

## **THE RIGHT STEPS TOWARDS GRADUATION: NB-PSO SMART COMBINATION FOR STUDENT GRADUATION PREDICTION**

**Ahmad Hafidzul Kahfi<sup>\*1</sup>, Titin Prihatin<sup>2</sup>, Yudhistira<sup>3</sup>, Adjat Sudradjat<sup>4</sup>, Ganda Wijaya<sup>5</sup>**

<sup>1,2</sup>Informatics, Faculty Of Information Technology, Universitas Nusa Mandiri, Indonesia

<sup>3</sup>Computer Science, Faculty of Informatics Engineering, Universitas Bina Sarana Informatika, Indonesia

<sup>4</sup>Information System, Faculty of Informatics Engineering, Universitas Bina Sarana Informatika, Indonesia

<sup>5</sup>Information System, Faculty Of Information Technology, Universitas Nusa Mandiri, Indonesia

Email: <sup>1</sup>[ahmad.azx@nusamandiri.ac.id](mailto:ahmad.azx@nusamandiri.ac.id), <sup>2</sup>[titin.tpn@nusamandiri.ac.id](mailto:titin.tpn@nusamandiri.ac.id), <sup>3</sup>[yudhistira.yht@bsi.ac.id](mailto:yudhistira.yht@bsi.ac.id),  
<sup>4</sup>[adjat.ajt@bsi.ac.id](mailto:adjat.ajt@bsi.ac.id), <sup>5</sup>[ganda.gws@nusamandiri.ac.id](mailto:ganda.gws@nusamandiri.ac.id)

(Article received: February 21, 2024; Revision: March 08, 2024; published: April 22, 2024)

### **Abstract**

*The current digital era demands a more innovative approach in predicting student campuses considering that campuses are not only important for students but also for lecturers, student guardians and higher education institutions. Previous studies have used various machine learning methods such as Decision Trees, Neural Networks, Support Vector Machines, etc. in these predictions. The problem that occurs is that even though various machine learning methods have been used, there are still limitations in the accuracy and efficiency of predicting student admissions, The problem in question can be given a real example of a case that occurred. So with this problem the aim is to develop a more effective methodology in predicting student permits, with recommendations from an intelligent combination of two computational techniques Naive Bayes (NB) and Particle Swarm Optimization (PSO). This research methodology includes data collection, NB model development and model partnership with PSO. Student graduation data is used in model testing with evaluation based on metrics such as accuracy and Area Under the Curve (AUC). The results showed a significant increase in accuracy to 86.94% from 83.30% and AUC value from 0.860 to 0.884 when using the combination of NB and PSO compared to NB without either. The integration of NB and PSO has been proven to increase effectiveness in classifying student graduation prediction cases. This research opens up opportunities for the practical application of technology in the education sector and emphasizes the importance of using effective optimization and feature selection techniques in improving prediction results.*

**Keywords:** *Combination of NB and PSO, Naive Bayes, Particle Swarm Optimization, Student Graduation.*

## **LANGKAH TEPAT MENUJU KELULUSAN: KOMBINASI CERDAS NB-PSO UNTUK PREDIKSI KELULUSAN MAHASISWA**

### **Abstrak**

Era digital saat ini menuntut pendekatan yang lebih inovatif dalam prediksi kelulusan mahasiswa mengingat kelulusan tidak hanya penting bagi mahasiswa tetapi juga bagi dosen, wali mahasiswa dan institusi pendidikan tinggi. Studi-studi sebelumnya telah menggunakan berbagai metode pembelajaran mesin seperti *Decision Tree*, *Neural Networks*, *Support Vector Machine*, dan lain-lain dalam prediksi ini. Permasalahan yang terjadi meskipun berbagai metode pembelajaran mesin telah digunakan, masih terdapat keterbatasan dalam akurasi dan efisiensi prediksi kelulusan mahasiswa, dimana terdapat mahasiswa yang lulus tidak tepat waktu masih banyak dan dapat menyebabkan rasio dosen dan mahasiswa menjadi berkurang. Maka dengan permasalahan tersebut memiliki tujuan untuk mengembangkan metodologi yang lebih efektif dalam prediksi kelulusan mahasiswa, dengan mengusulkan kombinasi cerdas dari dua teknik komputasi *Naive Bayes* (NB) dan *Particle Swarm Optimization* (PSO). Metodologi penelitian ini meliputi pengumpulan data, pengembangan model NB dan pengoptimalan model dengan PSO. Data kelulusan mahasiswa digunakan dalam pengujian model dengan evaluasi berdasarkan metrik seperti akurasi dan *Area Under Curve* (AUC). Hasil penelitian menunjukkan peningkatan signifikan dalam akurasi menjadi 86,94% dari 83,30% dan nilai AUC dari 0,860 menjadi 0,884 saat menggunakan kombinasi NB dan PSO dibandingkan dengan NB tanpa pengoptimalan. Integrasi NB dan PSO terbukti meningkatkan efektivitas dalam mengklasifikasikan kasus prediksi kelulusan mahasiswa. Penelitian ini membuka peluang untuk aplikasi praktis teknologi dalam sektor pendidikan dan menekankan pentingnya penggunaan teknik optimasi dan seleksi fitur yang efektif dalam meningkatkan hasil prediksi.

**Kata kunci:** Kelulusan Mahasiswa, Kombinasi NB dan PSO, Naïve Bayes, Particle Swarm Optimization.

## 1. PENDAHULUAN

Kelulusan merupakan puncak pencapaian bagi mahasiswa di seluruh tingkatan pendidikan tidak terkecuali bagi mereka yang menempuh pendidikan tinggi. Momen kelulusan ini tidak hanya merepresentasikan puncak prestasi akademik, tetapi juga menjadi simbol dedikasi dan usaha yang telah dilakukan oleh mahasiswa selama periode studi mereka. Dalam konteks ini kelulusan bukan hanya menjadi tujuan penting bagi mahasiswa, tetapi juga menjadi kebanggaan bagi dosen, wali mahasiswa dan institusi pendidikan tinggi secara keseluruhan.

Di era digital saat ini, menjadi semakin penting dimana data dan prediksi merupakan elemen kunci dalam kebutuhan pengambilan keputusan akan metodologi yang efektif untuk mengantisipasi dan memprediksi kelulusan mahasiswa. Prediksi kelulusan yang akurat dapat membantu institusi pendidikan tinggi dalam merancang kebijakan dan intervensi yang tepat guna serta meningkatkan tingkat keberhasilan mahasiswa. Selain itu, kemampuan untuk mengidentifikasi potensi hambatan yang mungkin dihadapi oleh mahasiswa selama masa studi mereka. Sehingga institusi dapat membantu dalam memberikan dukungan yang lebih terfokus dan efektif.

Penelitian terdahulu mengenai prediksi kelulusan mahasiswa menggunakan berbagai metode seperti *Decision Tree* (DT), *Neural Networks* (NN), *Support Vector Machine* (SVM), *Particle Swarm Optimization* (PSO), dan *Genetic Algorithm* (GA) menunjukkan hasil yang signifikan. Studi-studi ini penting dalam membantu institusi pendidikan tinggi untuk merancang strategi dan intervensi yang lebih efektif.

Berikut adalah ringkasan hasil dari penelitian-penelitian tersebut dalam bentuk tabel:

Tabel 1. Data penelitian kelulusan

Peneliti	Metode	Akurasi
[1]	DT	84,96%
	NN	84,68%
	SVM	85,18%
[2]	SVM	85,84%
	SVM+PSO	86,57%
[3]	SVM	85,81%
	SVM+GA	86,43%
[4]	DT	86,55%
	DT+PSO	87,56%
[5]	NN	84,55%
	NN+PSO	86,94%
[6]	NN	84,55%
	NN+GA	87,33%

Dari tabel tersebut, kita dapat melihat bahwa berbagai metode memiliki keefektifan yang berbeda dalam memprediksi kelulusan mahasiswa. Metode seperti PSO dan GA cenderung memberikan akurasi yang lebih tinggi dibandingkan dengan metode tradisional seperti DT, NN, dan SVM. Ini

menunjukkan pentingnya penggunaan teknik-teknik canggih dalam analisis data untuk meningkatkan akurasi prediksi kelulusan mahasiswa.

Penelitian kelulusan dengan dataset berbeda juga dilakukan oleh peneliti lain. Dimana peneliti menggunakan metode NB, SVM, KNN, DT dengan hasil akurasi yang tertinggi diraih oleh metode NB dengan nilai 76,79% [7]. Metode NB juga digunakan dengan hasil akurasi 75,00% [8], sampai hasil akurasi 80,72% dan model ini menjadi acuan oleh pihak universitas lain [9]. Peneliti lain dengan data kelulusan mahasiswa menggunakan metode DT dan hanya memakai entropy dan information gain dengan hasil metode DT membentuk pohon aturan sesuai keinginan [10].

Selain itu, penelitian ini juga menunjukkan bahwa tidak ada satu metode yang paling superior, berbagai metode memiliki kelebihan dan keterbatasannya masing-masing. Oleh karena itu, pendekatan yang menggabungkan beberapa metode mungkin dapat memberikan hasil yang lebih akurat dalam prediksi kelulusan mahasiswa.

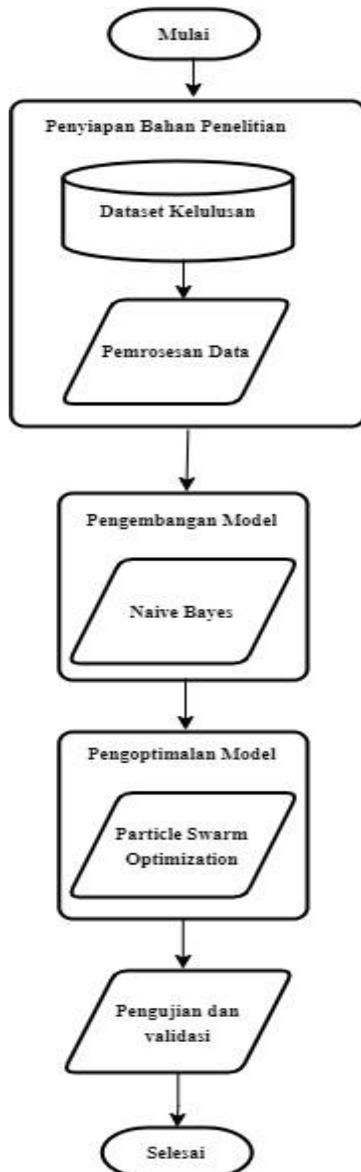
Walaupun telah ada berbagai metode pembelajaran mesin yang digunakan untuk deteksi kelulusan mahasiswa. Namun seperti masih ada keterbatasan model dalam akurasi dan efisiensi prediksi [11] contohnya model *Naïve Bayes* yang belum dilakukan pada pengujian kelulusan mahasiswa. Dalam konteks inilah, penelitian ini mengusulkan kombinasi cerdas dari dua teknik komputasi: *Naive Bayes* (NB) dan *Particle Swarm Optimization* (PSO). Metode ini diharapkan tidak hanya mampu mengklasifikasikan data mahasiswa dengan efisien, tetapi juga memperbaiki proses seleksi dan organisasi data yang heterogen.

NB terkenal dengan model klasifikasi dengan pendekatan pada prinsip probabilitistik dan asumsi independensi antar variabel prediktor dan menjadikannya pilihan yang efisien serta mudah diimplementasikan dalam berbagai studi [12]. Namun seperti masih ada keterbatasan model dalam akurasi dan efisiensi prediksi [11]. Sementara itu, PSO sebuah metode heuristik yang mengambil inspirasi dari perilaku sosial sekelompok hewan seperti burung atau ikan, belum sepenuhnya terexploitasi dalam konteks peningkatan kinerja model NB [13]. Penelitian ini berfokus pada pengembangan metodologi hibrida, yang mengkombinasikan dan mengintegrasikan kelebihan NB dalam klasifikasi yang cepat dan sederhana dengan keunggulan PSO dalam eksplorasi ruang pencarian secara global khususnya dalam optimasi parameter [14]. Pendekatan ini diharapkan dapat menghasilkan peningkatan signifikan dalam akurasi prediksi kelulusan mahasiswa, melalui peningkatan dalam seleksi fitur dan penyetelan parameter model [15].

Penelitian ini di masa mendatang akan menghasilkan metodologi baru yang belum pernah diaplikasikan sebelumnya dalam studi kelulusan mahasiswa menggunakan dataset yang sama seperti peneliti-peneliti sebelumnya dalam memprediksi kelulusan mahasiswa, yaitu menggunakan *Naive Bayes* dan *Particle Swarm Optimization*, yang mampu memberikan prediksi yang lebih akurat dan efisien. Selain itu, hasil penelitian ini berpotensi untuk dikembangkan menjadi sebuah alat bantu prediksi kelulusan mahasiswa yang dapat diintegrasikan ke dalam sistem basis data dunia pendidikan, sehingga membantu instansi pendidikan dalam pengambilan keputusan tentang kelulusan mahasiswa.

**2. METODE PENELITIAN**

Dalam metode penelitian yang dilakukan mempunyai beberapa tahap alur penelitian yang dapat dijelaskan dengan alur diagram pada Gambar 1.



Gambar 1. Alur Diagram Penelitian

Alur diagram pada Gambar 1 memberitahukan bahwa dalam penelitian melakukan proses terhadap data kelulusan dengan menggunakan model *naive bayes* dan *particle swarm optimization*. Berikut tahap-tahap yang terdapat pada alur penelitian.

**2.1. Penyiapan bahan penelitian**

Pada tahap penyiapan bahan penelitian di mulai dari mencari dan mendapatkan data [16]. Data yang didapat merupakan dataset berbentuk primer [17], dimana data didapatkan dari salah satu instansi pendidikan dengan data tentang kelulusan mahasiswa yang mencakup variabel-variabel seperti data mahasiswa dan nilai mahasiswa. Setelah pengumpulan data, langkah selanjutnya adalah pemrosesan data dengan pembersihan data untuk menghilangkan kesalahan dalam pengujian dan nilai yang hilang. Sehingga data menjadi penting untuk meningkatkan kualitas dan akurasi prediksi [18].

**2.2. Pengembangan model**

Pengembangan model *Naive Bayes* (NB) untuk pengujian pada data kelulusan melibatkan pembuatan sistem yang dapat menghitung probabilitas berdasarkan frekuensi dan gabungan nilai dari dataset yang ada. Metode NB menggunakan rumus berikut untuk menghitung probabilitas posterior[19]:

$$P(h|D) = \frac{P(D|h) \times P(h)}{P(D)} \tag{1}$$

Dalam konteks ini:

D adalah kumpulan data pelatihan. Ini merupakan data yang telah dikumpulkan dan akan digunakan untuk melatih model NB. Data ini berisi contoh-contoh sebelumnya yang akan membantu model mempelajari hubungan antara fitur *input* dan kelas *output*.

h adalah hipotesis atau kelas yang ingin kita prediksi. Dalam konteks kelulusan, ini bisa berarti apakah seorang siswa akan lulus atau tidak.

P(h|D) adalah probabilitas posterior. Ini adalah probabilitas hipotesis h (misalnya, seorang siswa lulus) setelah mempertimbangkan data pelatihan D. Probabilitas posterior ini merupakan inti dari model NB, yang mengupdate kepercayaan kita tentang kemungkinan hipotesis berdasarkan bukti yang ada.

P(D|h) adalah kemungkinan data pelatihan D diberikan hipotesis h. Ini mengukur seberapa baik data pelatihan cocok dengan hipotesis tertentu. Dalam kasus ini, ini bisa berarti seberapa sering siswa dengan karakteristik tertentu (sesuai dengan hipotesis h) lulus dalam data pelatihan.

P(h) adalah probabilitas awal dari hipotesis h. Ini adalah estimasi sebelumnya tentang seberapa sering hipotesis h terjadi, sebelum mempertimbangkan data pelatihan.

P(D) adalah probabilitas sebelumnya dari data pelatihan D. Ini sering kali merupakan faktor

normalisasi yang memastikan bahwa probabilitas posterior berada dalam rentang yang valid.

Dalam praktiknya, model NB mengasumsikan bahwa semua fitur dalam data pelatihan bersifat independen satu sama lain, yang sering tidak sepenuhnya akurat tetapi masih memberikan hasil yang baik dalam banyak kasus praktis. Model ini terutama berguna karena sederhana, efisien, dan sering memberikan hasil yang cukup akurat meskipun asumsi independensinya.

### 2.3. Pengoptimalan model

Dalam proses pengoptimalan model metode *Naive Bayes* (NB) dan *Particle Swarm Optimization* (PSO) digabungkan untuk mengembangkan model prediktif. NB, yang merupakan teknik pembelajaran mesin yang berbasis probabilitas yang digunakan untuk mengklasifikasikan data berdasarkan probabilitas kondisional. Sementara itu, PSO sebuah algoritma optimasi berdasarkan perilaku kawanan hewan yang berkelompok dan digunakan untuk menemukan kombinasi parameter optimal dalam model *Naive Bayes*. Kombinasi NB-PSO ini diharapkan dapat meningkatkan akurasi prediksi kelulusan mahasiswa dengan memanfaatkan kekuatan kedua metode tersebut.

### 2.4. Pengujian dan validasi

Penerapan dan Evaluasi Model Setelah model NB-PSO dikembangkan, langkah berikutnya adalah penerapan model pada dataset. Model ini diuji untuk menilai efektivitasnya dalam memprediksi kelulusan mahasiswa. Evaluasi dilakukan dengan menggunakan metrik seperti akurasi dan *confusion matrix*. Rumus akurasi dapat dilihat pada rumus (2) dan validasi pada rumus (3) dan rumus (4) sebagai berikut:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

$$specificity = \frac{TN}{TN+FP} \times 100\% \quad (4)$$

Uji coba dilakukan dengan membagi data menjadi dua set, yaitu data pelatihan dan data pengujian untuk memastikan bahwa model dapat bekerja dengan baik pada data yang belum pernah dilihat sebelumnya.

## 3. HASIL DAN PEMBAHASAN

### 3.1. Penyiapan bahan penelitian

Dataset yang dijadikan fokus pada penelitian ini adalah dataset dengan kumpulan data kelulusan dari mahasiswa sebuah perguruan tinggi swasta. Dataset ini terdiri dari sembilan atribut penting dan 1 atribut sebagai prediktor. Pertama, atribut jenis kelamin yang mencakup data mahasiswa laki-laki dan perempuan dengan tujuan untuk menganalisis pengaruh jenis kelamin terhadap motivasi dan kemampuan mencapai kelulusan tepat waktu. Kedua, jurusan SLTA yang berperan penting dalam menentukan kesesuaian antara jurusan di sekolah menengah atas dengan program studi di perguruan tinggi mahasiswa tersebut. Ketiga, Indeks Prestasi Kumulatif (IPK) dari semester pertama hingga semester keenam mahasiswa tersebut yang merupakan indikator penting dalam menilai performa akademis. Label atribut dalam dataset ini dibagi menjadi dua kategori, yaitu 1 untuk "Lulus Tepat Waktu" dan 2 untuk "Lulus Tidak Tepat Waktu", dengan total jumlah data sebanyak 796.

Tabel 2. Penyiapan Data Kelulusan

JK	JRS_SLTA	ASAL_SLTA	IPK1	IPK2	IPK3	IPK4	IPK5	IPK6	TEPAT WAKTU
2	7	5	3,55	3,56	3,61	3,59	3,62	3,57	YA
1	9	5	2,82	2,98	3,08	3,36	3,42	3,43	YA
1	5	3	2,27	2,27	2,52	2,41	2,77	2,93	TIDAK
2	4	2	3,41	3,49	3,44	3,47	3,52	3,54	YA
2	5	3	2,95	2,68	2,52	2,78	2,91	3,18	YA
1	9	3	1,23	0,73	0,59	0,69	2,48	2,94	TIDAK
2	7	3	2,95	3,05	2,95	2,96	3,16	3,21	YA
2	5	3	3,23	3,39	3,43	3,48	3,44	3,4	YA
1	7	3	3,27	3,12	2,46	3,05	3,17	3,25	TIDAK
2	5	3	2,82	2,51	2,51	2,59	2,76	2,94	YA
1	4	3	2,82	3,17	3,18	3,41	3,48	3,43	YA
..	...	...	.....	.....	....	.....	.....	.....	.....
2	10	5	2,68	2,59	2,23	2,14	2,03	1,92	TIDAK

Dalam penelitian ini, setelah melakukan proses awal. Selanjutnya, penyesuaian dilakukan pada beberapa atribut; jenis kelamin diubah menjadi angka (1 untuk laki-laki), jurusan SLTA dikategorikan ke dalam 10 kelompok (1-10), asal SLTA diklasifikasikan ke dalam 5 kelompok (1-5), dan nilai IPK untuk setiap semester hingga semester enam. Data tersebut dengan dirinci dalam tabel 1. Proses ini penting untuk memastikan analisis yang dilakukan

dapat diandalkan dan memberikan *insight* yang bermanfaat untuk prediksi kelulusan mahasiswa. maka akan dilanjutkan dengan pengujian menggunakan metode *Naive Bayes* (NB).

### 3.2. Pengembangan model

Data kelulusan yang telah diproses akan digunakan untuk pengujian dan validasi

menggunakan model *Naïve Bayes* (NB) melalui *software Rapid Miner*. Hasil dari pengujian tersebut akan ditunjukkan dalam *confusion matrix*, dengan detailnya disajikan di tabel 3.

Table 3. *Confussion Matrix* NB

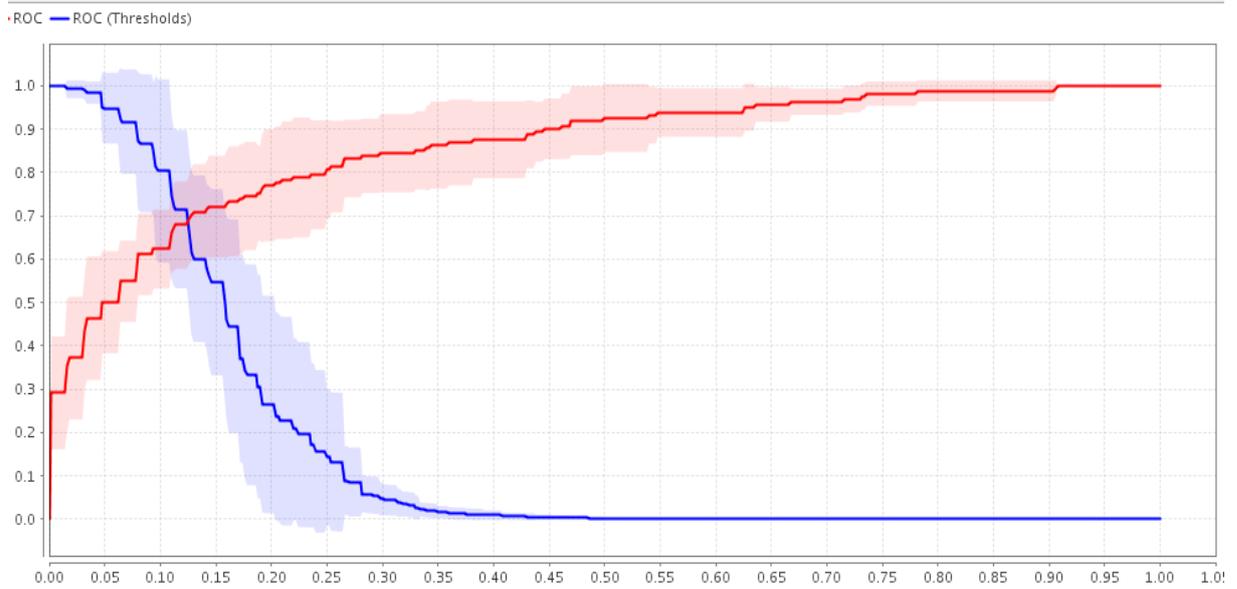
	<i>true ya</i>	<i>true tidak</i>
<i>pred ya</i>	543	42
<i>pred tidak</i>	91	120

Dari tabel 3 dapat diketahui bahwa Prediksi "ya" (kelulusan) yang benar sebanyak 543 mahasiswa dan hasil dari prediksi tersebut benar bahwa mahasiswa lulus tepat waktu dan prediksi salah terhadap

kelulusan sebanyak 42 mahasiswa. Prediksi "tidak" (tidak lulus) yang benar sebanyak 120 kali dan salah sebanyak 91 kali yang seharusnya lulus tepat waktu. Dari hasil tersebut dapat dirumuskan sebagai berikut untuk mencari akurasi:

$$\begin{aligned}
 \text{Akurasi} &= \frac{543 + 120}{543 + 120 + 42 + 91} \times 100\% \\
 &= 83,30\%
 \end{aligned}$$

Dari tabel 3 maka akan didapatkan kurva tentang kelulusan mahasiswa dengan model NB yang dapat dilihat pada gambar 2.



Gambar 2. Kurva Kelulusan Model NB

Gambar 2 Prediksi Kelulusan Mahasiswa dengan kurva tersebut dapat dikatakan bahwa model NB berhasil mencapai AUC (*Area Under the Curve*) sebesar 0,860% pada kurva ROC, menunjukkan efektivitas tinggi dari model prediksi yang digunakan dalam mengklasifikasikan kelulusan mahasiswa.

### 3.3. Pengoptimalan model

Data kelulusan yang telah diproses oleh model *Naïve Bayes* (NB) maka akan digunakan untuk pengujian dan validasi menggunakan model NB dengan metode pengoptimalan yaitu metode *Particle Swarm Optimization* (PSO). Hasil dari pengujian tersebut akan ditunjukkan dalam *confusion matrix*, dengan detailnya disajikan di tabel 3.

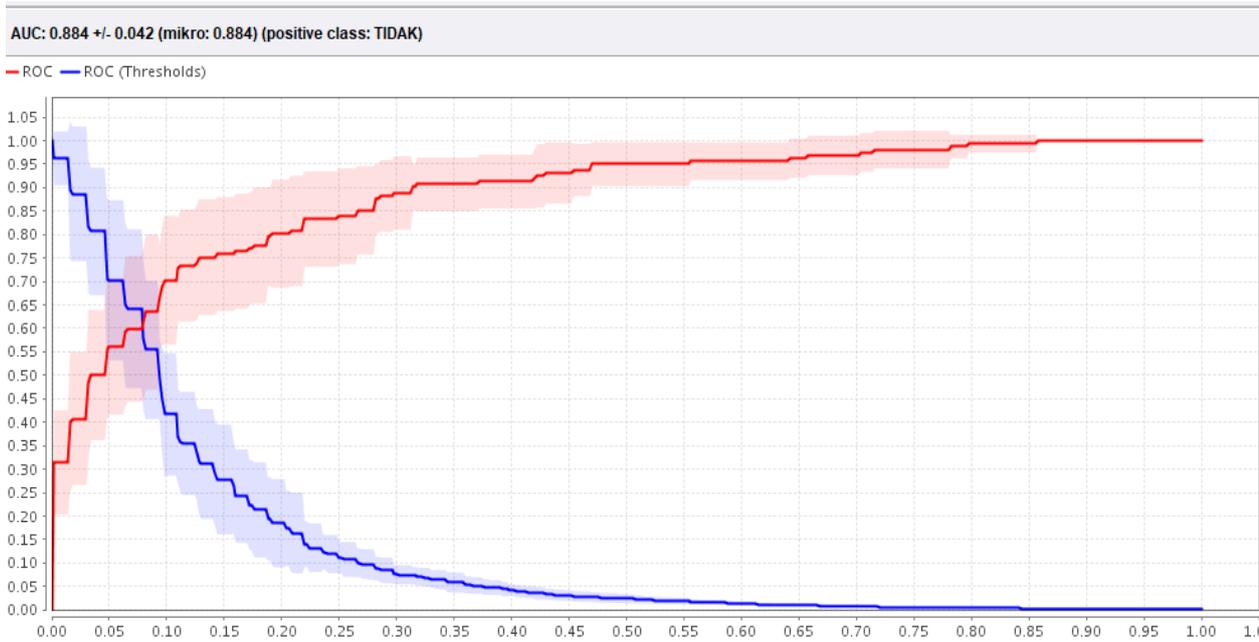
Table 4. *Confussion Matrix* NB+PSO

	<i>true ya</i>	<i>true tidak</i>
<i>pred ya</i>	584	54
<i>pred tidak</i>	50	108

Dari tabel 4 dapat diketahui bahwa Prediksi "ya" (kelulusan) yang benar sebanyak 584 mahasiswa dan hasil dari prediksi tersebut benar bahwa mahasiswa lulus tepat waktu dan prediksi salah terhadap kelulusan sebanyak 54 mahasiswa. Prediksi "tidak" (tidak lulus) yang benar sebanyak 108 kali dan salah sebanyak 50 kali yang seharusnya lulus tepat waktu. Dari tabel 4 maka akan didapatkan kurva tentang kelulusan mahasiswa dengan model NB yang dapat dilihat pada gambar 3.

Gambar 3 Prediksi Kelulusan Mahasiswa dengan kurva tersebut dapat dikatakan bahwa model NB dengan pengoptimalan PSO berhasil mencapai AUC (*Area Under the Curve*) sebesar 0,884% pada kurva ROC, menunjukkan efektivitas tinggi dari model prediksi yang digunakan dalam kelulusan mahasiswa.

Berdasarkan kedua eksperimen yang telah dilakukan, kita dapat menentukan nilai bobot untuk tiap atribut sesuai dengan yang ditampilkan pada Tabel 5.



Gambar 3. Kurva Kelulusan Model NB+PSO

Table 5. Bobot Nilai Kelulusan Mahasiswa NB+PSO

Attribute	Weight
JK	1.0
JRS_SLTA	0.0
ASAL_SLTA	1.0
IPK1	0.0
IPK2	0.0
IPK3	0.0
IPK4	1.0
IPK5	0.6488822536495003
IPK6	0.0

Dalam Tabel 5 tersebut kita dapat melihat bahwa PK1 memiliki bobot 1, PK2 memiliki bobot 0, PK3 memiliki bobot 0,654, dan PK4 memiliki bobot 1 serta PK5 memiliki bobot 0,648 dimana dari atribut tersebut memiliki pengaruh terhadap data kelulusan. Sedangkan atribut yang lain masing-masing memiliki bobot 0.

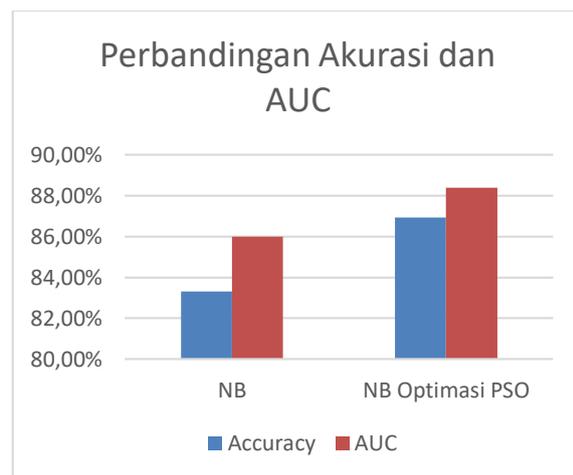
### 3.4. Pengujian dan validasi

Setelah model dioptimalkan, evaluasi akan dijalankan melalui pengujian dan validasi pada hasil dari kedua metode tersebut. Dari kedua algoritma tersebut baik NB atau NB+PSO akan menghasilkan nilai perbandingan yang terlihat tabel 6.

Table 6. Perbandingan Confussion Matrix

Algoritma	NB	NB Optimasi PSO
Accuracy	83.30%	86.94%
AUC	0.860%	0.884%

Dari hasil perbandingan pada tabel 6 bahwa metode NB mengalami kenaikan dari sisi akurasi dan AUC dengan pengoptimalan dengan menggunakan metode PSO yang dapat dilihat pada gambar 4 grafik perbandingan nilai tersebut.



Gambar 4. Perbandingan Nilai Akurasi dan AUC

Dari perbandingan yang ditunjukkan dalam grafik tersebut, terlihat bahwa terjadi peningkatan kinerja ketika *Particle Swarm Optimization* (PSO) diterapkan pada model *Naive Bayes* (NB). Khususnya untuk model NB yang dioptimasi dengan PSO (NB+PSO), terjadi peningkatan akurasi sebesar 3,64 persen dibandingkan dengan model *Naive Bayes* tanpa optimasi. Lebih jauh lagi, *Area Under Curve* (AUC) juga mengalami peningkatan yang signifikan sebesar 0,024 atau 2,40 persen. Kenaikan ini menunjukkan bahwa model yang dioptimasi tidak hanya memprediksi kelas dengan lebih akurat, tetapi juga memiliki kemampuan yang lebih baik dalam membedakan antara kelas positif dan negatif.

## 4. DISKUSI

Penelitian ini menunjukkan bahwa penggunaan metode *Naive Bayes* (NB) yang dioptimalkan dengan *Particle Swarm Optimization* (PSO) memberikan peningkatan yang signifikan dalam akurasi dari

83,30% naik menjadi 86,90% dan AUC (*Area Under Curve*) dari 0,860 naik menjadi 0,884. Peningkatan terjadi juga pada penelitian sebelumnya dimana metode PSO meningkatkan metode sebelumnya sebesar nilai akurasi 86,57%-87,56%. Sehingga dengan melakukan metode NB dan PSO menjadi akurasi tertinggi ketiga dari metode sebelumnya dengan PSO. Ini menegaskan pentingnya penggabungan teknik-teknik canggih dalam pembelajaran mesin untuk meningkatkan hasil prediksi, sesuai dengan temuan penelitian sebelumnya yang juga menunjukkan peningkatan akurasi dengan penggunaan teknik optimasi seperti PSO dan *Genetic Algorithm* (GA).

Penelitian sebelumnya menggunakan metode NN, SVM, DT dengan pengoptimalan PSO memberikan hasil nilai yang beragam. Akan tetapi, tidak ada yang menggunakan metode NB pada data kelulusan tersebut serta tidak ada pengoptimalan pada metode NB dengan PSO untuk membandingkan hasil nilai langsung. Penelitian dimasa mendatang lebih baik membandingkan pemilihan fitur yang lain.

## 5. KESIMPULAN

Penelitian ini mengembangkan sebuah model prediksi kelulusan mahasiswa yang inovatif dengan mengintegrasikan algoritma *Naive Bayes* (NB) dan *Particle Swarm Optimization* (PSO). Hasil penelitian ini menunjukkan peningkatan yang signifikan dalam akurasi dan nilai AUC (*Area Under Curve*) saat menggunakan kombinasi NB dan PSO dibandingkan dengan penggunaan NB secara independen. Spesifiknya akurasi prediksi meningkat menjadi 86,94% untuk model NB+PSO dari 83,30% untuk NB saja dan nilai AUC meningkat dari 0,860 menjadi 0,884 meskipun peningkatan ini tidak terlalu signifikan. Hal ini membuktikan bahwa integrasi NB dan PSO mampu meningkatkan efektivitas dalam mengklasifikasikan kasus Prediksi Kelulusan mahasiswa dengan lebih akurat.

Lebih lanjut, penelitian ini membuka peluang untuk aplikasi praktis teknologi dalam sektor pendidikan, khususnya dalam hal peningkatan akurasi prediksi kelulusan mahasiswa. Implementasi teknik-teknik seperti NB dan PSO dalam analisis data pendidikan menjanjikan potensi yang signifikan untuk pengembangan lebih jauh, termasuk dalam peningkatan keakuratan prediksi dalam berbagai aspek lainnya. Studi ini menggarisbawahi pentingnya penggunaan teknik optimasi dan seleksi fitur yang efektif untuk meningkatkan hasil prediksi. Meskipun demikian, diperlukan evaluasi dan peningkatan model yang lebih mendalam untuk mengatasi keterbatasan yang ada dan memperluas efektivitas model dalam berbagai kondisi aplikasi praktis.

## DAFTAR PUSTAKA

- [1] V. Riyanto, A. Hamid, and Ridwansyah, "Prediction of Student Graduation Time

Using the Best Algorithm," *Indones. J. Artif. Intell. Data Min.*, vol. 2, no. 2, pp. 1–9, 2019, doi: <http://dx.doi.org/10.24014/ijaidm.v2i1.6424>.

- [2] Suhardjono, G. Wijaya, and A. Hamid, "PREDIKSI WAKTU KELULUSAN MAHASISWA MENGGUNAKAN SVM BERBASIS PSO," *Bianglala Inform.*, vol. 7, no. 2, pp. 97–101, 2019, doi: <https://doi.org/10.31294/bi.v7i2.6654.g3731>.
- [3] R. Ridwansyah, G. Wijaya, and J. J. Purnama, "Hybrid Optimization Method Based on Genetic Algorithm for Graduates Students," *J. Pilar Nusa Mandiri*, vol. 16, no. 1, pp. 53–58, 2020, doi: 10.33480/pilar.v16i1.1180.
- [4] Hendra, M. A. Azis, and Suhardjono, "ANALISIS PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN DECISION TREE BERBASIS PARTICLE SWARM OPTIMIZATION," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 1, pp. 102–107, 2020, doi: <https://doi.org/10.32736/sisfokom.v9i1.756>.
- [5] H. Nurdin, Sartini, Sumarna, Y. I. Maulana, and V. Riyanto, "Prediction of Student Graduation with the Neural Network Method Based on Particle Swarm Optimization," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. 4, pp. 2353–2362, 2023, doi: 10.33395/sinkron.v8i4.12973.
- [6] W. E. Pangesti, I. Ariyati, Priyono, Sugiono, and R. Suryadithia, "Utilizing Genetic Algorithms To Enhance Student Graduation Prediction With Neural Networks," *Sink. J. dan Penelit. Tek. Inform.*, vol. 9, no. 1, pp. 276–284, 2024, doi: <https://doi.org/10.33395/sinkron.v9i1.13161> e-ISSN.
- [7] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm," *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [8] Hartatik, "Optimasi Model Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Naive Bayes," *Indones. J. Appl. Informatics*, vol. 5, no. 1, pp. 32–38, 2020.
- [9] S. Salmu and A. Solichin, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naive Bayes: Studi Kasus UIN Syarif Hidayatullah Jakarta," *Semin. Nas. Multidisiplin Ilmu 2017*, no. April, pp. 701–709, 2017.
- [10] J. H. Jaman, "Prediksi Kelulusan Mahasiswa Dengan Metode Algoritma C4.5," *Syntax*, vol. 2, no. 2, pp. 1–6, 2013, doi:

- 10.1002/jhet.1722. no. 5, pp. 1239–1243, 2022, doi: 10.20884/1.jutif.2022.3.5.328.
- [11] M Hafidz Ariansyah, Esmi Nur Fitri, and Sri Winarno, “Improving Performance of Students’ Grade Classification Model Uses Naïve Bayes Gaussian Tuning Model and Feature Selection,” *J. Tek. Inform.*, vol. 4, no. 3, pp. 493–501, 2023, doi: 10.52436/1.jutif.2023.4.3.737.
- [12] N. Widiastuti, A. Hermawan, and D. Avianto, “KOMPARASI ALGORITMA KLASIFIKASI DATAMINING UNTUK PREDIKSI MINAT PENCARI KERJA,” *TEKNOINFO*, vol. 17, no. 1, 2023.
- [13] Ridwansyah, I. Ariyati, and S. Faizah, “PARTICLE SWARM OPTIMIZATION BERBASIS CO-EVOLUSIONER DALAM EVALUASI KINERJA ASISTEN DOSEN,” *J. SAINTEKOM*, vol. 9, no. 2, pp. 165–177, 2019, doi: <https://doi.org/10.33020/saintekom.v9i2.96>.
- [14] I. Ariyati, S. Rosyida, K. Ramanda, V. Riyanto, S. Faizah, and Ridwansyah, “Optimization of the Decision Tree Algorithm Used Particle Swarm Optimization in the Selection of Digital Payments,” in *Journal of Physics: Conference Series*, 2020, vol. 1641, no. 1, doi: 10.1088/1742-6596/1641/1/012090.
- [15] M. Iqbal *et al.*, “Implementation of Particle Swarm Optimization Based Machine Learning Algorithm for Student Performance Prediction,” *JITK (Jurnal Ilmu Pengetah. Dan Teknol. Komputer)*, vol. 6, no. 2, pp. 195–204, 2020, doi: 10.33480/jitk.v6i2.1695.IMPLEMENTATION.
- [16] J. J. Purnama, H. M. Nawawi, S. Rosyida, Ridwansyah, and Risandar, “Klasifikasi Mahasiswa Her Berbasis Algoritma Svm Dan Decision Tree,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1253–1260, 2019, doi: 10.25126/jtiik.202073080.
- [17] R. Ridwansyah, V. Riyanto, A. Hamid, S. Rahayu, and J. J. Purnama, “Grouping Data in Predicting Infant Mortality Using K-Means and Decision Tree,” *Paradigma*, vol. 24, no. 2, pp. 168–174, 2022, doi: 10.31294/paradigma.v24i2.1399.
- [18] Ridwansyah and E. Purwaningsih, “Mengukur Tingkat Error Ketahanan Beton dengan Metode Klasifikasi Neural Network dan Support Vector Machine,” *J. Tek. Inform.*, vol. 5, no. 1, 2019, doi: <https://doi.org/10.51998/jti.v5i1.295>.
- [19] A. S. Yudistira and A. Nugroho, “Prediction of the English Premier League Champion Team for the 2021/2022 Season Using the Naïve Bayes Method,” *J. Tek. Inform.*, vol. 3,