Journal of Artificial Intelligence and Software Engineering
Vol. 6, No. 2, Juni 2026, pp. 185~194
E-ISSN: 2777-001X, DOI: 10.30811/jaise.v6i2.9179

Prediction Analysis of Kip Student Placement at Universitas Prima Indonesia Using SVM Hyperparameter Optimization Method and Comparative Study

Oloan Sihombing[1]*, Calvin Wahyu Febrian Sidabutar[2], Angelica[3], Rico Halim[4] , Daniel Agus Towi Sitompul[5]
[12345] Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Prima Indonesia, Medan, 20118, Indonesia

Informasi Artikel

Diterima : 20 Mei 2026
Revisi : 29 Mei 2026
Publikasi : 30 Juni 2026

Kata Kunci:

Support Vector Machine
K-Nearest Neigbor
KIP
Classification
Machine Learning

ABSTRAK

Penelitian ini bertujuan membangun model prediksi penempatan mahasiswa penerima Kartu Indonesia Pintar (KIP) di Fakultas Sains dan Teknologi Universitas Prima Indonesia menggunakan algoritma Support Vector Machine (SVM) yang dioptimasi hyperparameternya dan dibandingkan dengan K-Nearest Neighbor (KNN). Penelitian menggunakan 269 data mahasiswa yang diperoleh dari instansi akademik dan kuesioner. Tahapan penelitian meliputi pra-pemrosesan data, pembentukan label target, encoding, normalisasi, serta pembagian data 80:20. Optimasi SVM dilakukan menggunakan Grid Search dan 5-fold cross-validation, sedangkan KNN digunakan tanpa optimasi. Hasil penelitian menunjukkan bahwa SVM dengan kernel RBF, parameter C=10 dan gamma=0,1 memperoleh akurasi 88,89%, lebih tinggi dibandingkan KNN sebesar 83,33%. Selain itu, SVM juga menunjukkan performa lebih baik pada metrik precision, recall, dan F1-score, sehingga lebih efektif digunakan untuk mendukung pengambilan keputusan penempatan mahasiswa KIP secara objektif dan tepat sasaran.

ABSTRACT

This study aims to develop a prediction model for the placement of Kartu Indonesia Pintar (KIP) recipient students at the Faculty of Science and Technology, Universitas Prima Indonesia, using the optimized Support Vector Machine (SVM) algorithm and comparing it with the K-Nearest Neighbor (KNN) algorithm. The study used 269 student records obtained from academic institutions and questionnaires. The research stages included data preprocessing, target label construction, encoding, normalization, and data splitting with an 80:20 ratio. SVM optimization was carried out using Grid Search combined with 5-fold cross-validation, while KNN was used without optimization. The results showed that SVM with the RBF kernel, C = 10, and gamma = 0.1 achieved an accuracy of 88.89%, outperforming KNN, which achieved 83.33%. In addition, SVM demonstrated better performance in precision, recall, and F1-score metrics, making it more effective for supporting objective and accurate decision-making in KIP student placement.

Penulis Koresponden

Email: oloansihombing@unprimdn.ac.id

Cara sitasi IEEE:

[1] O. Sihombing, C. W. F. Sidabutar, A. Angelica, R. Halim, D. A. T. Sitompul, ”Analisis Prediksi Penempatan Mahasiswa KIP Pada Fakultas Science dan Technology Universitas Prima Indonesia Metode Optimasi Hyperparameter SVM dan Studi Komparatif,” *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 6, no. 2, p. 185-194, Juni 2026. doi:10.30811/jaise.v6i2.9179

1. PENDAHULUAN

Pendidikan adalah cara yang sistematis untuk mentransfer ilmu dari satu individu kepada individu lainnya dengan kriteria yang telah ditetapkan. Pendidikan berpotensi menjadi sarana untuk memperbaiki mutu sumber daya manusia. Mendapatkan akses terhadap pendidikan adalah hak yang dimiliki oleh setiap individu dan telah diatur dalam Pasal 31 Undang-Undang Dasar Tahun 1945 [1]. Beberapa elemen seperti kekurangan dana, minimnya akses ke fasilitas pendidikan, dan kurangnya dukungan pendidikan dari keluarga juga merupakan salah satu halangan dalam pertumbuhan anak-anak ini [2]. Untuk mengatasi permasalahan tersebut, Pemerintah memperkenalkan sebuah program beasiswa Kartu Indonesia Pintar (KIP) yang merupakan pengembangan dari program Bantuan Siswa Miskin (BSM), KIP adalah salah satu jenis beasiswa yang disediakan oleh pemerintah untuk mendukung calon mahasiswa yang memiliki keterbatasan finansial dalam meneruskan studi mereka [3].

Program KIP dilaksanakan berdasarkan regulasi Permendikbud No. 10 Tahun 2020 tentang Program Indonesia Pintar. Regulasi ini menetapkan bahwa bantuan pendidikan diberikan untuk memastikan individu dari latar belakang ekonomi rendah dapat melanjutkan studi mereka hingga perguruan tinggi [4]. Selain itu, regulasi tersebut menegaskan bahwa perguruan tinggi penyelenggara Program Indonesia Pintar (PIP) melalui KIP kuliah harus memastikan penerima bantuan mendapatkan layanan pendidikan yang sesuai dan tepat sasaran, sesuai dengan minat serta kompetensinya [5].

Universitas Prima Indonesia (UNPRI), terutama fakultas Science and Technology, merupakan tempat yang menyelenggarakan program KIP. Akan tetapi, ketika melakukan proses penempatan mahasiswa di dalam program studi yang dituju, cara pengambilan keputusan selama ini belum sepenuhnya mengandalkan analisis data yang baik. Masalah ini bisa saja menyebabkan penempatan yang kurang tepat, sehingga mempengaruhi semangat belajar, capaian akademis, dan juga kelulusan para mahasiswa yang tidak tepat waktu atau pada keadaan yang paling buruk dapat meningkatkan resiko putus kuliah dikarenakan ketidakcocokan dalam program studi yang telah ditentukan

Support Vector Machine (SVM) merupakan algoritma pembelajaran mesin yang sering digunakan untuk mengklasifikasikan data dan menganalisis sentimen, serta dapat mencapai tingkat akurasi yang tinggi dalam berbagai kajian. Metode ini dapat diterapkan pada berbagai jenis data seperti data yang berbentuk linear atau nonlinear dengan memanfaatkan fungsi kernel untuk membawa data ke dimensi yang lebih tinggi. Ini membuat batas pemisah (*classifier*) menjadi lebih baik [6]. Berdasarkan penelitian sebelumnya SVM juga mampu digunakan dalam memprediksi kemampuan akademik mahasiswa berdasarkan data karakteristik individu seperti gaya belajar dan multiple intelligence, sehingga dapat mendukung proses penentuan kelayakan atau penempatan mahasiswa pada program tertentu [7]. Selain itu, SVM mampu mengklasifikasikan periode studi dan predikat kelulusan mahasiswa dengan tingkat akurasi yang tinggi, sehingga berpotensi dimanfaatkan sebagai alat bantu evaluasi akademik yang lebih sistematis dan objektif. Berdasarkan temuan tersebut, penerapan SVM berpotensi besar untuk diadaptasi penempatan mahasiswa penerima KIP [8].

Namun, kinerja SVM sangat bergantung pada parameter yang digunakan dan juga membutuhkan penyesuaian. Tujuannya adalah untuk meningkatkan akurasi model. Misalnya, kita dapat menggunakan metode Grid Search. Metode Random Search juga digunakan dan teknik ini terbukti mampu meningkatkan performa yang didapatkan model SVM [9]. Untuk memperoleh model yang lebih optimal dan benar-benar cocok untuk keperluan akademis, perlu dilakukan studi komparatif atau bisa disebut dengan perbandingan antara metode seperti Random Forest dan K-Nearest Neighbor (KNN) [10]. Tujuan Penelitian ini dilakukan untuk menguraikan dan mengoptimalkan model prediksi penentuan penempatan mahasiswa KIP kuliah yang lebih akurat, tidak memihak, dan sejalan dengan objektif aturan pendidikan yang berlaku di Indonesia.

2. METODE

2.1 Tahapan Penelitian

Tahapan riset ini dilaksanakan untuk mengevaluasi performa algoritma SVM dan KNN dalam menentukan metode klasifikasi dengan tingkat akurasi terbaik. Proses dimulai dengan penetapan dataset, kemudian dilanjutkan dengan pra-pemrosesan data. Data yang sudah siap selanjutnya dibagi menjadi dua bagian, yaitu data untuk latihan juga data untuk penguji. Kedua algoritma diterapkan secara terpisah untuk melakukan klasifikasi. Hasil yang diperoleh dievaluasi menggunakan metrik pengukuran kinerja, khususnya akurasi. Analisis perbandingan dilakukan untuk menentukan algoritma yang menunjukkan performa paling optimal.

2.2 Metode Pengumpulan Data

Melakukan pengumpulan informasi agar bisa memperoleh data yang diperlukan yang dibutuhkan untuk mendukung kebutuhan penelitian. Setiap penelitian memerlukan data yang valid agar hasilnya dapat dibuktikan dan dipertanggungjawabkan di kemudian hari. Terdapat beberapa metode - metode yang diterapkan dalam proses pengumpulan data, antara lain:

1. Observasi

Pengamatan dilakukan secara langsung terhadap objek penelitian guna menemukan informasi yang relevan. Pada tahap ini, peneliti melakukan observasi terhadap kondisi dan karakteristik data yang digunakan sebagai bahan penelitian [11].

2. Studi Dokumentasi

Data dikumpulkan melalui analisis dokumen dengan cara memperoleh data langsung dari unit akademik di kampus. Teknik dokumentasi digunakan karena dapat menyediakan data yang tepat dan terstruktur sehingga mendukung proses analisis dalam penelitian [12].

3. Kuesioner

Kuesioner dimanfaatkan untuk mengumpulkan informasi secara tertulis dari respon dengan cara menyebarkan kepada responden yang bersangkutan dengan objek penelitian dengan tujuan mendapatkan data tambahan terkait karakteristik dan faktor yang memengaruhi pengolahan data. Data yang didapat melalui kuesioner berfungsi sebagai dukungan dalam analisis penelitian untuk memperkuat temuan dan diskusi yang ada [13].

4. Pengumpulan Data

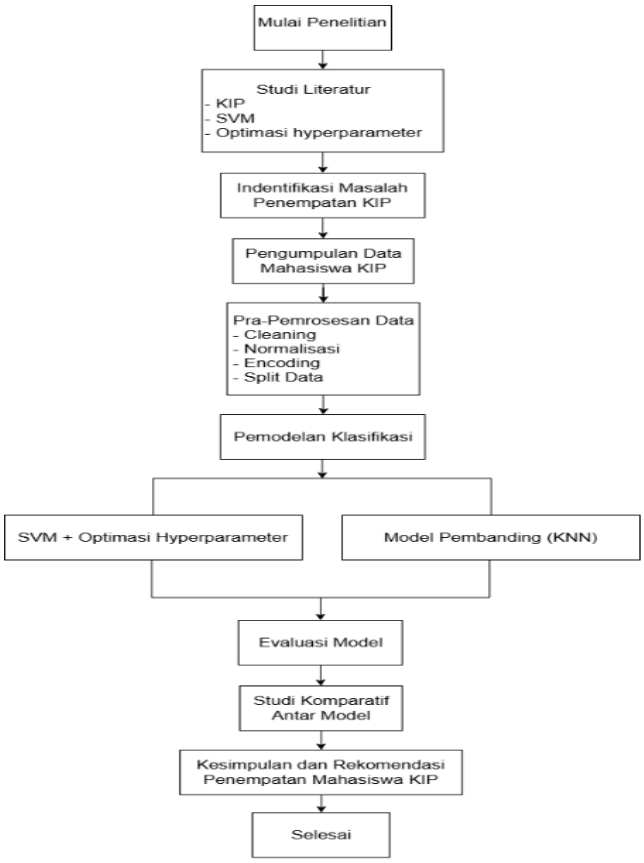
Data dalam penelitian ini diperoleh dari melalui analisis dokumen dengan mengumpulkan data secara langsung dari unit akademik UNPRI. Data juga diperoleh dengan mengedarkan kuesioner kepada responden guna memperoleh informasi tambahan yang mendukung penelitian ini.

5. Analisis Data

Setelah data diperoleh, peneliti melanjutkan ke tahap analisis data dengan menyesuaikan data sesuai dengan kebutuhan penelitian. Analisis dilakukan untuk mengetahui algoritma yang menghasilkan performa terbaik berdasarkan hasil pengujian.

6. Pengolahan Data

Pengolahan data dilakukan dengan menerapkan algoritma SVM dan KNN. Hasil pengolahan data dari kedua algoritma kemudian dibandingkan untuk menentukan metode yang paling optimal.



Gambar 1. Diagram Penelitian

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Data Penelitian

Penelitian ini memanfaatkan data mahasiswa penerima KIP pada Fakultas Science dan Technology UNPRI yang diperoleh dari Unit Akademik Kampus serta Kuesioner yang telah disebarkan kepada sejumlah mahasiswa penerima KIP UNPRI. Dataset terdiri dari 269 record mahasiswa yang disusun dalam format . xlsx menggunakan Microsoft Excel dan selanjutnya diimpor ke Google Colab menggunakan Library Pandas untuk proses analisis. Dataset yang dipergunakan mencakup berbagai atribut, yaitu Program Studi, Jenjang, Semester, Tahun, Penghasilan orang tua, Indeks Prestasi, Absensi, dan Tempat Tinggal.Berdasarkan eksplorasi awal, tidak ditemukan nilai hilang (missing values) pada seluruh atribut, sehingga data siap diproses lebih lanjut

3.2 Pra-Pemrosesan Data

3.2.1 Pembersihan data

Tahap awal pembersihan data dilakukan dengan menghilangkan kolom yang tidak penting untuk proses klasifikasi. Kolom UNIVERSITAS, NAMA, dan NIM dieliminasi karena bersifat unik dan tidak berkorelasi dengan target prediksi.

3.2.2 Pembuatan Label Target

Kolom target penempatan tidak tersedia secara eksplisit dalam dataset, sehingga dilakukan rekayasa fitur berdasarkan kriteria kelayakan penerimaan KIP. Berdasarkan kebijakan umum program KIP Kuliah, kriteria utama kelayakan adalah penghasilan orang tua kurang dari Rp3.000.000 per bulan. Penelitian tentang hasil memprediksi penerima KIP di Universitas Adzkia juga menerapkan batas penghasilan Rp3.000.000 sebagai kriteria utama kelayakan [14]. Berdasarkan kriteria tersebut, ditetapkan mahasiswa layak menerima KIP adalah mereka yang memiliki penghasilan orang tua kurang dari Rp3.000.000. Indeks prestasi tidak digunakan sebagai kriteria dalam penelitian ini karena fokus utama adalah pada aspek ekonomi sebagai determinan utama kelayakan KIP.Distribusi label yang diperoleh:

Table 1. Kategori Kelayakan Manusia

Kategori	Label	Jumlah Mahasiswa	Persentase
Layak	1	98	36,43%
Tidak Layak	0	171	63,57%
Total		**269**	**100%**

Distribusi ini menunjukkan ketidakseimbangan kelas (*class imbalance*) di mana kelas Layak hanya mewakili 36% dari total dataset. Kondisi serupa juga ditemukan dalam penelitian penerimaan beasiswa KIP di berbagai perguruan tinggi, di mana jumlah penerima umumnya lebih sedikit dibanding total pendaftar [15].

3.2.3 Encoding Data Kategorikal

Variabel kategorikal yang terdiri dari PROGRAM STUDI (10 kategori), JENJANG (hanya S1), dan TEMPAT TINGGAL (kota/desa) ditransformasi ke bentuk numerik menggunakan teknik Label Encoding. Pemilihan metode ini didasarkan pada efisiensi komputasi dan kesesuaian untuk variabel dengan jumlah kategori terbatas.

3.2.4 Normalisasi Data

Proses normalisasi diterapkan menggunakan metode *StandardScaler* yang mengubah data agar memiliki nilai rata-rata 0. dan standar deviasi 1. Tahapan ini menjadi krusial khususnya bagi algoritma SVM yang sangat peka terhadap skala data. Fitur yang dinormalisasi meliputi SEMESTER, TAHUN, PENGHASILAN ORANG TUA, INDEKS PRESTASI, ABSENSI, dan ketiga fitur hasil encoding.

3.2.5 Pembagian Data

Dataset dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian dengan rasio 80:20 menggunakan teknik *stratified split* agar proporsi kelas target tetap terjaga. Pembagian ini menghasilkan:

Table 2. Pembagian data Training dan Testing

Jenis Data	Jumlah Data	Persentase
Data Latih	215	80%
Data Uji	54	20%
Total	**269**	**100%**

Penggunaan stratified split penting dilakukan mengingat ketidakseimbangan kelas, sehingga model tetap mendapatkan representasi proporsional dari kedua kelas selama pelatihan. Penelitian prediksi penerimaan

beasiswa KIP di Universitas Adzkia menunjukkan bahwa pembagian data dengan rasio 80:20 menghasilkan kinerja yang paling optimal dibanding rasio lainnya [14].

3.3 Pemodelan dengan SVM dan Optimasi Hyperparameter

3.3.1 Konfigurasi Grid Search

Optimasi hyperparameter pada algoritma SVM diterapkan dengan metode grid search dengan 5-fold *cross-validation*. Rentang nilai hyperparameter yang diuji meliputi:

Table 3. Parameter dan Nilai Pengujian Model SVM

No	Parameter	Nilai Parameter
1	C (Regularization Parameter)	0.1, 1, 10, 100
2	Gamma	1, 0.1, 0.01, 0.001
3	Kernel	Rbf, poly, sigmoid

Pemilihan rentang parameter dalam studi ini merujuk pada penelitian terdahulu yang mengindikasikan bahwa kinerja SVM sangat bergantung pada angka - angka seperti C dan Gamma, sehingga penting untuk mengeksplorasi berbagai kombinasi nilai dalam rentang tertentu agar dapat memperoleh hasil yang optimal [16].

3.3.2 Hasil Optimasi SVM

Setelah melalui proses grid search, diperoleh kombinasi hyperparameter terbaik:

Table 4. Hasil Penentuan Parameter Terbaik pada Model SVM

Parameter	Nilai Parameter
Kernel	Radial Basic Function (RBF)
C (Regularization Parameter)	10
Gamma	0.1

Pemilihan kernel RBF sebagai kernel optimal konsisten dengan temuan penelitian sebelumnya yang menyatakan bahwa kernel RBF mampu mengkategorikan data ke dalam dimensi lebih tinggi dan mengatasi hubungan non-linear antar fitur dengan lebih baik dibanding kernel linear atau polinomial [17]. Kernel RBF juga terbukti efektif dalam berbagai konteks prediksi akademik karena fleksibilitasnya dalam menangani pola data yang kompleks [18].

Akurasi rata-rata cross-validation terbaik yang dicapai model SVM adalah 87,44%. Hasil ini menunjukkan bahwa model SVM dengan konfigurasi optimal mampu menggeneralisasi pola data dengan baik dan konsisten di berbagai fold data. Penelitian tentang prediksi kelulusan mahasiswa menggunakan SVM dengan optimasi grid search juga melaporkan bahwa model yang dihasilkan cukup konsisten dan mencegah overfitting [18].

3.3.3 Evaluasi Model SVM pada Data Uji

Pengujian model SVM terbaik terhadap data uji (54 sampel) menghasilkan akurasi sebesar 88,89%, yang berarti model berhasil memprediksi dengan benar sebanyak 48 dari 54 sampel uji. Hasil ini melampaui target penelitian yaitu akurasi di atas 80%. Disajikan pada Tabel 3.5 untuk confusion matrix model SVM:

Table 5. Confusion Matrix Model SVM

Kategori Kelayakan	Prediksi Tidak Layak	Prediksi Layak
Aktual Tidak Layak	30 (TN)	4 (FP)
Aktual Layak	2 (FN)	18 (TP)

Berdasarkan *confusion matrix* tersebut, metrik evaluasi yang diperoleh:

- **Presisi (Layak)** = TP/(TP+FP) = 18/(18+4) = 81,82%
- **Recall (Layak)** = TP/(TP+FN) = 18/(18+2) = 90,00%
- **F1-Score (Layak)** = 2 × (Presisi × Recall)/(Presisi + Recall) = 85,71%

- **Spesifisitas** = TN/(TN+FP) = 30/(30+4) = 88,24%

Nilai *recall* 90,00% untuk kelas Layak mengindikasikan bahwa model mampu mengidentifikasi 18 dari 20 mahasiswa yang benar-benar layak menerima KIP. Dalam konteks penempatan KIP, *recall* menjadi metrik kritis karena kesalahan *false negative* (2 mahasiswa layak tidak terdeteksi) berimplikasi pada calon penerima manfaat yang berhak justru tidak direkomendasikan. Penelitian tentang prediksi penerimaan beasiswa KIP di Universitas Adzkia melaporkan nilai *recall* yang lebih rendah (5,13%) untuk kelas minoritas, yang menunjukkan bahwa model dalam penelitian ini memiliki kinerja lebih baik dalam mendeteksi mahasiswa layak [14].

Spesifisitas yang tinggi (88,24%) menunjukkan bahwa model sangat baik dalam mengidentifikasi mahasiswa yang tidak layak, sehingga meminimalkan kesalahan pemberian rekomendasi kepada yang tidak berhak. Penelitian tentang klasifikasi penerima beasiswa KIP menggunakan algoritma Naïve Bayes juga menekankan pentingnya keseimbangan antara kemampuan mendeteksi kelas layak dan tidak layak [19].

3.4 Model Pembanding KNN (Tanpa Optimasi)

3.4.1 Konfigurasi Model KNN

Sebagai model pembanding, algoritma KNN digunakan dengan parameter default dari library scikit-learn akan ditampilkan pada table 3.6:

Table 6. Parameter Terbaik Model KNN

Parameter	Nilai Parameter
n_neighbors	5
weights	uniform
metric	Minkowski (p = 2, setara Euclidean)

Pemilihan nilai K=5 sebagai default didasarkan pada praktik umum dalam klasifikasi KNN dan penelitian terdahulu tentang seleksi penerima beasiswa yang menunjukkan bahwa K=5 memberikan akurasi optimal [20]. Penelitian implementasi KNN untuk seleksi beasiswa Aceh Carong juga melaporkan bahwa nilai K=5 menghasilkan akurasi tertinggi (86%) dibanding nilai K lainnya [21].

3.4.2 Evaluasi Model KNN pada Data Uji

Pengujian model KNN pada data uji (54 sampel) menghasilkan akurasi sebesar **83,33%** (45 dari 54 sampel diprediksi benar).

Table 7. Menyajikan Confusion Matrix model KNN

Kategori Kelayakan	Prediksi Tidak Layak	Prediksi Layak
Aktual Tidak Layak	28 (TN)	6 (FP)
Aktual Layak	3 (FN)	17 (TP)

Metrik evaluasi untuk model KNN:

- **Presisi (Layak)** = 17/(17+6) = 73,91%
- **Recall (Layak)** = 17/(17+3) = 85,00%
- **F1-Score (Layak)** = 2×(0,7391×0,8500)/(0,7391+0,8500) = 79,07%
- **Spesifisitas** = 28/(28+6) = 82,35%

Nilai *recall* untuk kelas Layak pada model KNN (85,00%) lebih rendah dibanding SVM (90,00%), yang berarti SVM lebih unggul dalam mengidentifikasi mahasiswa layak. KNN juga menghasilkan lebih banyak *false positive* (6 berbanding 4 pada SVM) dan *false negative* (3 berbanding 2 pada SVM).Penelitian implementasi KNN untuk penerimaan beasiswa KIP di Politeknik Siber Cerdika Internasional melaporkan akurasi tertinggi 76,67% dengan nilai K=1 dan K=2 [15]. Hasil penelitian ini (83,33%) lebih tinggi, kemungkinan disebabkan oleh perbedaan karakteristik dataset dan kriteria kelayakan yang digunakan.

3.5 Studi Komparatif Antar Model

3.5.1 Perbandingan Kinerja Berdasarkan Metrik Evaluasi

Pada tabel 3.8 akan disajikan perbandingan lengkap kinerja antar kedua model, yaitu

Table 8. Perbandingan Kinerja Model SVM dan KNN

Metrik	SVM (Optimasi)	KNN (Default)
Akurasi Cross-validation Rata-rata	87,44%	82,79%
Akurasi Data Uji	88,89%	83,33%
Presisi (Kelas Layak)	81,82%	73,91%
Recall (Kelas Layak)	90,00%	85,00%
F1-Score (Kelas Layak)	85,71%	79,07%
Spesifisitas	88,24%	82,35%
False Positive	4	6
False Negative	2	3

Berdasarkan Tabel 3.8, SVM unggul dalam seluruh metrik evaluasi. Akurasi data uji SVM (88,89%) lebih tinggi 5,56% dibanding KNN (83,33%).Keunggulan SVM dapat dijelaskan melalui karakteristik algoritma yang berfungsi untuk mengidentifikasi *hyperplane* terbaik yang membedakan berbagai kelas dengan jarak yang paling besar. Pendekatan ini membuat SVM lebih robust terhadap data baru dan memiliki kemampuan generalisasi lebih baik, terutama ketika data memiliki struktur yang kompleks [17]. Kernel RBF yang terpilih dalam optimasi memungkinkan SVM menangkap pola non-linear dalam data akademik dengan lebih efektif. Sementara itu, KNN yang berbasis jarak sangat bergantung pada distribusi data lokal dan pemilihan nilai K. Dengan parameter default K=5, KNN menunjukkan kinerja yang cukup baik namun kurang optimal dibanding SVM yang telah dioptimasi. Penelitian tentang klasifikasi penerimaan beasiswa Aceh Carong juga menganalisis bahwa nilai K yang optimal (K=5) tetap menghasilkan keakuratan yang lebih rendah dibanding SVM pada beberapa skenario pengujian [21].

3.5.2 Analisis False Negative sebagai Metrik Kritis

Dalam konteks penempatan mahasiswa KIP, *false negative* (FN) menjadi metrik paling kritis karena berkaitan langsung dengan keadilan sosial. Mahasiswa yang sebenarnya layak namun tidak terdeteksi oleh model (FN) berisiko kehilangan kesempatan memperoleh bantuan pendidikan.Model SVM menghasilkan FN sebanyak 2 mahasiswa layak tidak terdeteksi dari total 20 mahasiswa layak aktual pada data uji, sehingga *false negative rate* SVM adalah 10,00%. Sementara itu, model KNN menghasilkan FN sebanyak 3 mahasiswa (15,00%). SVM terbukti lebih unggul dalam meminimalkan kesalahan kritis ini. Penelitian tentang prediksi penerimaan beasiswa KIP di Universitas Adzkia melaporkan permasalahan serupa, di mana *recall* untuk kelas minoritas hanya mencapai 5,13% akibat ketidakseimbangan kelas yang ekstrem [14]. Penelitian tersebut menyarankan penggunaan teknik penyeimbangan data seperti SMOTE (*Synthetic Minority Oversampling*
Technique) untuk meningkatkan sensitivitas model terhadap kelas minoritas.

3.5.3 Analisis Stabilitas Model Melalui Cross-Validation

Evaluasi stabilitas model dilakukan dengan menganalisis skor *5-fold cross-validation* terdapat pada tabel 3.9, yaitu:

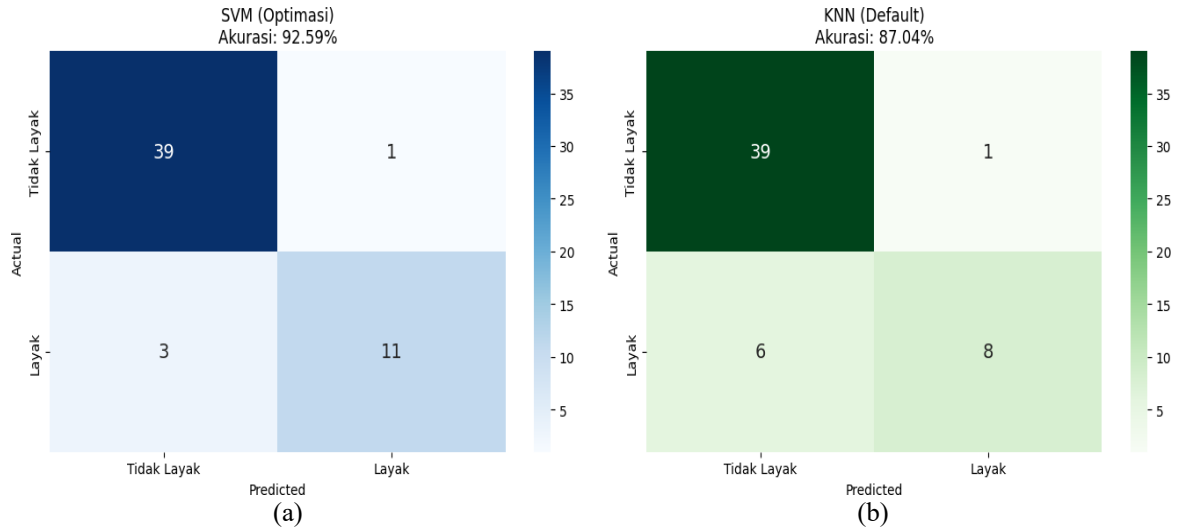
Table 9. Perbandingan Hasil Cross Validation

Model	Fold 1 (%)	Fold 2 (%)	Fold 3 (%)	Fold 4 (%)	Fold 5 (%)	Standar Deviasi (%)
SVM	87,44	87,44	87,44	87,44	87,44	0,00
KNN	82,33	80,47	84,19	82,79	84,19	1,52

Standar deviasi yang lebih kecil pada SVM (0,00%) menunjukkan bahwa model SVM memiliki stabilitas sangat baik dan konsisten di berbagai *fold* data. Sebaliknya, KNN menunjukkan variabilitas lebih tinggi, dengan akurasi bervariasi antara 80,47% hingga 84,19%. Hal ini mengindikasikan bahwa KNN lebih sensitif terhadap komposisi data latih. Penelitian optimasi SVM dengan grid search juga melaporkan bahwa model SVM yang dihasilkan cukup konsisten dan mencegah *overfitting* karena proses validasi silang yang terintegrasi dalam pencarian hyperparameter [19].

3.5.4 Confusion Matrix

Selanjutnya, untuk melakukan evaluasi lebih mendalam terhadap performa dari masing-masing algoritma, kita dapat mengamati hasil Confusion Matrix dari kedua algoritma yang disajikan pada gambar 2 berikut :



Gambar 2. Confusion Matrix (a) Model SVM dan (b) Model KNN

3.5.5 Keterbatasan Penelitian

Keterbatasan yang perlu diakui oleh peniliti:

1. **Ketidakseimbangan Kelas**: Proporsi kelas Layak (36,43%) yang lebih kecil dibanding Tidak Layak (63,57%) berpotensi menyebabkan model bias ke kelas mayoritas, meskipun SVM telah menunjukkan kinerja baik dengan *recall* 90,00%. Teknik *oversampling* seperti SMOTE dapat diterapkan untuk penelitian lanjutan.

2. **Kriteria Berdasarkan Penghasilan Saja**: Penggunaan hanya kriteria penghasilan (< Rp3.000.000) untuk menentukan label "Layak" mungkin belum mempertimbangkan faktor lain seperti prestasi akademik atau kondisi khusus lainnya. Penelitian tentang klasifikasi penerima beasiswa KIP menekankan pentingnya validasi kriteria dengan pihak terkait [19].

3. **Fitur Terbatas**: Hanya 8 fitur yang digunakan, padahal faktor-faktor seperti jumlah tanggungan orang tua, kondisi tempat tinggal, atau status DTKS juga berpotensi memengaruhi kelayakan. Penelitian tentang seleksi beasiswa KIP Kuliah menggunakan enam atribut utama termasuk Status DTKS dan Status P3KE yang tidak tersedia dalam dataset ini [14].

3.5.6 Perbandingan dengan Peneliti Terdahulu

Hasil analisis data ini sejalan dengan beberapa penelitian sebelumnya yang mengaplikasikan SVM dan KNN untuk prediksi akademik dan seleksi penerima bantuan: Penelitian pada klasifikasi kualitas air menunjukkan bahwa model SVM dengan kernel RBF yang dioptimasi mampu menghasilkan kinerja klasifikasi yang lebih baik dibandingkan metode SVM yang tidak dioptimasi [16]. Hal ini menguatkan temuan penelitian ini bahwa kernel RBF merupakan pilihan robust untuk berbagai domain masalah klasifikasi. Penelitian tentang implementasi metode K-NN untuk penerimaan beasiswa KIP di Politeknik Siber Cerdika Internasional menghasilkan akurasi tertinggi 76,67% dengan nilai K=1 dan K=2 menggunakan dataset 80 mahasiswa [15]. Hasil penelitian ini (83,33% untuk KNN dengan 269 data) lebih tinggi, menunjukkan bahwa jumlah sampel yang lebih besar berkontribusi pada peningkatan akurasi model.

Penelitian mengenai pengelompokan mahasiswa yang akan menerima beasiswa KIP dengan memanfaatkan algoritma Naïve Bayes di Universitas Tomakaka Mamuju memperoleh tingkat akurasi sebesar 95,35% dengan 171 data mahasiswa [19]. Akurasi yang lebih tinggi ini mengindikasikan bahwa penentuan algoritma sangat kontekstual dan dipengaruhi oleh atribut data. Penelitian mengenai pemilihan fitur dan perbandingan metode klasifikasi untuk memprediksi kelulusan mahasiswa menggunakan dataset 210 mahasiswa menunjukkan bahwa SVM dengan optimasi hyperparameter menghasilkan model yang konsisten dan mencegah overfitting [18]. Temuan ini konsisten dengan stabilitas model SVM yang ditunjukkan dalam

penelitian ini. Penelitian tentang implementasi algoritma KNN dalam klasifikasi penerimaan beasiswa Aceh Carong menggunakan dataset 249 mahasiswa melaporkan bahwa nilai K=5 menghasilkan akurasi tertinggi 86% [21]. Hasil ini sejalan dengan penelitian ini yang menggunakan K=5 dan menghasilkan akurasi 83,33%.

3.5.7 Rekomendasi Penempatan Mahasiswa KIP

Berdasarkan hasil analisis komparatif terhadap 269 data mahasiswa, model SVM dengan konfigurasi kernel RBF, C=10, dan gamma=0,1 direkomendasikan untuk diimplementasikan dalam sistem pendukung keputusan penempatan mahasiswa KIP di Fakultas Science and Technology UNPRI. Rekomendasi ini didasarkan pada keunggulan SVM dalam:

1. Akurasi tertinggi (88,89%)
2. Recall tertinggi untuk kelas Layak (90,00%)
3. Stabilitas model yang sangat baik (standar deviasi CV=0%)
4. Tingkat false positive (4) dan false negative (2) terendah

Dengan menggunakan model ini, dari 269 mahasiswa yang diproses, hanya sekitar 2 mahasiswa layak yang berisiko tidak terdeteksi (false negative rate 10%), yang masih dalam batas toleransi untuk sistem pendukung keputusan.

Namun demikian, implementasi model ini sebaiknya tidak menggantikan keputusan manusia sepenuhnya, melainkan berperan sebagai:

- **Penyaring awal** (*pre-screening*) untuk mengidentifikasi kandidat potensial.
- **Validasi silang** untuk memverifikasi keputusan manual.
- **Peringatan dini** untuk kasus-kasus *borderline* yang memerlukan investigasi lebih lanjut.

Sebagai langkah pengembangan ke depan, disarankan untuk:

1. Menggunakan teknik penyeimbangan data seperti SMOTE untuk meningkatkan *recall* kelas Layak.
2. Menambah fitur relevan seperti Status DTKS, Status P3KE, atau jumlah tanggungan orang tua [14].
3. Memperbesar ukuran dataset dengan menambah data dari tahun ajaran berbeda
4. Menguji algoritma lain seperti Random Forest atau XGBoost untuk perbandingan lebih komprehensif
5. Mengintegrasikan model dengan *dashboard* interaktif yang memungkinkan tim seleksi melakukan eksplorasi data dan melihat justifikasi di balik setiap rekomendasi.

4. KESIMPULAN

Dalam penelitian, prediksi penempatan mahasiswa penerima KIP pada Fakultas Science and Technology UNPRI menggunakan metode SVM dengan optimasi hyperparameter melalui metode grid search dan studi komparatif KNN terbukti mampu meningkatkan kinerja model Support Vector Machine (SVM). Hal ini ditunjukkan oleh hasil cross validation yang konsisten, dengan rata-rata akurasi mencapai 87,44% dan deviasi standar sebesar 0,00%, yang mencerminkan performa model yang sangat baik. Pada tahap pengujian, model SVM mencapai akurasi sebesar 88,89% dengan nilai recall untuk kategori layak sebesar 90,00%, yang menunjukkan bahwa model mampu mendeteksi sebagian besar mahasiswa KIP yang sesuai dalam penempatannya.Berdasarkan studi komparatif, model SVM juga menunjukkan kinerja yang lebih unggul dibandingkan dengan model K-Nearest Neighbor (KNN) pada seluruh indikator evaluasi, seperti presisi, akurasi, F1-score, recall, dan spesifisitas. Model KNN hanya memperoleh akurasi sebesar 83,33% serta memiliki tingkat kesalahan yang lebih tinggi. Secara keseluruhan, penerapan algoritma SVM mampu menghasilkan model klasifikasi dengan kinerja yang baik dalam memprediksi penempatan mahasiswa KIP. Dengan menggunakan kernel Radial Basis Function (RBF) serta parameter terbaik (C = 10 dan gamma = 0,1), model ini mampu menangkap pola data baik yang bersifat linear maupun non-linear.

UCAPAN TERIMAKASIH

Penulis menyampaikan ucapan terima kasih kepada semua pihak yang telah memberikan dukungan dalam pelaksanaan penelitian ini, baik berupa bantuan teknis maupun non-teknis, sehingga penelitian ini dapat diselesaikan dengan baik.

REFERENSI

[1] A. D. Larasati, D. Dinda, N. A. Aidah, R. Gustiputri, and S. N. R. Isyak, "Analisis Kebijakan Program Beasiswa Kartu Indonesia Pintar-Kuliah (KIP-K) di Universitas Diponegoro," *Jurnal Ilmu Administrasi dan Studi Kebijakan (JIASK)*, vol. 5, no. 1, pp. 1–22, Sep. 2022.

[2] Dewi Ulfah Ningsih and Hasbiah, "Pengaruh Beasiswa Kartu Indonesia Pintar. (KIP) Terhadap Kesenjangan Sosial Melalui Aksesibilitas Pendidikan," *At-Thariqah J. Ekon.*, vol. 4, no. 2, pp. 89–99, 2024, doi: 10.47945/at-thariqah.v4i2.1626.

[3] Irene Dwi Ardianty, Syafiq Febriyanto, R.Aj Cahya Dira Aulia Putri, Ardiya Pramesti Regita Cahyani. and Mohamad Djasuli, "Penyalahgunaan Dana Kartu Indonesia Pintar Kuliah Dan Dampaknya Terhadap Pendidikan Di Indonesia," *Coll. Stud. J.*, vol. 7, no. 1, pp. 28–36, 2024, doi: 10.56301/csj.v7i1.1193.

[4] Menteri Pendidikan dan Kebudayaan Republik Indonesia, "Permendikbud Nomor 8 Tahun 2020," *Kementeri. Pendidik. dan Kebud. Republik Indones.*, vol. 58, no. 12, pp. 7250–7257, 2020.

[5] Puslapdik, "Kuliah Bukan Sebuah Mimpi Program Indonesia Pintar (Pip) Pendidikan Tinggi Kartu Indonesia Pintar Kuliah (Kip Kuliah)," p. 4, 2021.

[6] F. Andrianto, A. Fadlil, and I. Riadi, "Linear Kernel Optimization of Support Vector Machine Algorithm on Online Marketplace Sentiment Analysis," *Komputasi J. Ilm. Ilmu Komput. dan Mat.*, vol. 21, no. 1, pp. 68–82, 2024, doi: 10.33751/komputasi.v21i1.9266.

[7] B. I. Nugroho, N. A. Santoso, and A. A. Murtopo, "Prediksi Kemampuan Akademik Mahasiswa dengan Metode Support Vector Machine," *Remik*, vol. 7, no. 1, pp. 177– 188, 2023, doi: 10.33395/remik.v7i1.12010.

[8] Sumiyatun, Y. Cahyadi, and E. Faizal, "Implementation of Support Vector Machine Algorithm for Classification of Study Period and Graduation Predicate of Students," *Indones. J. Data Sci.*, vol. 6, no. 1, pp. 56–64, 2025, doi: 10.56705/ijodas.v6i1.214.

[9] M. Fajri and A. Primajaya, "Komparasi Teknik Hyperparameter Optimization pada SVM untuk Permasalahan Klasifikasi dengan Menggunakan Grid Search dan Random Search," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 14–19, 2023, doi: 10.30871/jaic.v7i1.5004.

[10] M. Hariyanto, M. Kholiq, A. Yani, and Narti, "Inti nusa mandiri," *Inti Nusa Mandiri*, vol. 14, no. 2, pp. 133–138, 2020.

[11] S. Zanariyah, "Teknik Observasi Yang Efektif Dan Efisien Pada KegiatanKuliah Kerja Nyata (KKN)," *J. Pengabdi. Multidisiplin*, vol. 4, p. 2024, 2024.

[12] Ardiansyah, Risnita, and M. S. Jailani, "IHSAN: Jurnal Pendidikan Islamhttp://ejournal.yayasanpendidikandzurriyatulquran.id/index.php/ihsanVolume 1 Nomor 2Juli20231Ardiansyah, Risnita, M.Syahran JailaniTeknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif ," *J. IHSAN J. Pendidik. Islam*, vol. 1, no. 2, pp. 1–9, 2023.

[13] N. Nursalam and A. S. A. D. Djaha, "Pelatihan Pembuatan Kuesioner Penelitian Bagi Mahasiswa Prodi Administrasi Negara Fisip Universitas Nusa Cendana," *Jdistira*, vol. 3, no. 1, pp. 25–31, 2023, doi: 10.58794/jdt.v3i1.433.

[14] Andre Febrianto, Rusdy A Siroj, and Hartatiana, "Studi Literatur: Landasan Dalam Memilih Metode Penelitian Yang Tepat," *J. Educ. Res. Dev. | E-ISSN 3063-9158*, vol. 1, no. 2, pp. 259–263, 2024, doi: 10.62379/jerd.v1i2.142.

[15] M. A. Thoriq, M., Maulana, F., Eirlangga, Y. S., Hayati, N., & Madani, "Implementasi Algoritma Naïve Bayes dalam Prediksi Penerimaan Mahasiswa Penerima Beasiswa KIPdi Universitas Adzkia," *FASILKOM*, vol. 15, no. 1, pp. 108–114, 2025, [Online]. Available: https://ejurnal.umri.ac.id/index.php/JIK/article/view/8959/3506

[16] A. Wijaya, L. Magdalena, and C. Nas, "Implementation of the Smart Indonesia Card Scholarship (KIP) Acceptance Using the K-NN Method (Case Study: Politeknik Siber Cerdika Internasional)," *J. Indones. Sos. Teknol.*, vol. 5, no. 9, pp. 3222–3234, 2024, doi: 10.59141/jist.v5i9.1382.

[17] T. Seviya and L. Hakim, "Information Technology Water Quality Classification Using SVM with PSO Based Parameter Optimization," vol. 4, no. 3, pp. 389–401, 2025.

[18] S. P. Haryatmi, E., & Hervianti, "Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu," *J. RESTI*, vol. 5, no. 2, pp. 386– 392, 2021.

[19] J. Zeniarja, A. Salam, and F A. Ma'ruf, "Seleksi Fitur dan Perbandingan Algoritma Klasifikasi untuk Prediksi Kelulusan Mahasiswa," *J. Rekayasa Elektro.*, vol. 18, no. 2, pp. 102–108, 2022, doi: 10.17529/jre.v18i2.24047.

[20] Hidayat, M. KH, I. Kusmanto, M. Kadir, and Kristian, "Identifikasi Mahasiswa Calon Penerima Beasiswa KIP Menggunakan Algoritma Naive Bayes di Universitas Tomakaka Mamuju," *J. FASILKOM*, vol. 15, no. 3, pp. 456–464, 2025.

[21] M. Kholil, Kusrini, and Henderi, "Penerapan Metode K Nearest Neighbord Dalam Proses Seleksi Penerima Beasiswa," *Proceeding Semin. Nas. Sist. Inf. dan Teknol. Inf.*, vol. 1, no. 1, pp. 13–18. 2018,[Online].Available: http://www.sisfotenika.stmikpontianak.ac.id/index.php/sensitek/article/view/235

[22] R. Yanti, S. Retno, and B. Yafis, "Implementation of KNN Algorithm to classify the Scholarship Recipients of Aceh Carong at Universitas Malikussaleh." *J. Adv. Comput. Knowl. Algorithms*, vol. 1, no. 1, pp. 5–9, 2024, doi: 10.29103/jacka.v1i1.1453