

PENINGKATAN AKURASI KNN DALAM PREDIKSI KELULUSAN MAHASISWA MELALUI OPTIMASI PARAMETER PSO

Yessica Fara Desvia^{1*}; Febryawan Yuda Pratama²; Ganda Wijaya³

Bisnis Digital^{1,2}
Politeknik Jatiluhur, Jawa Barat, Indonesia^{1,2}
<https://polijati.ac.id/>^{1,2}
yessicadesvia@polijati.ac.id^{1*}, febryawanpratama@polijati.ac.id²

Sistem Informasi³
Universitas Nusa Mandiri, Jakarta, Indonesia³
<https://www.nusamandiri.ac.id/>³
ganda.gws@nuasmandiri.ac.id³

(*) Corresponding Author



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-Non Komersial 4.0 Internasional.

Abstract— Predicting student graduation is a crucial aspect in supporting academic planning and ensuring timely completion of studies. However, no prior research has specifically applied the integration of K-Nearest Neighbor (KNN) and Particle Swarm Optimization (PSO) for graduation prediction using student data. This study aims to evaluate the effectiveness of combining KNN and PSO in improving classification accuracy. The KNN algorithm is used for classification, while PSO is implemented as a feature selection technique to identify the most relevant attributes. A dataset of 750 student records was processed through data preprocessing and attribute weighting using PSO, followed by model training and evaluation with 10-fold cross-validation. The evaluation results show that the KNN+PSO model improves accuracy from 80.91% to 84.31%, along with increases in precision and recall. These findings indicate that PSO enhances the performance of KNN, particularly in identifying students likely to graduate on time. This research is to provide a basis for academic management to carry out earlier interventions, such as developing learning assistance programs or more targeted academic policies, so as to increase student graduation rates.

Keywords: classification, feature selection, graduation prediction, K-NN, PSO.

Abstrak— Prediksi kelulusan mahasiswa menjadi aspek penting dalam menunjang perencanaan akademik dan strategi penyelesaian studi. Namun, belum ada penelitian yang secara spesifik mengintegrasikan metode K-Nearest Neighbor (KNN) dengan Particle Swarm Optimization (PSO) untuk menganalisis data kelulusan mahasiswa. Penelitian ini bertujuan untuk mengevaluasi efektivitas kombinasi KNN dan PSO dalam meningkatkan akurasi klasifikasi kelulusan. Metode yang digunakan adalah algoritma KNN untuk klasifikasi dan PSO sebagai teknik seleksi fitur untuk mengidentifikasi atribut paling relevan. Dataset berisi 750 data mahasiswa diproses melalui tahapan *preprocessing* dan pembobotan atribut menggunakan PSO, dilanjutkan dengan pelatihan dan pengujian menggunakan *10-fold cross-validation*. Hasil evaluasi menunjukkan bahwa model KNN+PSO menghasilkan peningkatan akurasi dari 80,91% menjadi 84,31%, serta peningkatan *precision* dan *recall*. Temuan ini menunjukkan bahwa penerapan PSO dapat meningkatkan performa prediksi KNN, khususnya dalam mengidentifikasi mahasiswa yang berpotensi lulus tepat waktu. penelitian ini adalah memberikan dasar bagi pihak manajemen akademik untuk melakukan intervensi lebih dini, seperti penyusunan program pendampingan belajar atau kebijakan akademik yang lebih tepat sasaran, sehingga dapat meningkatkan tingkat kelulusan mahasiswa.

Kata kunci: klasifikasi, seleksi fitur, kelulusan mahasiswa, K-NN, PSO.

PENDAHULUAN

Pendidikan tinggi merupakan salah satu indikator penting dalam peningkatan kualitas sumber daya manusia (Ong & Pedrasa, 2021). Salah satu tantangan yang dihadapi institusi pendidikan adalah mendeteksi dan memprediksi kelulusan mahasiswa secara tepat agar dapat dilakukan intervensi lebih dini dalam proses akademik (Liao, 2023). Kemampuan untuk memprediksi kelulusan mahasiswa menjadi penting karena berdampak langsung terhadap manajemen akademik, efisiensi studi, serta akreditasi institusi (Lin, 2023).

Prediksi kelulusan mahasiswa dapat digunakan sebagai dasar pengambilan keputusan dalam kebijakan akademik, seperti pemberian bimbingan, beasiswa, atau penyesuaian kurikulum (Rawatlal et al., 2022). Dalam praktiknya, pendekatan manual cenderung tidak akurat dan memakan waktu, sehingga dibutuhkan pendekatan berbasis data yang lebih efisien dan sistematis (Karmagatri et al., 2023).

Prediksi kelulusan mahasiswa telah menjadi topik penting dalam pengelolaan akademik berbasis data (Kurniawati & Maulidevi, 2022). Berbagai pendekatan telah dilakukan untuk meningkatkan akurasi prediksi, terutama dengan menggabungkan algoritma klasifikasi dengan metode optimasi (Hamid & Ridwansyah, 2024).

Penelitian menggabungkan Neural Network dengan PSO untuk meningkatkan kinerja model prediktif. Penelitian ini menunjukkan bahwa PSO dapat membantu mencari bobot optimal pada jaringan saraf (Nurdin et al., 2023). Kombinasi Neural Network dan GA dalam prediksi status kelulusan mahasiswa. Hasilnya menunjukkan peningkatan akurasi, tetapi masih menghadapi tantangan dalam konvergensi awal (Pangesti et al., 2024). Dalam konteks lain, Decision Tree juga dioptimasi menggunakan algoritma Genetika, namun model yang dihasilkan cenderung overfitting terhadap data latih (Rukiastiandari et al., 2024).

Naive Bayes juga menjadi alternatif dalam penelitian klasifikasi kelulusan dengan menggunakan NB yang dioptimasi PSO dan menemukan bahwa pendekatan ini dapat meningkatkan presisi klasifikasi (Kahfi et al., 2024). Hal serupa juga dilakukan melalui penerapan Naive Bayes dengan optimasi GA, meskipun pada kasus ini tingkat recall masih rendah (Sumarna et al., 2024).

Pendekatan lain dilakukan dengan mengeksplorasi berbagai teknik machine learning tanpa optimasi, dengan hasil yang beragam tergantung pada atribut input (Ridwansyah et al., 2024). Sementara itu, pendekatan deep learning juga mulai digunakan, namun membutuhkan

dataset besar dan sumber daya komputasi tinggi (Iqbal et al., 2025).

Dari berbagai studi sebelumnya, terlihat bahwa metode optimasi seperti PSO dan GA telah banyak diterapkan untuk SVM, DT, NN, dan NB, namun sangat minim penelitian yang menerapkan PSO secara khusus pada algoritma KNN dalam konteks prediksi kelulusan mahasiswa. Hal ini menjadi celah penelitian yang ingin diisi dalam penelitian ini, yaitu dengan mengembangkan model KNN yang dioptimasi parameter dan fitur-fiturnya menggunakan PSO untuk meningkatkan akurasi dan efisiensi prediksi.

Salah satu algoritma yang cukup efektif untuk klasifikasi adalah K-Nearest Neighbor (KNN) karena kesederhanaannya dan kemampuannya menangani data non-linear. Namun, performa KNN sangat dipengaruhi oleh parameter nilai k dan pemilihan fitur, yang apabila tidak dioptimalkan dapat menyebabkan akurasi menurun. Oleh karena itu, dibutuhkan metode optimasi yang mampu mencari kombinasi parameter terbaik dari KNN agar prediksi lebih akurat.

Salah satu pendekatan yang menjanjikan dalam optimasi parameter adalah Particle Swarm Optimization (PSO). PSO merupakan algoritma metaheuristik yang terinspirasi dari perilaku sosial burung dalam kawanan, dan terbukti efektif dalam mencari solusi optimal di ruang parameter yang kompleks. Penelitian oleh Aryani dkk. (2020) menunjukkan bahwa PSO mampu meningkatkan kinerja model klasifikasi seperti SVM dan Decision Tree dengan cara menyesuaikan parameter dan pemilihan fitur secara dinamis.

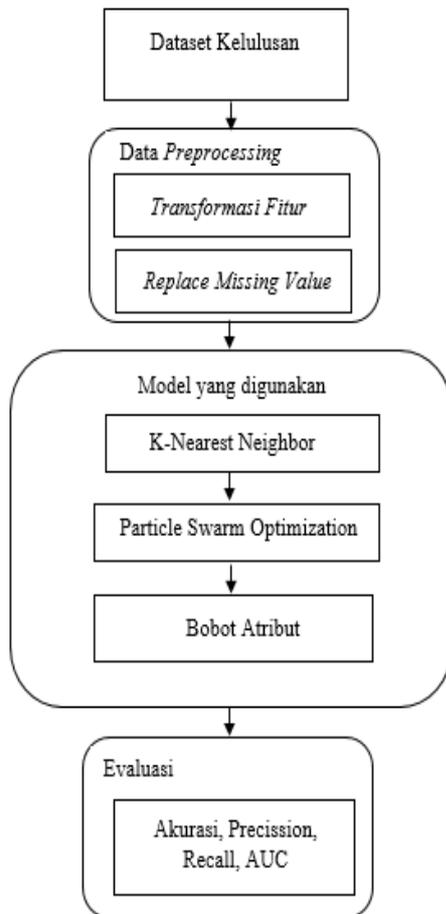
Meskipun banyak algoritma telah diterapkan, masih jarang penelitian yang secara khusus mengintegrasikan KNN dengan PSO untuk prediksi kelulusan mahasiswa. Belum banyak eksplorasi terhadap pengaruh nilai parameter k KNN terhadap akurasi klasifikasi dalam konteks ini, sehingga dibutuhkan pendekatan yang mampu mengoptimasi parameter secara otomatis.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk meningkatkan akurasi prediksi kelulusan mahasiswa dengan mengoptimasi parameter algoritma KNN menggunakan PSO. Optimasi ini diharapkan dapat meningkatkan kinerja klasifikasi dibandingkan dengan KNN standar dan mengkaji sejauh mana kontribusi PSO dalam pengoptimalan metode KNN untuk meningkatkan akurasi dan efisiensi prediksi kelulusan mahasiswa.

BAHAN DAN METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasi

untuk membangun model prediksi kelulusan mahasiswa. Dimana penelitian ini yang nantinya meningkatkan akurasi dan efisiensi proses prediksi. Penelitian ini dilaksanakan melalui tahapan preprocessing data, pelatihan model, pemilihan parameter dan fitur optimal dengan PSO, dan evaluasi kinerja. Alur Penelitian dapat dilihat pada Gambar 1.



Sumber: (Hasil Penelitian, 2025)
Gambar 1. Alur Penelitian

Dataset

Data yang digunakan dalam penelitian ini bersumber dari data akademik mahasiswa suatu perguruan tinggi swasta di Indonesia. Dataset terdiri atas 750 entri mahasiswa dengan 12 atribut input yang mencakup informasi akademik dan administrasi seperti indeks prestasi kumulatif (IPK), Jenis Kelamin, Jurusan dan Asal sekolah sebelum masuk ke perguruan tinggi dalam menentukan status kelulusan. Selain itu, terdapat satu atribut target berupa status kelulusan (tepat waktu atau tidak). Data diperoleh melalui dokumentasi dari sistem informasi akademik

kampus dan telah melalui proses deidentifikasi untuk menjaga kerahasiaan.

Data Preprocessing

Data dikumpulkan dari sistem informasi akademik institusi swasta secara resmi, proses preprocessing data mencakup dua tahap utama yaitu penanganan *missing value* dengan metode imputasi yang sesuai, serta transformasi fitur melalui normalisasi data numerik dan encoding atribut kategorikal agar siap digunakan dalam algoritma KNN.

Model yang digunakan

Dalam proses analisis, algoritma KNN digunakan sebagai model dasar klasifikasi (Rukiastiandari et al., 2025). Untuk mengoptimasi performa model, dilakukan seleksi fitur menggunakan algoritma PSO (Widodo et al., 2024). Seluruh proses implementasi dilakukan menggunakan perangkat lunak RapidMiner Studio, dengan memanfaatkan operator-operator yang mendukung workflow KNN dan PSO (Ridwansyah et al., 2025). Model KNN dijalankan dalam dua skenario: tanpa optimasi (*baseline model*) dan dengan optimasi menggunakan PSO. Dalam penelitian ini, PSO digunakan untuk menentukan bobot fitur (*feature weighting*) sehingga atribut yang lebih relevan memperoleh kontribusi lebih besar dalam proses klasifikasi, sementara atribut yang kurang signifikan memiliki bobot yang lebih kecil. Evaluasi performa model dilakukan menggunakan metrik seperti akurasi, precision, recall, dan AUC (*Area Under Curve*) dengan validasi silang (*cross-validation*) untuk memastikan keandalan hasil. Perbandingan hasil dari kedua skenario ini menjadi dasar dalam menilai efektivitas pendekatan optimasi PSO terhadap KNN dalam konteks prediksi kelulusan mahasiswa.

Evaluasi Model

Evaluasi terhadap performa model dilakukan menggunakan metode 10-fold cross-validation guna memastikan hasil yang diperoleh bersifat generalizable dan tidak bergantung pada satu subset data saja. Dalam pendekatan ini, data dibagi menjadi 10 bagian yang masing-masing secara bergantian menjadi data uji dan data latih. Beberapa metrik evaluasi yang digunakan untuk menilai kinerja model meliputi:

Akurasi (Accuracy)

Mengukur proporsi prediksi yang benar terhadap keseluruhan prediksi:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Presisi (Precision)

Menilai ketepatan model dalam memprediksi kelas positif:

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Recall (Sensitivity)

Mengukur kemampuan model dalam mendeteksi seluruh data yang relevan atau positif:

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

AUC (Area Under Curve)

Merupakan ukuran keseluruhan performa model dalam membedakan kelas. AUC diperoleh dari kurva *Receiver Operating Characteristic* (ROC) yang menggambarkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR).

Seluruh pengujian dilakukan menggunakan perangkat komputer dengan spesifikasi Intel Core i7, RAM 16 GB, dan sistem operasi Windows 10. Evaluasi dilakukan secara terintegrasi dalam platform RapidMiner, di mana perbandingan kinerja model sebelum dan sesudah optimasi dengan PSO dicatat secara menyeluruh. Melalui pendekatan ini, penelitian diharapkan dapat menghasilkan model klasifikasi yang lebih akurat, efisien, dan dapat diandalkan dalam mendukung pengambilan keputusan akademik berbasis data.

HASIL DAN PEMBAHASAN

Dataset

Pada penelitian ini, dataset yang digunakan berasal dari sistem data akademik mahasiswa di sebuah perguruan tinggi swasta di Indonesia. Dataset ini terdiri dari 750 sampel mahasiswa dengan 12 atribut input yang mencakup berbagai informasi terkait akademik dan administrasi, seperti Indeks Prestasi Kumulatif (IPK), jenis kelamin, jurusan, dan asal sekolah sebelum memasuki perguruan tinggi, yang digunakan untuk memprediksi status kelulusan mahasiswa. Atribut target yang digunakan dalam penelitian ini adalah status kelulusan mahasiswa, yaitu apakah mereka lulus tepat waktu atau tidak. Data yang digunakan diperoleh dari dokumentasi sistem informasi

akademik perguruan tinggi yang telah melalui proses deidentifikasi untuk melindungi kerahasiaan data mahasiswa. Dataset dapat dilihat pada Tabel 1.

Tabel 1. Atribut Dataset Kelulusan

Role	Nama	Tipe
Label	JK	binomial
Regular	JRS_SLTA	integer
Regular	ASAL_SLTA	Integer
Regular	IPK1	numeric
Regular	IPK2	numeric
Regular	IPK3	numeric
Regular	IPK4	numeric
Regular	IPK5	numeric
Regular	IPK6	numeric

Sumber: (Hasil Penelitian, 2025)

Dataset ini mengandung dua kelompok utama, yaitu atribut *input* dan atribut target. Atribut *input* berisi informasi mengenai karakteristik akademik dan administratif mahasiswa, seperti IPK, jenis kelamin, jurusan, dan asal sekolah. Semua atribut ini memainkan peran penting dalam proses pembuatan model prediksi kelulusan mahasiswa. Atribut target yang digunakan adalah status kelulusan mahasiswa, yang terbagi menjadi dua kategori: "tepat waktu" atau "tidak tepat waktu".

Selain itu, atribut dalam dataset ini terdiri dari berbagai tipe data, termasuk data numerik, integer, dan binomial. Sebagai contoh, atribut "TEPAT WAKTU" merupakan data binomial yang hanya memiliki dua nilai, yaitu "Ya" atau "Tidak", sedangkan atribut seperti IPK1 hingga IPK6 adalah data numerik yang menunjukkan skor IPK mahasiswa pada berbagai periode.

Data Preprocessing

Proses preprocessing data mencakup dua langkah utama: penanganan missing value dan transformasi fitur. Missing value pada data diatasi menggunakan metode imputasi yang sesuai, sehingga data yang hilang dapat diisi dengan nilai yang relevan. Selain itu, dilakukan juga normalisasi data numerik untuk menyamakan skala setiap fitur numerik, serta encoding untuk atribut kategorikal seperti jenis kelamin, sehingga data dapat diterima oleh algoritma K-Nearest Neighbors (KNN) yang digunakan dalam penelitian ini. Berikut hasil dari *preprocessing data* pada Tabel 2.

Tabel 2. *Sample Dataset* Kelulusan

No	TEPAT WAKTU	JK	JRS_SLTA	ASAL_SLTA	IPK1	IPK2	IPK3	IPK4	IPK5	IPK6
1	YA	1	4	4	2.550	2.610	2.640	2.740	2.970	3.370
2	YA	2	1	3	3.230	3.150	3.070	3.170	3.250	3.170
3	YA	2	4	2	2.730	2.730	2.620	2.640	2.890	3.330
4	YA	1	3	4	3.100	3.100	2.970	3.030	3.330	3.330
5	TIDAK	1	5	2	2.180	2.440	2.120	2.360	2.660	2.570
6	TIDAK	2	9	3	2.320	2.390	2.430	2.740	2.670	2.950
7	YA	1	3	7	3.640	2.220	3.200	3.110	3.490	2.740

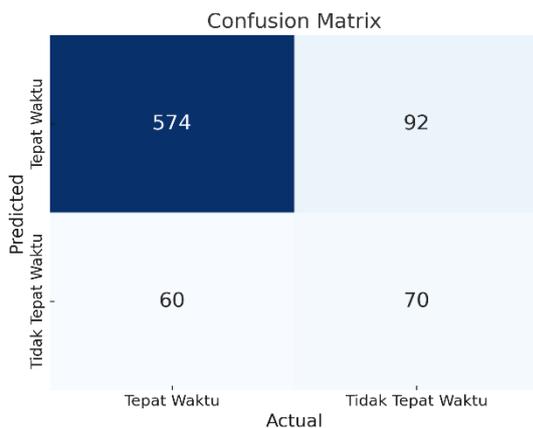
No	TEPAT WAKTU	JK	JRS_SLTA	ASAL_SLTA	IPK1	IPK2	IPK3	IPK4	IPK5	IPK6
8	YA	2	7	3	3.640	3.460	3.330	3.530	3.560	3.670
...
750	YA	1	8	5	3.090	3.120	3.050	3.280	3.320	3.370

Sumber: (Hasil Penelitian, 2025)

Pada Tabel 2. Merupakan hasil dari data yang sudah melalui *processing data*. Proses preprocessing ini penting untuk memastikan kualitas data yang baik dan mempersiapkan data agar siap digunakan dalam membangun model prediksi kelulusan mahasiswa.

Model yang digunakan

Model K-Nearest Neighbor (KNN) digunakan untuk memprediksi kelulusan mahasiswa berdasarkan dataset yang telah melalui tahapan preprocessing dan optimasi fitur. Evaluasi model dilakukan menggunakan confusion matrix untuk mengetahui performa klasifikasi pada masing-masing kelas, yaitu "Tepat Waktu" dan "Tidak Tepat Waktu". Gambar 4 berikut menyajikan *confusion matrix* hasil prediksi model KNN terhadap data uji



Sumber: (Hasil Penelitian, 2025)

Gambar 2. *Confusion Matrix* KNN

Berdasarkan Gambar 2 confusion matrix di atas, diperoleh hasil sebagai berikut: True Positive, dengan hasil 574 mahasiswa berhasil diprediksi tepat waktu dengan benar. False Positive, dengan hasil 92 mahasiswa diprediksi tepat waktu, padahal sebenarnya tidak tepat waktu. True Negative, dengan hasil 70 mahasiswa diprediksi tidak tepat waktu dan benar adanya. False Negative, dengan hasil 60 mahasiswa yang sebenarnya tepat waktu justru diprediksi tidak tepat waktu.

Hasil ini menunjukkan bahwa model KNN cenderung lebih baik dalam mengidentifikasi mahasiswa yang lulus tepat waktu dibandingkan mereka yang tidak tepat waktu. Hal ini terlihat dari jumlah true positive yang tinggi (574), sementara

true negative relatif lebih rendah (70). Nilai false positive yang cukup besar (92) juga menunjukkan adanya kesalahan klasifikasi yang masih perlu diminimalkan.

Secara keseluruhan, model KNN menunjukkan performa yang cukup baik, namun masih memiliki ruang untuk perbaikan, khususnya kesalahan klasifikasi yang perlu dimaksimalkan. Sebagai bagian dari evaluasi performa model KNN, dilakukan juga analisis menggunakan kurva ROC untuk mengetahui seberapa baik model mampu membedakan antara dua kelas, yaitu mahasiswa yang lulus tepat waktu dan tidak lulus tepat waktu. Kurva ROC menggambarkan keseimbangan antara TPR dan FPR pada berbagai nilai ambang batas (*threshold*), yang penting untuk menilai keandalan model dalam klasifikasi biner. Berikut adalah kurva ROC yang dihasilkan oleh model KNN pada gambar 3.



Sumber: (Hasil Penelitian, 2025)

Gambar 3. Kurva ROC KNN

Gambar tersebut menunjukkan bahwa kurva ROC (berwarna merah) hampir mengikuti garis diagonal, yang merepresentasikan model dengan performa setara dengan tebakan acak. Hal ini diperkuat oleh nilai AUC sebesar 0,500, menandakan bahwa model tidak memiliki kemampuan yang signifikan dalam membedakan antara dua kelas target. Selain itu, bayangan merah di sekitar garis ROC mencerminkan ketidakpastian (*confidence interval*) dari prediksi model, yang semakin menunjukkan ketidakstabilan performa KNN dalam eksperimen ini.

Sebagai langkah perbaikan, perlu dilakukan optimasi berbasis PSO. Dengan mengintegrasikan PSO, algoritma KNN dapat melakukan seleksi fitur secara otomatis sehingga hanya atribut-atribut yang paling relevan yang digunakan dalam proses klasifikasi. Hal ini bertujuan untuk mengurangi kompleksitas model, meningkatkan akurasi, serta memperbaiki kemampuan generalisasi terhadap data uji. Evaluasi terhadap model KNN+PSO akan memberikan pemahaman yang lebih dalam mengenai seberapa besar kontribusi dari optimasi ini terhadap kinerja klasifikasi kelulusan mahasiswa.

Tabel 3. Bobot Atribut

Attribute	Waeight
JK	1
JRS_SLTA	0
ASAL_SLTA	1
IPK1	0
IPK2	0.506
IPK3	1
IPK4	0.964
IPK5	1
IPK6	0.164

Sumber: (Hasil Penelitian, 2025)

Tabel 3 menunjukkan hasil bobot atribut yang dihasilkan dari proses seleksi fitur menggunakan KNN yang dioptimasi dengan PSO. Nilai bobot berkisar antara 0 hingga 1, di mana semakin mendekati angka 1 menunjukkan bahwa atribut tersebut lebih relevan atau signifikan dalam proses klasifikasi.

Dari tabel tersebut terlihat bahwa beberapa atribut memiliki bobot 0, seperti JRS_SLTA, ASAL_SLTA, dan IPK1, yang mengindikasikan bahwa atribut-atribut tersebut tidak memberikan kontribusi berarti terhadap performa klasifikasi model. Sementara itu, atribut dengan bobot penuh (1), yaitu JK, IPK3, dan IPK5, dianggap sebagai fitur paling penting dan berpengaruh dalam memprediksi kelulusan mahasiswa.

Atribut lainnya, seperti IPK2 (0.506), IPK4 (0.964), dan IPK6 (0.164), masih memiliki tingkat kontribusi yang bervariasi, namun tetap dipertahankan oleh algoritma PSO dalam proses seleksi karena nilainya berada di atas nol. Hasil ini mengindikasikan bahwa informasi akademik mahasiswa seperti indeks prestasi semester akhir (IPK3, IPK4, IPK5) berperan dominan dibandingkan latar belakang pendidikan sebelumnya.

Dengan menggunakan bobot ini, model KNN dapat difokuskan hanya pada fitur-fitur yang signifikan, sehingga diharapkan mampu meningkatkan performa klasifikasi dibandingkan model KNN tanpa seleksi fitur. Dari bobot tersebut dapat menghasilkan *confusion matrix* yang dapat dilihat pada Gambar 4.

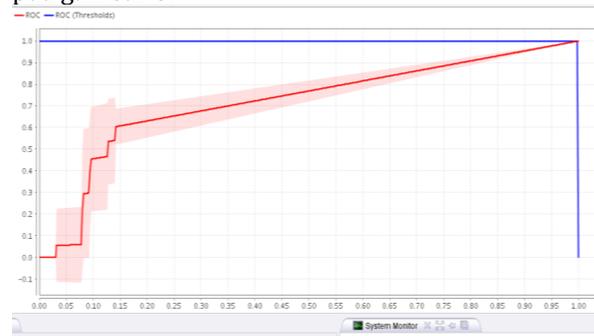


Sumber: (Hasil Penelitian, 2025)

Gambar 4. *Confusion Matrix* KNN+PSO

Confusion matrix gambar 4 menunjukkan hasil klasifikasi model KNN setelah dilakukan proses seleksi fitur menggunakan PSO. Dari hasil prediksi terhadap data uji, diperoleh 577 mahasiswa berhasil diprediksi tepat waktu secara benar (True Positive), dan 94 mahasiswa diprediksi tidak tepat waktu dengan benar pula (True Negative). Sebaliknya, terdapat 68 mahasiswa yang diprediksi tepat waktu, padahal sebenarnya tidak tepat waktu (False Positive), serta 57 mahasiswa yang sebenarnya tepat waktu, namun diprediksi tidak tepat waktu oleh model (False Negative).

Proses klasifikasi KNN dengan bantuan optimasi PSO, hanya fitur-fitur yang paling relevan digunakan dalam perhitungan jarak antar data. Hal ini membuat model KNN+PSO mampu menghitung kedekatan antar mahasiswa secara lebih tepat, sehingga klasifikasi menjadi lebih akurat. Berikut adalah kurva ROC yang dihasilkan oleh model KNN pada gambar 5.



Sumber: (Hasil Penelitian, 2025)

Gambar 5. Kurva ROC KNN+PSO

Pada Gambar 5, meskipun nilai AUC yang diperoleh untuk model KNN yang dioptimasi dengan PSO tetap berada pada -0.500, yang mengindikasikan bahwa model tidak dapat membedakan dengan baik antara kelas positif dan negatif, terdapat peningkatan akurasi yang signifikan dibandingkan dengan model KNN tanpa optimasi. Hal ini menunjukkan bahwa meskipun PSO tidak meningkatkan kemampuan model dalam

hal AUC, optimasi ini berhasil memperbaiki pengklasifikasian secara keseluruhan.

Evaluasi

Untuk mengukur performa model KNN dan KNN yang telah dioptimasi menggunakan Particle Swarm Optimization (PSO), dilakukan evaluasi berdasarkan metrik. Tabel 4 berikut menyajikan perbandingan hasil evaluasi antara model KNN dengan model KNN yang telah dioptimasi menggunakan PSO:

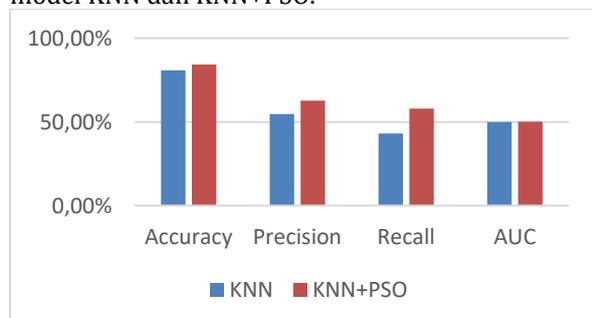
Tabel 4. Perbandingan Evaluasi Model

Metode	KNN	KNN+PSO
Accuracy	80,91%	84,31%
Precision	54,64%	62,79%
Recall	43,12%	57,94%
AUC	50,00%	50,20%

Sumber: (Hasil Penelitian, 2025)

Dari hasil evaluasi di atas, dapat disimpulkan bahwa penerapan PSO pada algoritma KNN memberikan peningkatan performa pada sebagian besar metrik evaluasi: Akurasi meningkat dari 80,91% menjadi 84,31%, menandakan bahwa model hasil optimasi PSO mampu memprediksi label secara lebih tepat. Precision naik dari 54,64% menjadi 62,79%, menunjukkan bahwa model KNN+PSO lebih akurat dalam memprediksi mahasiswa yang lulus tepat waktu. Recall meningkat dari 43,12% ke 57,94%, mengindikasikan kemampuan model dalam menemukan kasus yang benar-benar tepat waktu menjadi lebih baik.

Meskipun nilai AUC pada kedua model masih sama (0,500), hal ini menunjukkan bahwa performa pemisahan kelas oleh model masih perlu ditingkatkan. PSO secara efektif membantu KNN dalam memilih atribut yang lebih relevan, sehingga menghasilkan klasifikasi yang lebih akurat secara umum. Untuk memberikan visualisasi yang lebih jelas, grafik batang pada Gambar 6 yang menunjukkan perbandingan metrik evaluasi antara model KNN dan KNN+PSO.



Sumber: (Hasil Penelitian, 2025)

Gambar 6. Diagram Perbandingan Metode

Dari grafik tersebut tampak bahwa model KNN+PSO memiliki performa yang unggul secara konsisten dibandingkan KNN murni. Perbedaan paling mencolok terlihat pada metrik recall, yang menunjukkan peningkatan kemampuan model dalam mengidentifikasi mahasiswa yang lulus tepat waktu. Hasil ini memperkuat temuan bahwa pendekatan optimasi berbasis PSO mampu meningkatkan kualitas prediksi model KNN secara signifikan, meskipun aspek AUC tetap menjadi catatan untuk pengembangan lebih lanjut.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan metode K-Nearest Neighbor (KNN) yang dioptimasi dengan Particle Swarm Optimization (PSO) mampu meningkatkan performa prediksi kelulusan mahasiswa. Model KNN murni menunjukkan akurasi yang cukup baik, namun masih kurang optimal dalam aspek precision dan recall. Setelah dilakukan optimasi menggunakan PSO, kinerja model mengalami peningkatan, terutama pada kemampuan dalam mengenali mahasiswa yang lulus tepat waktu. Hal ini dibuktikan dengan meningkatnya nilai akurasi menjadi 84,31%, precision sebesar 62,79%, dan recall sebesar 57,94%. Meskipun nilai AUC tidak mengalami peningkatan signifikan, penggunaan PSO berhasil membantu proses seleksi fitur yang lebih relevan, sehingga menghasilkan prediksi yang lebih tepat dan stabil. Temuan ini menunjukkan bahwa integrasi metode optimasi dalam model klasifikasi konvensional dapat menjadi pendekatan yang efektif dalam pengambilan keputusan berbasis data, khususnya dalam konteks prediksi akademik. Penelitian selanjutnya dapat diterapkan dengan menggabungkan PSO dengan algoritma lain selain KNN.

REFERENSI

- Hamid, A., & Ridwansyah. (2024). Optimizing Heart Failure Detection: A Comparison between Naive Bayes and Particle Swarm Optimization. *Paradigma*, 26(1), 30–36. <https://doi.org/https://doi.org/10.31294/p.v26i1.3284>
- Iqbal, M., Yumi Novita Dewi, Lisnawanty, Maisyaroh, & Suhardjono. (2025). Optimalisasi Prediksi Dalam Kelulusan Berbasis Deep Learning: Perbandingan Kinerja Multi-Layer Perceptron dan Deep Neural Network. *Infotek: Jurnal Informatika Dan Teknologi*, 8(2), 630–641. <https://doi.org/DOI>

- 10.29408/jit.v8i2.30756
- Kahfi, A. H., Prihatin, T., Yudhistira, Sudradjat, A., & Wijaya, G. (2024). THE RIGHT STEPS TOWARDS GRADUATION: NB-PSO SMART COMBINATION FOR STUDENT GRADUATION PREDICTION. *Jurnal Teknik Informatika (Jutif)*, 5(2), 607–614. <https://doi.org/https://doi.org/10.52436/1.jutif.2024.5.2.1889>
- Karmagatri, M., Kurnianingrum, D., Suciana, M. R., & Utami, S. A. (2023). Predicting Factors Related to Student Performance Using Decision Tree Algorithm. *International Conference on Cybernetics and Intelligent System (ICORIS)*. <https://doi.org/10.1109/ICORIS60118.2023.10352269>
- Kurniawati, G., & Maulidevi, N. U. (2022). Multivariate Sequential Modelling for Student Performance and Graduation Prediction. *International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. <https://doi.org/10.1109/ICITACEE55701.2022.9923971>
- Liao, X. (2023). Research and enlightenment on the Graduation Rate of American Universities Based on Panel Data. *International Conference on Management Engineering, Software Engineering and Service Sciences (ICMSS)*. <https://doi.org/10.1109/ICMSS56787.2023.10118079>
- Lin, K. (2023). Research on Graduation Destination Prediction Algorithm Based on Students' Learning Behavior Data. *Asian Conference on Artificial Intelligence Technology (ACAIT)*. <https://doi.org/10.1109/ACAIT60137.2023.10528473>
- Nurdin, H., Sartini, Sumarna, Maulana, Y. I., & Riyanto, V. (2023). Prediction of Student Graduation with the Neural Network Method Based on Particle Swarm Optimization. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 8(4), 2353–2362. <https://doi.org/10.33395/sinkron.v8i4.12973>
- Ong, D. P., & Pedrasa, J. R. I. (2021). Student Risk Assessment: Predicting Undergraduate Student Graduation Probability Using Logistic Regression, SVM, and ANN. *TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON)*. <https://doi.org/10.1109/TENCON54134.2021.9707322>
- Pangesti, W. E., Ariyati, I., Priyono, Sugiono, & Suryadithia, R. (2024). Utilizing Genetic Algorithms To Enhance Student Graduation Prediction With Neural Networks. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 9(1), 276–284. <https://doi.org/https://doi.org/10.33395/sinkron.v9i1.13161> e-ISSN
- Rawatlal, R., Chetty, M., & Naicker, A. K. (2022). Latent Factors for Consistently Predicting Student Success. *World Engineering Education Forum - Global Engineering Deans Council (WEEF-GEDC)*. <https://doi.org/10.1109/WEEF-GEDC54384.2022.9996209>
- Ridwansyah, Andharsaputri, R. L., Yudhistira, Irmawati Carolina, & Suharjanti. (2025). K-Nearest Neighbors Optimization using Particle Swarm Optimization in Selection Digital Payments. *Jurnal Teknologi Informasi Dan Terapan (J-TIT)*, 12(1), 1–8. <https://doi.org/https://doi.org/10.25047/jtit.v12i1.5911>
- Ridwansyah, Iqbal, M., Destiana, H., Sugiono, & Hamid, A. (2024). Data Mining Berbasis Machine Learning Untuk Analitik Prediktif Dalam Kelulusan. *SemanTIK*, 10(2), 1–10. <https://doi.org/https://doi.org/10.55679/semantik.v10i2.67>
- Rukiastindari, S., Luthfia Rohimah, Aprillia, Chodidjah, & Fara Mutia. (2025). Model Hibrida K-Nearest Neighbors Berbasis Genetic Algorithm untuk Prediksi Penyakit Ginjal Kronis. *Infotek: Jurnal Informatika Dan Teknologi*, 8(1), 44–55. <https://doi.org/10.29408/jit.v8i1.27918>
- Rukiastindari, S., Rohimah, L., Aprillia, A., & Mutia, F. (2024). Predicting Graduation Outcomes: Decision Tree Model Enhanced with Genetic Algorithm. *Paradigma - Jurnal Komputer Dan Informatika*, 26(1), 1–6. <https://doi.org/10.31294/p.v26i1.3165>
- Sumarna, Nawawi, I., Suhardjono, Hari Sugiarto, & Yuliandari, D. (2024). MENINGKATKAN AKURASI PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN METODE ALGORITMA GENETIKA. *Jurnal Informatika, Manajemen Dan Komputer*, 16(2). <https://doi.org/http://dx.doi.org/10.36723/juri.v16i2.706>
- Widodo, S., Brawijaya, H., & Samudi, S. (2024). Building a Predictive Model for Chronic Kidney Disease: Integrating KNN and PSO. *Paradigma - Jurnal Komputer Dan Informatika*, 26(1), 58–64. <https://doi.org/10.31294/p.v26i1.3282>