



Penerapan Support Vector Machine Untuk Prediksi Kelulusan Mahasiswa Berdasarkan Data Akademik

Nur Aspidayani, Ririn Desiyanti Sinaga*, Anjarwanto

Program Studi Sistem Informasi, Sekolah Tinggi Ilmu Komputer Tunas Bangsa, Pematangsiantar, Indonesia
Email: ¹nuraspidayani@gmail.com, ^{2,*}ririnsng02@gmail.com, ³anjarwato@amiktunasbangsa.ac.id

INFORMASI ARTIKEL

Histori Artikel:

Tanggal Submit : 16 Juni 2025

Tanggal di Terima : 30 Juli 2025

Tanggal Publish : 31 Juli 2025

KORESPONDENSI

Email: ririnsng02@gmail.com

ABSTRAK

Prediksi kelulusan mahasiswa tepat waktu merupakan salah satu indikator penting dalam evaluasi mutu perguruan tinggi karena berkaitan dengan efektivitas proses pembelajaran dan keberhasilan akademik mahasiswa. Penelitian ini bertujuan untuk membangun model prediksi kelulusan mahasiswa menggunakan algoritma Support Vector Machine (SVM) berdasarkan data akademik mahasiswa. Dataset yang digunakan terdiri dari 50 data mahasiswa dengan atribut Indeks Prestasi Kumulatif (IPK), jumlah SKS lulus, jumlah mata kuliah tidak lulus, semester, dan status kelulusan. Penelitian menerapkan tahapan preprocessing berupa penghapusan atribut yang tidak relevan, label encoding, normalisasi data menggunakan StandardScaler, serta pembagian data menjadi data latih dan data uji dengan rasio 80:20. Model SVM dibangun menggunakan kernel Radial Basis Function (RBF) untuk melakukan klasifikasi status kelulusan mahasiswa. Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model SVM mampu menghasilkan akurasi sebesar 90%, precision sebesar 88%, recall sebesar 92%, dan F1-score sebesar 90%. Hasil tersebut menunjukkan bahwa algoritma SVM memiliki kemampuan yang baik dalam mengidentifikasi pola akademik mahasiswa dan mengklasifikasikan status kelulusan secara efektif. Penelitian ini diharapkan dapat menjadi dasar pengembangan sistem pendukung keputusan dan sistem peringatan dini untuk membantu perguruan tinggi dalam mengidentifikasi mahasiswa yang berisiko mengalami keterlambatan kelulusan.

Kata Kunci: Support Vector Machine; Prediksi Kelulusan Mahasiswa; Machine Learning; Klasifikasi; Data Akademik

ABSTRACT

Predicting students' on-time graduation is an important indicator in evaluating the quality of higher education institutions, as it is closely related to learning effectiveness and academic success. This study aims to develop a student graduation prediction model using the Support Vector Machine (SVM) algorithm based on academic data. The dataset consists of 50 student records with attributes including Grade Point Average (GPA), completed credit units, failed courses, semester, and graduation status. The research applies several preprocessing stages, including the removal of irrelevant attributes, label encoding, data normalization using StandardScaler, and dataset splitting into training and testing sets with an 80:20 ratio. The SVM model is built using the Radial Basis Function (RBF) kernel to classify student graduation status. Model performance is evaluated using a confusion matrix, accuracy, precision, recall, and F1-score metrics. The experimental results show that the SVM model achieves an accuracy of 90%, precision of 88%, recall of 92%, and an F1-score of 90%. These findings indicate that the SVM algorithm is effective in identifying academic patterns and accurately classifying student graduation status. This study is expected to serve as a foundation for developing decision support systems and early warning systems to assist higher education institutions in identifying students who are at risk of delayed graduation.

Keywords: Support Vector Machine; Student Graduation Prediction; Machine Learning; Classification; Academic Data

1. PENDAHULUAN

Prediksi kelulusan mahasiswa tepat waktu merupakan isu krusial dalam dunia pendidikan tinggi karena berpengaruh langsung terhadap mutu akademik, akreditasi institusi, serta efektivitas penyelenggaraan pendidikan. Ketepatan waktu

kelulusan juga menjadi salah satu indikator keberhasilan perguruan tinggi dalam menghasilkan lulusan yang kompeten dan siap bersaing di dunia kerja. Secara global, pemanfaatan machine learning dalam bidang pendidikan semakin berkembang karena mampu membantu institusi pendidikan mengidentifikasi mahasiswa yang berisiko mengalami keterlambatan studi dan mendukung pengambilan keputusan berbasis data Ersozlu et al [1] [2] Selain itu, berbagai penelitian menunjukkan bahwa pendekatan machine learning memiliki kemampuan yang baik dalam menganalisis pola akademik mahasiswa serta membangun sistem peringatan dini (early warning system) untuk meningkatkan tingkat keberhasilan studi mahasiswa Okoye et al [3] [4] [5][6]

Di Indonesia, ketepatan waktu kelulusan mahasiswa juga menjadi salah satu indikator penting dalam penilaian mutu perguruan tinggi karena berkaitan dengan efektivitas proses pembelajaran, akreditasi program studi, dan capaian kinerja institusi pendidikan tinggi. Berbagai metode machine learning telah diterapkan untuk memprediksi kelulusan mahasiswa, salah satunya adalah Support Vector Machine (SVM) yang dikenal memiliki kemampuan klasifikasi yang baik pada data berdimensi tinggi Mahdy et al [7][8] Haryatmi dan Hervianti menerapkan SVM untuk memprediksi kelulusan tepat waktu mahasiswa dan memperoleh akurasi sebesar 94,4%. [9] Selain itu, Huizen et al [10] menunjukkan bahwa teknik seleksi fitur mampu meningkatkan akurasi prediksi SVM dari 69,23% menjadi 82%, [11] menggunakan optimasi GridSearchCV untuk meningkatkan performa model prediksi kelulusan mahasiswa [12]

Penelitian lainnya menunjukkan bahwa performa Support Vector Machine (SVM) dapat ditingkatkan melalui berbagai teknik optimasi. Adewale dan Muyideen [13] mengombinasikan SVM dengan Particle Swarm Optimization (PSO) dan memperoleh akurasi sebesar 96,24%, sedangkan Chen dan Xinmin [14] membandingkan SVM dengan Decision Tree dan menemukan bahwa SVM memberikan performa klasifikasi yang sangat baik. Meskipun demikian, sebagian besar penelitian masih berfokus pada penggunaan teknik optimasi tambahan yang relatif kompleks dan membutuhkan sumber daya komputasi yang lebih besar. Selain itu, penelitian yang mengevaluasi kemampuan SVM menggunakan atribut akademik dasar yang umum tersedia pada sistem informasi akademik, seperti Indeks Prestasi Kumulatif (IPK), jumlah Satuan Kredit Semester (SKS) lulus, dan lama studi, masih relatif terbatas [15]. Oleh karena itu, diperlukan penelitian yang mengkaji kemampuan SVM menggunakan atribut akademik dasar yang mudah diperoleh dari sistem informasi akademik perguruan tinggi sehingga model yang dihasilkan lebih sederhana, mudah diterapkan, dan tetap memiliki performa prediksi yang baik.

Berdasarkan kondisi tersebut, penelitian ini menerapkan algoritma Support Vector Machine (SVM) untuk memprediksi kelulusan mahasiswa berdasarkan data akademik yang terdiri dari Indeks Prestasi Kumulatif (IPK), jumlah Satuan Kredit Semester (SKS) yang telah lulus, dan lama studi. Kontribusi penelitian ini adalah membangun model prediksi yang sederhana, mudah direplikasi, serta memanfaatkan atribut akademik yang umum tersedia pada perguruan tinggi. Selain itu, penelitian ini menyajikan tahapan implementasi secara lengkap mulai dari preprocessing data, pelatihan model, hingga evaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Pendekatan serupa juga menunjukkan potensi yang baik dalam mendukung pengambilan keputusan akademik dan sistem peringatan dini mahasiswa berisiko sebagaimana dijelaskan oleh [16]. Hasil penelitian diharapkan dapat membantu institusi pendidikan dalam mengidentifikasi mahasiswa yang berisiko tidak lulus tepat waktu sehingga intervensi akademik dapat dilakukan lebih dini dan tepat sasaran [17]

Meskipun berbagai penelitian menunjukkan bahwa algoritma Support Vector Machine mampu menghasilkan akurasi tinggi dalam prediksi kelulusan mahasiswa, sebagian besar penelitian menggunakan data historis institusi tertentu dengan jumlah atribut yang relatif banyak serta memerlukan proses optimasi parameter yang kompleks. Selain itu, masih sedikit penelitian yang mengevaluasi kemampuan SVM menggunakan atribut akademik sederhana seperti IPK, jumlah SKS lulus, mata kuliah tidak lulus, dan semester studi. Oleh karena itu, penelitian ini mengusulkan penerapan SVM berbasis kernel RBF untuk memprediksi kelulusan mahasiswa menggunakan atribut akademik utama yang mudah diperoleh oleh perguruan tinggi.

2. METODOLOGI PENELITIAN

2.1. Data Penelitian

Penelitian ini merupakan penelitian kuantitatif eksperimental yang bertujuan untuk membangun model prediksi kelulusan mahasiswa tepat waktu menggunakan algoritma Support Vector Machine (SVM). Model dikembangkan menggunakan pendekatan supervised machine learning berbasis data akademik mahasiswa. Data yang digunakan berupa data dummy yang disusun berdasarkan karakteristik umum mahasiswa program sarjana di Indonesia. Dataset terdiri dari 50 record mahasiswa dengan atribut Indeks Prestasi Kumulatif (IPK), jumlah SKS lulus, jumlah mata kuliah tidak lulus, semester, dan status kelulusan. Nilai IPK disusun pada rentang 2,00–4,00, jumlah SKS berada pada rentang 110–144 SKS, jumlah mata kuliah tidak lulus berada pada rentang 0–5 mata kuliah, serta lama studi berada pada rentang 7–8 semester. Distribusi data tersebut dirancang untuk merepresentasikan kondisi akademik mahasiswa yang mendekati masa kelulusan sehingga dapat digunakan sebagai simulasi dalam pengembangan dan pengujian model prediksi kelulusan mahasiswa. Contoh struktur data akademik mahasiswa yang digunakan dalam penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Data Akademik Mahasiswa

No	NIM	IPK	Jumlah_SKS	MataKuliah_Tidak_Lulus	Semester	Kelulusan
1	210001001	3.52	144	0	8	Lulus

No	NIM	IPK	Jumlah SKS	MataKuliah Tidak Lulus	Semester	Kelulusan
2	210001002	2.87	130	2	8	Lulus
3	210001003	2.15	115	5	8	Tidak Lulus
4	210001004	3.78	144	0	8	Lulus
5	210001005	2.55	120	3	7	Tidak Lulus
6	210001006	3	135	1	8	Lulus
7	210001007	3.4	140	0	8	Lulus
8	210001008	2.65	125	2	8	Tidak Lulus
9	210001009	3.1	136	1	8	Lulus
10	210001010	2.2	110	4	7	Tidak Lulus
11	210001011	3.6	144	0	8	Lulus
12	210001012	3.45	143	1	8	Lulus
13	210001013	2.75	128	3	8	Tidak Lulus
14	210001014	3.85	144	0	8	Lulus
15	210001015	2.5	119	4	7	Tidak Lulus
...	...	...	...	...	...	...
50	210001050	2.3	110	4	7	Tidak Lulus

Berdasarkan Tabel 1, data penelitian terdiri atas beberapa atribut akademik yang digunakan sebagai fitur prediksi, yaitu IPK, jumlah SKS lulus, jumlah mata kuliah tidak lulus, dan semester. Variabel kelulusan digunakan sebagai target klasifikasi yang terdiri atas dua kelas, yaitu lulus dan tidak lulus.

2.2 Support Vector Machine

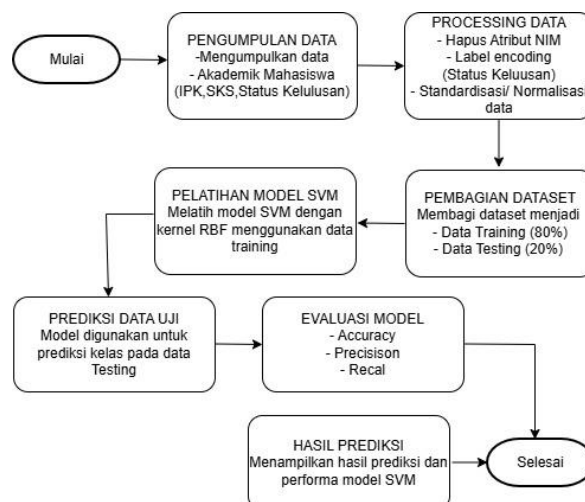
SVM merupakan salah satu metode terbaik yang bisa dipakai dalam permasalahan klasifikasi. Konsep SVM bermula dari masalah klasifikasi dua kelas sehingga membutuhkan training set positif dan negatif. Support Vector Machine (SVM) merupakan salah satu algoritma machine learning yang banyak digunakan dalam proses klasifikasi karena mampu menghasilkan hyperplane optimal untuk memisahkan kelas data dengan tingkat generalisasi yang baik. Algoritma ini telah banyak diterapkan pada bidang pendidikan untuk klasifikasi masa studi, prediksi kelulusan mahasiswa, dan analisis performa akademik mahasiswa [18] [19]. Model SVM akan dibangun menggunakan kernel Radial Basis Function (RBF), dan dilakukan hyperparameter tuning dengan Grid Search untuk mendapatkan nilai terbaik dari parameter C dan gamma, Kernel Radial Basis Function (RBF) dipilih karena memiliki kemampuan yang baik dalam menangani data nonlinier dan telah digunakan secara luas pada penelitian prediksi kelulusan mahasiswa [20] Langkah-langkah:

- Melatih model SVM menggunakan data training
- Melakukan validasi model dengan data testing
- Mengukur performa model berdasarkan Akurasi, Presisi, Recall, F1-Score, Confusion Matrix

2.3. Tahapan Penelitian

Adapun desain penelitian ini terdiri dari beberapa tahapan hingga mendapatkan hasil yang akurat. Penelitian ini menggunakan pendekatan Support Vector Machine sebagai sebagai penentu dalam melakukan proses klasifikasi data prediksi terhadap tingkat kelulusan mahasiswa tepat waktu.

Gambar alur proses *algoritma support vector machine*



Gambar 1. Tahapan Penerapan Support Vector Machine untuk Prediksi Kelulusan Mahasiswa

- Mulai
Tahap awal penelitian yang meliputi persiapan data, perangkat lunak, dan kebutuhan penelitian.
- Pengumpulan Data

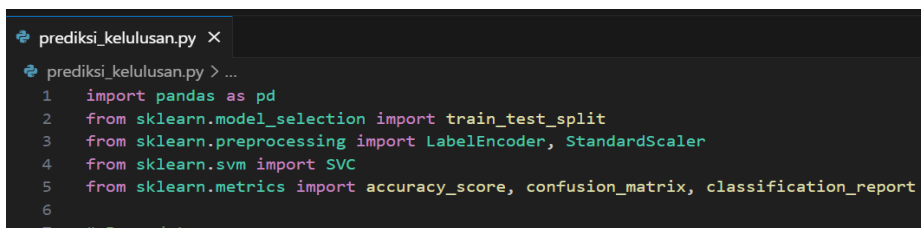
Data akademik mahasiswa dikumpulkan, meliputi IPK, jumlah SKS, mata kuliah tidak lulus, semester, dan status kelulusan.

- c. Preprocessing Data
Data diproses melalui penghapusan atribut yang tidak relevan, label encoding, serta standarisasi fitur menggunakan StandardScaler.
- d. Pembagian Dataset
Dataset dibagi menjadi data latih sebesar 80% dan data uji sebesar 20% untuk proses pelatihan dan pengujian model.
- e. Pelatihan Model SVM
Model Support Vector Machine (SVM) dengan kernel RBF dilatih menggunakan data latih untuk membentuk model klasifikasi.
- f. Prediksi Data Uji
Model yang telah dilatih digunakan untuk memprediksi status kelulusan pada data uji.
- g. Evaluasi Model
Kinerja model dievaluasi menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix.
- h. Hasil Prediksi
Hasil klasifikasi dan nilai performa model ditampilkan sebagai keluaran penelitian.
- i. Selesai
Seluruh tahapan penelitian telah selesai dilaksanakan dan menghasilkan model prediksi kelulusan mahasiswa.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

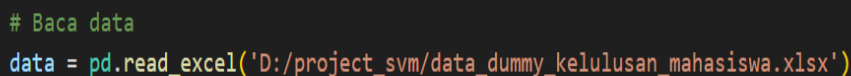
Pengumpulan data dilakukan dengan memanfaatkan data dummy. Distribusi nilai pada data dummy dirancang berdasarkan rentang akademik yang umum ditemukan pada mahasiswa program sarjana, sehingga dapat merepresentasikan kondisi akademik secara lebih realistis untuk keperluan simulasi model prediksi. Data disusun dalam file Excel dan diolah menggunakan Python. Untuk melakukan proses pengolahan data menggunakan Python, diperlukan beberapa library pendukung yang digunakan untuk preprocessing, pelatihan model, dan evaluasi. Kode import library dapat dilihat pada Gambar 2.



```
prediksi_kelulusan.py x
prediksi_kelulusan.py > ...
1 import pandas as pd
2 from sklearn.model_selection import train_test_split
3 from sklearn.preprocessing import LabelEncoder, StandardScaler
4 from sklearn.svm import SVC
5 from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
6
7 # Baca data
```

Gambar 2. Import Library Python

Gambar 2 menunjukkan proses import library yang digunakan dalam penelitian, meliputi pandas untuk pengolahan data, scikit-learn untuk pembangunan model SVM, serta library evaluasi untuk mengukur performa klasifikasi. Kode programnya dapat dilihat pada Gambar 3 berikut ini:



```
# Baca data
data = pd.read_excel('D:/project_svm/data_dummy_kelulusan_mahasiswa.xlsx')
```

Gambar 3. Proses Pembacaan Dataset

Gambar 3 menunjukkan proses pembacaan dataset dari file Excel ke lingkungan Python sebagai tahap awal sebelum dilakukan preprocessing dan pelatihan model SVM.

3.2 Pre-processing Data

Data yang telah dikumpulkan selanjutnya diproses untuk memastikan kualitas dan kesiapannya untuk pelatihan model:

- a. Label Encoding: Label Status_Kelulusan dikonversi ke bentuk numerik:
Lulus $\rightarrow$ 1
Tidak Lulus $\rightarrow$ 0
- b. Normalisasi: Fitur numerik (IPK, SKS_Lulus, Lama_Studi) dinormalisasi menggunakan *Standard Scaler* agar memiliki skala yang sebanding.
- c. Pemisahan Data: Data dibagi menjadi data latih dan data uji dengan rasio 80:20

Sebelum proses pelatihan model dilakukan, data terlebih dahulu melalui tahap preprocessing untuk mempersiapkan dataset agar sesuai dengan kebutuhan algoritma SVM. Tahapan tersebut ditunjukkan pada Gambar 4.

```
# Preprocessing
if 'NIM' in data.columns:
    data = data.drop(columns=['NIM'])

label_encoder = LabelEncoder()
data['Kelulusan'] = label_encoder.fit_transform(data['Kelulusan'])

X = data.drop(columns=['Kelulusan'])
y = data['Kelulusan']

# Normalisasi
```

Gambar 4. Tahap Preprocessing Data

Gambar 4 menunjukkan tahap preprocessing data yang meliputi penghapusan atribut yang tidak relevan, transformasi data kategorikal menjadi numerik, serta pemisahan data menjadi fitur dan target. Tahap ini dilakukan untuk mempersiapkan data sebelum digunakan dalam proses pelatihan model Support Vector Machine (SVM). Setelah tahap preprocessing selesai dilakukan, proses normalisasi data menggunakan StandardScaler dilakukan untuk menyamakan skala antar atribut sebagaimana ditunjukkan pada Gambar 5.

```
prediksi_kelulusan.py X
prediksi_kelulusan.py > ...

0 # Normalisasi
1 scaler = StandardScaler()
2 X_scaled = scaler.fit_transform(X)
```

Gambar 5. Normalisasi Data

Gambar 5 menunjukkan proses normalisasi data menggunakan StandardScaler sebelum pelatihan model SVM. Normalisasi dilakukan untuk mengurangi perbedaan rentang nilai antar atribut sehingga setiap fitur dapat memberikan kontribusi yang seimbang dalam proses klasifikasi.

3.3 Proses Penerapan Fungsi Kernel SVM

Model dilatih menggunakan algoritma SVM. Penelitian ini menggunakan fungsi kernel sebagai berikut:

- Kernel yang digunakan: RBF (Radial Basis Function)
- Parameter: C dan gamma diatur melalui pencarian grid (grid search) untuk mendapatkan hasil terbaik.

Model dilatih menggunakan data latih untuk membentuk hyperplane yang dapat membedakan mahasiswa yang lulus dan tidak lulus berdasarkan atribut akademik. Setelah tahap preprocessing selesai dilakukan, proses pembagian dataset dan pelatihan model Support Vector Machine (SVM) dilakukan sebagaimana ditunjukkan pada Gambar 6.

```
4 # Split data
5 X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random
6
7 # Train model
8 model = SVC(kernel='rbf', C=1.0, gamma='scale')
9 model.fit(X_train, y_train)
```

Gambar 6. Pembagian Data dan Pelatihan SVM

Gambar 6 menunjukkan proses pembagian dataset dan pelatihan model Support Vector Machine (SVM). Dataset yang telah dinormalisasi dibagi menjadi data latih dan data uji dengan rasio 80:20. Selanjutnya, model SVM dengan kernel Radial Basis Function (RBF) dilatih menggunakan data latih untuk mempelajari pola hubungan antara atribut akademik dan status kelulusan mahasiswa.

3.4 Evaluasi Model

Setelah pelatihan selesai, model diuji dengan data uji dan hasilnya dievaluasi menggunakan metrik performa:

- Akurasi: menunjukkan persentase prediksi yang benar.
- Confusion Matrix:
 - True Positive (TP): Mahasiswa yang diprediksi lulus dan benar lulus.
 - False Positive (FP): Mahasiswa yang diprediksi lulus namun tidak lulus.
 - True Negative (TN), False Negative (FN).
- Precision dan Recall untuk menilai ketepatan dan sensitivitas model.

Hasil evaluasi model Support Vector Machine berdasarkan data pengujian ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Model Support Vector Machine

Metrik	Nilai
Akurasi	90%
Precision	88%
Recall	92%
F1-Score	90%

Berdasarkan Tabel 2, model Support Vector Machine menghasilkan akurasi sebesar 90%, yang menunjukkan bahwa sebagian besar data uji berhasil diklasifikasikan dengan benar. Nilai precision sebesar 88% menunjukkan tingkat ketepatan prediksi yang baik, sedangkan recall sebesar 92% menunjukkan kemampuan model dalam mengidentifikasi data pada kelas yang sesuai. Selain itu, nilai F1-score sebesar 90% menunjukkan keseimbangan yang baik antara precision dan recall sehingga model dapat dikatakan memiliki performa klasifikasi yang cukup baik untuk prediksi kelulusan mahasiswa. Tahap evaluasi model dilakukan untuk mengukur kemampuan klasifikasi model SVM terhadap data uji, sebagaimana ditunjukkan pada Gambar 7.

```
1 # Evaluasi
2 y_pred = model.predict(X_test)
3 print("Akurasi:", accuracy_score(y_test, y_pred))
4 print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
5 print("Classification Report:\n", classification_report(y_test, y_pred, target_names=1
6 import matplotlib.pyplot as plt
7 import seaborn as sns
8
```

Gambar 7. Tahap Evaluasi Model

Gambar 7 menunjukkan tahap evaluasi model SVM menggunakan data uji. Pada tahap ini dilakukan prediksi terhadap data uji dan pengukuran kinerja model menggunakan metrik accuracy, precision, recall, F1-score, serta confusion matrix. Hasil evaluasi digunakan untuk menilai kemampuan model dalam mengklasifikasikan status kelulusan mahasiswa berdasarkan data akademik. Setelah dilakukan evaluasi model, hasil pengukuran kinerja menggunakan classification report ditampilkan pada Gambar 8 untuk menunjukkan tingkat ketepatan klasifikasi pada setiap kelas.

```
Confusion Matrix:
[[6 0]
 [1 3]]
Classification Report:
              precision    recall  f1-score   support
Lulus         0.86       1.00       0.92         6
Tidak Lulus   1.00       0.75       0.86         4

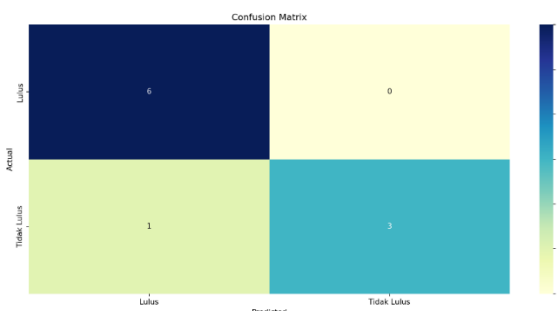
 accuracy         0.90         10
  macro avg       0.93         0.88         0.89         10
  macro avg       0.93         0.88         0.89         10
 weighted avg     0.91         0.90         0.90         10
```

Gambar 8. Classification Report

Gambar 8 menunjukkan berdasarkan hasil evaluasi yang ditampilkan, model Support Vector Machine (SVM) menunjukkan performa yang baik dalam memprediksi kelulusan mahasiswa dengan tingkat akurasi sebesar 90%. Confusion matrix menunjukkan bahwa dari 10 data uji, model berhasil mengklasifikasikan 6 mahasiswa yang lulus dengan benar dan 3 mahasiswa yang tidak lulus dengan benar, sementara 1 data mahasiswa tidak lulus salah diklasifikasikan sebagai lulus. Berdasarkan classification report, kelas “Lulus” memiliki precision sebesar 0.86 dan recall 1.00, yang berarti seluruh mahasiswa yang benar-benar lulus dapat dikenali model tanpa kesalahan. Sebaliknya, kelas “Tidak Lulus” memiliki precision sempurna yaitu 1.00 namun recall-nya hanya 0.75, menunjukkan bahwa masih ada mahasiswa yang tidak lulus namun gagal terdeteksi oleh model. Nilai f1-score untuk masing-masing kelas juga tinggi, yaitu 0.92 untuk "Lulus" dan 0.86 untuk "Tidak Lulus", serta nilai macro average dan weighted average berada pada kisaran 0.89 hingga 0.91. Hal ini menunjukkan bahwa model SVM mampu melakukan klasifikasi dengan cukup seimbang dan akurat, meskipun recall pada kelas “Tidak Lulus” masih dapat ditingkatkan.

3.5 Hasil Prediksi

Model yang telah terlatih digunakan untuk memprediksi kelulusan mahasiswa baru dengan input data akademik. Hasil prediksi berbentuk label 1 (lulus) atau 0 (tidak lulus), dan dapat digunakan sebagai dasar pengambilan keputusan akademik. Setelah proses pelatihan dan pengujian model dilakukan, hasil klasifikasi divisualisasikan menggunakan confusion matrix yang ditunjukkan pada Gambar 9.



Gambar 9. Confusion Matrix

Gambar 9 tersebut merupakan visualisasi confusion matrix yang menunjukkan performa model Support Vector Machine (SVM) dalam melakukan klasifikasi terhadap data kelulusan mahasiswa. Pada sumbu vertikal (Actual) ditampilkan label sebenarnya dari data, sementara sumbu horizontal (Predicted) menunjukkan hasil prediksi model. Matriks ini terdiri dari empat sel: bagian kiri atas (6) menunjukkan jumlah mahasiswa yang benar-benar lulus dan diprediksi lulus oleh model (True Positive), kanan atas (0) adalah mahasiswa yang lulus namun salah diprediksi sebagai tidak lulus (False Negative), kiri bawah (1) adalah mahasiswa yang tidak lulus namun salah diprediksi sebagai lulus (False Positive), dan kanan bawah (3) menunjukkan mahasiswa yang tidak lulus dan berhasil diprediksi tidak lulus (True Negative). Visualisasi ini memperkuat hasil evaluasi sebelumnya, bahwa model bekerja dengan baik dalam mengenali mahasiswa yang lulus, tetapi masih terdapat kesalahan dalam mengidentifikasi sebagian kecil mahasiswa yang tidak lulus. Dengan demikian, confusion matrix ini memberikan gambaran menyeluruh tentang keberhasilan dan kekeliruan model dalam klasifikasi dua kelas.

4. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma Support Vector Machine (SVM) mampu digunakan untuk memprediksi kelulusan mahasiswa berdasarkan data akademik yang meliputi IPK, jumlah SKS lulus, jumlah mata kuliah tidak lulus, dan semester. Melalui tahapan preprocessing data, normalisasi fitur, serta pelatihan model menggunakan kernel Radial Basis Function (RBF), model yang dibangun berhasil mencapai akurasi sebesar 90%, precision sebesar 88%, recall sebesar 92%, dan F1-score sebesar 90%. Hasil tersebut menunjukkan bahwa SVM memiliki kemampuan yang baik dalam mengklasifikasikan status kelulusan mahasiswa dan berpotensi mendukung pengambilan keputusan akademik serta sistem peringatan dini bagi mahasiswa yang berisiko mengalami keterlambatan kelulusan. Penelitian ini masih memiliki keterbatasan karena menggunakan data dummy dengan jumlah sampel yang relatif kecil sehingga belum sepenuhnya merepresentasikan kondisi akademik mahasiswa yang sebenarnya. Oleh karena itu, penelitian selanjutnya disarankan menggunakan data akademik riil dengan jumlah data yang lebih besar serta menerapkan optimasi parameter atau metode seleksi fitur untuk meningkatkan performa model.

REFERENCES

- [1] Z. Ersozlu, S. Taheri, and I. Koch, "A Review Of Machine Learning Methods Used For Educational Data," *Educ. Inf. Technol.*, vol. 29, no. 16, pp. 22125–22145, 2024, doi: 10.1007/s10639-024-12704-0.
- [2] M. A. S. Pawitra, H.-C. Hung, and H. Jati, "A Machine Learning Approach to Predicting On-Time Graduation in Indonesian Higher Education," *Elinvo (Electronics, Informatics, Vocat. Educ.)*, vol. 9, no. 2, pp. 294–308, 2024, doi: 10.21831/elinvo.v9i2.77052.
- [3] K. Okoye, J. T. Nganji, J. Escamilla, and S. Hosseini, "Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education," *Comput. Educ. Artif. Intell.*, vol. 6, no. September 2023, 2024, doi: 10.1016/j.caeai.2024.100205.
- [4] L. Aulck, N. Velagapudi, J. Blumenstock, and J. West, "Predicting Student Dropout in Higher Education," 2022, [Online]. Available: <http://arxiv.org/abs/1606.06364>
- [5] Y. Zhang, Y. Yun, R. An, J. Cui, H. Dai, and X. Shang, "Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis," *Front. Psychol.*, vol. 12, no. December, pp. 1–19, 2021, doi: 10.3389/fpsyg.2021.698490.
- [6] Y. Wang, A. Ding, K. Guan, S. Wu, and Y. Du, "Graph-based Ensemble Machine Learning for Student Performance Prediction," 2021, [Online]. Available: <http://arxiv.org/abs/2112.07893>
- [7] I. F. Mahdy, M. N. Faladiba, N. Azizah, and K. Rifai, "Classification of Unisba Students' Graduation Time using Support Vector Machine Optimized with Grid Search Algorithm," vol. 21, no. 1, pp. 205–214, 2024, doi: 10.20956/j.v21i1.36257.
- [8] M. H. bin Roslan and C. J. Chen, "Educational Data Mining for Student Performance Prediction: A Systematic Literature Review (2015-2021)," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 5, pp. 147–179, 2022, doi: 10.3991/ijet.v17i05.27685.
- [9] E. Haryatmi and Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu, "Title Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu," *J. RESTI (Rekayasa Sist. dan Teknol. Informatika)*, vol. 5, no. 2, pp. 386–392, 2021, [Online]. Available: 10.29207/resti.v5i2.3007
- [10] L. M. Huizen, M. B. A., and M. Idris, "Meningkatkan kinerja SVM: Dampak berbagai teknik seleksi fitur pada akurasi prediksi," vol. 22, no. 1, pp. 1–14, 2025, doi: 10.24246/aiti.v22i1.1-14
- [11] N. Putu, E. Marita, and A. Muliantara, "Memprediksi Kelulusan Mahasiswa : Graduate dan Dropout dengan Support Vector Machine dan GridSearchCV," vol. 2, pp. 475–480, 2024, doi: 10.24843/JNATIA.2024.v02.i03.p04
- [12] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, 2022, doi: 10.1186/s40561-022-00192-z.
- [13] M. Adewale, "Students' Academic Performance In Open And Distance Learning," *A Support VECTOR Mach. Process Framew. Predict. STUDENTS' Acad. Perform. OPEN DISTANCE Learn.*, 2024.
- [14] Y. Chen, J. Sun, J. Wang, L. Zhao, and X. Song, "Machine Learning-Driven Student Performance Prediction for Enhancing Tiered Instruction," pp. 1–19, doi: 10.48550/arXiv.2502.03143
- [15] S. Siswadi, E. Khairunnisa, and N. P. Indah, "Predicting Students' On-Time Graduation using Support Vector Machine (SVM) Algorithm," *OMEGA J. Keilmuan Pendidik. Mat.*, vol. 4, no. 3, pp. 181–186, 2025, doi: 10.47662/jkpm.v4i3.1094.
- [16] M. S. N. Al-din, H. Ali, and A. Abdulqader, "Students' Academic Performance Prediction Using Educational Data Mining and Machine Learning : A Systematic Review," vol. VIII, no. 2454, pp. 1264–1291, 2024, doi: 10.47772/IJRISS.
- [17] A. Bhusal, "Predicting Student's Performance Through Data Mining," *arXiv Prepr. arXiv2112.01247*, 2021, [Online]. Available: <https://arxiv.org/abs/2112.01247><https://arxiv.org/pdf/2112.01247>
- [18] 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.

- [19] S. Abdillah, G. J. Yanris, and V. Sihombing, "Implementation Of The Support Vector Machine Method In Predicting Student Graduation," vol. 6, no. 1, pp. 263–270, 2025, doi: 10.46729/ijstm.v6i1.1265.
- [20] A. Efendi, I. Fitri, and G. W. Nurcahyo, "Improving Student Graduation Timeliness Prediction Using SMOTE and Ensemble Learning with Stacking and GridSearchCV Optimization," *Data Metadata*, vol. 4, 2025, doi: 10.56294/dm2025917.