve Bayes + PSO pada tools RapidMiner



Gambar 6. Pemodelan NB + PSO

3.5 Evaluasi dan Validasi

Evaluasi dan validasi hasil pada penelitian ini dilakukan menggunakan confusion matrix sebagai *performance* dan kurva ROC (*Receiver Operating Characteristic*) yang menghasilkan nilai AUC (*Area Under Curve*).

a. Metode Naïve Bayes

Diketahui :

Tabel 11. Hasil *Performance Vector NB*

Classification	Class	
	Ya	Tidak
Ya	175	41
Tidak	20	9
Accuracy	75.10%	

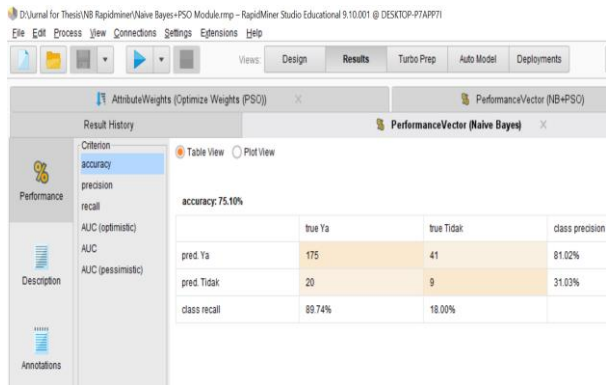
Berdasarkan hasil performance vector pada tabel 11, dapat disimpulkan bahwa dari dataset sebanyak 245 record sebagai data testing dari total seluruh data sebanyak 911 record didapatkan hasil prediksi benar sebanyak 216 record dengan klasifikasi benar kelas “Ya” = 175 dan klasifikasi benar dengan kelas “Tidak” = 41, kemudian klasifikasi salah sebanyak 29 record dengan rincian klasifikasi salah kelas “Ya” = 20 dan klasifikasi salah kelas “Tidak” = 9. Sehingga dihasilkan :

- $$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\%$$

$$= \frac{(175 + 9)}{(175 + 9 + 41 + 20)} \times 100\%$$

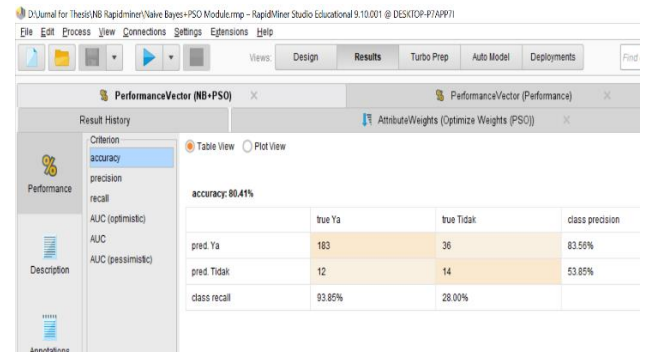
$$= \frac{(184)}{(245)} \times 100\% = 75,10\%$$

Berdasarkan pengujian diatas dapat disimpulkan bahwa persentase *accuracy* yang dihasilkan adalah sebesar 75,10%. Berikut adalah hasil *performance accuracy* pada software RapidMiner:



Gambar 7. Hasil Accuracy Naive Bayes

Berdasarkan pengujian diatas dapat disimpulkan bahwa persentase *accuracy* yang dihasilkan adalah sebesar 80,41%. Berikut adalah hasil *performance accuracy* pada software RapidMiner:



Gambar 8. Hasil Accuracy NB + PSO

b. Metode Naive Bayes + PSO

Diketahui :

Tabel 12. Hasil Performance Vector NB + PSO

Classification	Class	
	Ya	Tidak
Ya	183	36
Tidak	12	14
Accuracy	80.41 %	

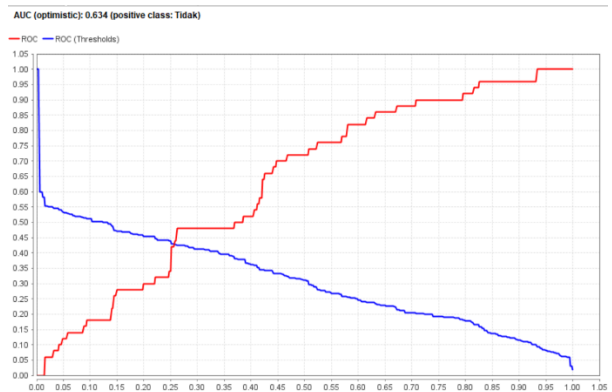
Berdasarkan hasil performance vector pada tabel 12, dapat disimpulkan bahwa dari dataset sebanyak 245 record sebagai data testing dari total seluruh data sebanyak 911 record didapatkan hasil prediksi benar sebanyak 219 record dengan klasifikasi benar kelas “Ya” = 183 dan klasifikasi benar dengan kelas “Tidak” = 36, kemudian klasifikasi salah sebanyak 26 record dengan rincian klasifikasi salah kelas “Ya” = 12 dan klasifikasi salah kelas “Tidak” = 14. Sehingga dihasilkan :

- $$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100 \%$$

$$= \frac{(183 + 14)}{(183 + 14 + 36 + 12)} \times 100 \%$$

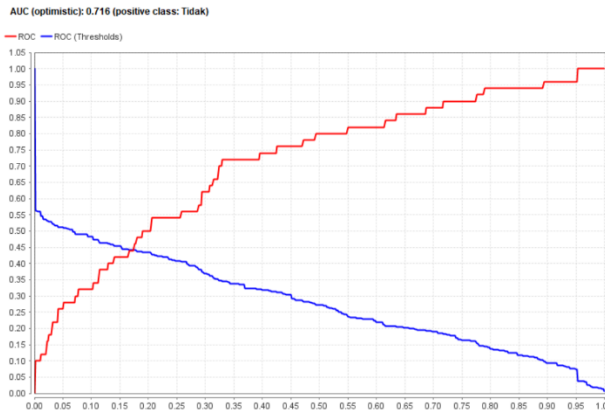
$$= \frac{(197)}{(245)} \times 100 \% = 80,41 \%$$

Berdasarkan pengujian yang telah dilakukan, metode Naive Bayes dihasilkan nilai AUC (Area Under Curve) dari kurva ROC (Receiver Operating Characteristic) sebesar 0.634 sehingga dapat disimpulkan bahwa algoritma mempunyai kinerja yang kurang baik. Berikut adalah hasil performance AUC untuk algoritma Naive Bayes pada software RapidMiner :



Gambar 9. ROC Pengujian Naive Bayes

Berdasarkan pengujian yang telah dilakukan, metode Naive Bayes + PSO dihasilkan nilai AUC (Area Under Curve) dari kurva ROC (Receiver Operating Characteristic) sebesar 0.716 sehingga dapat disimpulkan bahwa algoritma mempunyai kinerja yang cukup baik. Berikut adalah hasil performance AUC untuk algoritma Naive Bayes + PSO pada software RapidMiner :



Gambar 10. ROC Pengujian Naïve Bayes + PSO

4. Kesimpulan

Optimalisasi Particle Swarm Optimization (PSO) pada algoritma Naïve Bayes memiliki kinerja baik dalam memprediksi calon pendonor darah tetap dengan akurasi 80,41% dan nilai AUC 0,716 jika dibandingkan perhitungan algoritma Naïve Bayes dengan akurasi sebesar 75,10%. Hasil akurasi yang dihasilkan setelah diolah dengan aplikasi RapidMiner versi 9.10 menjadi 80,41 %, mengalami peningkatan sebesar 5,31 %. Atribut yang mempengaruhi dalam menentukan prediksi calon pendonor darah tetap berjumlah 7 (tujuh) atribut. Atribut – atribut tersebut yaitu nama, umur, golongan darah, jenis kelamin, hemoglobin, dan berat badan dan satu atribut label yaitu Status Pendonor Rutin.

Referensi

- [1] Kementerian Kesehatan RI., Situasi Donor Darah di Indonesia. *InfoDatin: Pusat Data dan Informasi Kementerian Kesehatan RI*, 2019.
- [2] Amalia, H., “Perbandingan Metode Data Mining SVM dan NN Untuk Klasifikasi Penyakit Ginjal Kronis”. *Jurnal PILAR Nusa Mandiri*, vol. 14, no. 1, p. 1., 2018.
- [3] Hayatin, N., Marthasari, G. I., & Nuraini, L., “Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization”. *Jurnal Online Informatika*, 5(1), 81–88, 2020. <https://doi.org/10.15575/join.v5i1.558>
- [4] Handayani, K., Latif, A., Rifqi Firdaus, M., & Nur Hasan, F., “Komparasi Algoritma C4.5 dan Naïve Bayes dalam Penentuan Status Kelayakan Donor Darah”. *SISTEMASI: Jurnal Sistem Informasi*, vol. 10, no. 3, 676-687, 2021. <http://sistemasi.ftik.unisi.ac.id>
- [5] Adha, A. C., Yuhandri, Y., & Nurcahyo, G. W., “Prediksi Potensi Relawan Pendonor Darah Menjadi Pendonor Darah Tetap dengan Penerapan Metode Klasifikasi Decision Tree”. *Jurnal Informasi Dan Teknologi*, 233–238, 2021. <https://doi.org/10.37034/jidt.v3i4.158>
- [6] Priliani Hutrim, H., “Penerapan Data Mining Untuk Penentuan Kelayakan Calon Pendonor Darah Menggunakan Metode Naive Bayes (Studi Kasus PMI Kabupaten Kotawaringin Timur)”. *JUISI*, vol. 06, no. 02, 2020.
- [7] Abdul Manan, M., Wahyu Nyipto Wibowo, G., & Sarwido., “Penerapan Algoritma Naïve Bayes Untuk Prediksi Heregistrasi Calon Mahasiswa Baru”. *Jurnal Teknik Informatika*, vol. 1, no. 1, hal. 1 – 10, 2022.
- [8] Septiani, T., K., N., Sukrisno, M., M, A., Ghani., “Penerapan Particle Swarm Optimization (PSO) Dalam Pemilihan Atribut Untuk Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis Dengan Metode Naive Bayes”. *Journal Speed-Sentra Penelitian Engineering Dan Edukasi*, vol. 12. no.1, 2020.
- [9] Hadi, S., *Pengertian Darah dan Fungsinya*. Bandung: Informatika, 2015.
- [10] Magdalena, L., & Mulyasari, H., “Rancangan Sistem Informasi PMI Dengan Mengintegrasikan Data Pendonor dan Stok Darah Antar Cabang PMI Di Wilayah III Cirebon”. *Prosiding SENDI_U*. hal. 68-75, 2018.
- [11] Kuncoro, S., *Efek Donor Darah*. Yogyakarta: Graha Ilmu. 2015.
- [12] Lutfi, M, Zuryaty, and Mayangsari. M., “Donor Darah : Selamatkan Jiwa dan Sehatkan Raga di Masa Pandemi COVID-19”. *Jurnal Paradigma (Pemberdayaan & Pengabdian Kepada Masyarakat)*. vol. 4, no. 1, hal 27-35. 2022.
- [13] Septiana D., Astuti Y., dan Barokah L., “Gambaran Karakteristik Pendonor Darah Yang Lolos Seleksi Donor Di Unit Transfusi Darah Palang Merah Indonesia Kabupaten Gunungkidul”. *Jurnal Ilmiah Cerebral Medika*. vol. 3, no. 2, 2021.
- [14] Castaka A., S., Teguh, M., Z., “Rancang Bangun Aplikasi Donor Darah Berbasis Mobile di PMI Kabupaten Bandung”. *KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer*. vol. 01, no 01, pp 11-18, 2018.
- [15] Palang Merah Indonesia, “*Tentang Palang Merah Indonesia (PMI)*”, 2015.
- [16] Peraturan Menteri Kesehatan Republik Indonesia Nomor 91 tahun 2015 tentang Standar Pelayanan Transfusi Darah., 2015.

- [17] Efori B., *Data Mining untuk Perguruan Tinggi*. Yogyakarta: Deepublish, 2020.
- [18] Y, Mardi., "Data Mining : Klasifikasi Menggunakan Algoritma C4.5". *Jurnal Edik Informatika*. vol. 2, no. 2, pp. 213-219, 2020. doi: 10.22202/ei.2016.v2i2.1465.
- [19] Nurachim, R. I., "Pemilihan Model Prediksi Indeks Harga Saham Yang Dikembangkan Berdasarkan Algoritma Support Vector Machine (SVM) Atau Multilayer Perceptron (MLP) Studi Kasus : Saham PT Telekomunikasi Indonesia Tbk". *Jurnal Teknologi Informatika & Komputer* . vol. 5, Issue 1, 2019.
- [20] Hadapiningradja Kusumodestoni, R., Sarwido., "Komparasi Model Support Vector Machines (Svm) dan Neural Network Untuk Mengetahui Tingkat Akurasi Prediksi Tertinggi Harga Saham". *JURNAL INFORMATIKA UPGRIS*. vol. 3, issue 1, 2017.
- [21] Rohman, F. F., & Fauziah, A., "Rancang Bangun Aplikasi Sistem Pakar Untuk Menentukan Jenis Gangguan Perkembangan Pada Anak". *Media Informatika*, vol. 6, no. 1, 2018.
- [22] Idris, M., "Implementasi Data Mining Dengan Algoritma Naïve Bayes Untuk Memprediksi Angka Kelahiran". *Jurnal Pelita Informatika*. vol. 7, no.3, 2019.
- [23] Riswanto, I., & Laluma, R. H., "Klasifikasi Kelayakan Pinjaman Pada Koperasi Karyawan Menggunakan Metode Naive Bayes Classifier Berbasis Web". *Infotronik : Jurnal Teknologi Informasi dan Elektronika*, vol. 5, no. 1, hal. 11-16, 2020.
- [24] James, G., "An Introduction to Statistical Learning with Application in R". *Springer*. no. 176, 2013.
- [25] Rumini, dan Norhikmah., "Prediksi Kegagalan Siswa Dalam Data Mining Dengan Menggunakan Metode Naïve Bayes". *Jurnal Mantik Penusa*, vol. 3, no. 1, pp 42-46. 2019.
- [26] Muhammad, H., Prasojo, C. A., Sugianto, N. A., Surtiningsih, L., & Cholissodin, I., "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris". *Jurnal Teknologi Informasi Dan Ilmu Komputer*, vol. 4, no. 3, pp 180, 2018.
- [27] Yusup, N., Zain, A. M., Zaiton, S., & Hashim, M., "Prosedia Engeneering Overview of PSO for Optimizing Process Parameter of Machining". *Elsevier*, 2012.
- [28] Husniah, H. F., & Arifin, T., "Implementasi Algoritma Naïve Bayes Berbasis Particle Swarm Optimization Untuk Memprediksi Penyakit Hepatitis". *Jurnal Ilmu Komputer*, vol. 14, no. 1, 2021.
- [29] Dinata, R. K., Akbar, H., & Hasdyna, N., "Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus". *ILKOM Jurnal Ilmiah*, vol. 12. no. 2, 104-111,2020.
- [30] Kesuma Dinata, R., & Hasdyna, N., "Klasifikasi Sekolah Menengah Pertama/Sederajat Wilayah Bireuen Menggunakan Algoritma K-Nearest Neighbor Berbasis Web". *CESS (Journal of Computer Engineering System and Science)*. vol. 5, no.1, 2020.
- [31] Fatmawati, K., & Windarto, A. P., "Data Mining : Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi". *CESS (Journal of Computer Engineering System and Science)*. vol. 3, no. 2, 2018.
- [32] A. Supratikna, *Metodologi Penelitian Kuantitatif dan Kualitatif dalam Psikologi*. Sleman : Universitas Sanata Darma, 2015.
- [33] Achyani, E., Y., "Penerapan Metode Particle Swarm Optimization Pada Optimasi Prediksi Pemasaran Langsung". *Jurnal Informatika*. vol. 5 no. 1, 2018.
- [34] E. Prasetyo, *Data Mining Mengolah Data, Menjadi Informasi menggunakan Matlab*. Yogyakarta: Andi, 2018.