

Aplikasi Cerdas Prediksi Kelulusan Mahasiswa Berbasis *Website* Menggunakan Metode *Support Vector Machine* (SVM)

Fajar Budi Kurniawan¹, Lia Farokhah²

¹Teknik Informatika, Fakultas Teknologi dan Desain, Institut Teknologi dan Bisnis ASIA Malang

²Teknik Informatika, Fakultas Teknologi dan Desain, Institut Teknologi dan Bisnis ASIA Malang

¹jarbud29@gmail.com, ²farokhah@asia.ac.id *

Abstract

Student graduation is an important measure for universities achieving accreditation and optimizing student abilities. Furthermore, late handling makes this indicator difficult to attain and affects a variety of factors. The goal of this work is to create a website for a student graduation prediction model utilizing the Support Vector Machine (SVM) approach. The data used covers a variety of characteristics, including age, gender, work position, leave history, and cumulative grade point average (GPA) from semesters 1 through 4. To address class imbalance, data is normalized and oversampled during the analysis phase. Model evaluation employs cross-validation and confusion matrix approaches. The evaluation results indicate that the SVM model has an average accuracy of 97.21%. This conclusion suggests that the created model is reliable for forecasting student graduation, which contributes significantly to academic management in higher education institutions.

Keywords: support vector machine, graduation prediction, machine learning, data analysis, higher education

Abstrak

Kelulusan mahasiswa merupakan salah satu indikator penting bagi perguruan tinggi dalam akreditasi dan pemaksimalan kemampuan mahasiswa. Selain itu, penanganan yang terlambat membuat indikator ini sulit dicapai dan berdampak pada berbagai hal. Penelitian ini bertujuan untuk mengembangkan *website* model prediksi kelulusan mahasiswa menggunakan metode *Support Vector Machine* (SVM). Data yang digunakan mencakup berbagai fitur, seperti umur, jenis kelamin, status pekerjaan, riwayat cuti, dan Indeks Prestasi Semester (IPS) dari semester 1 hingga 4. Proses analisis meliputi normalisasi data, *oversampling* untuk mengatasi ketidakseimbangan kelas. Evaluasi model menggunakan teknik *cross-validation* dan *confusion matrix*. Hasil evaluasi menunjukkan bahwa model SVM mencapai rata-rata akurasi sebesar 97,21%. Temuan ini menunjukkan bahwa model yang dikembangkan dapat diandalkan untuk memprediksi kelulusan mahasiswa, memberikan kontribusi signifikan terhadap pengelolaan akademik di institusi pendidikan tinggi.

Kata kunci: *support vector machine*, prediksi kelulusan, *machine learning*, analisis data, pendidikan tinggi

©This work is licensed under a Creative Commons Attribution -ShareAlike 4.0International License

1. Pendahuluan

Pembangunan sumber daya manusia yang unggul bergantung pada pendidikan tinggi. Seseorang dapat memperoleh derajat, pangkat, atau jabatan yang lebih tinggi dengan mendapatkan gelar atau ijazah dari perguruan tinggi. [1]. Perguruan tinggi tidak hanya bertanggung jawab dalam menyediakan pendidikan yang bermutu, tetapi juga harus memastikan bahwa mahasiswa dapat menyelesaikan studi mereka tepat waktu. Ini juga menjadi salah satu kriteria penting untuk akreditasi institusi pendidikan tinggi. Di Indonesia, Akreditasi merupakan salah satu jaminan mutu dalam sebuah perguruan tinggi khususnya di program studi. Tantangan utama yang dihadapi oleh program studi adalah ketika penanganannya terlambat maka akan lebih sulit untuk dicapai. Hal ini dikarenakan kelulusan tepat waktu melibatkan setidaknya jenjang delapan semester atau 4 tahun. Sangat penting untuk melakukan evaluasi dan pemantauan kelulusan siswa; salah satu cara untuk memastikan bahwa siswa menyelesaikan kuliah tepat waktu adalah dengan mengumpulkan data sejarah kelulusan mereka [2]. Disisi lain, ada keterbatasan sks yang diambil dalam setiap satu semester sesuai dengan aturan [3]. Dalam sebuah kurikulum program studi,

terdapat mata kuliah prasyarat untuk mata kuliah yang lain. Hal ini akan membuat penanganan yang terlambat lebih rumit. Tingkat kelulusan tepat waktu pada Program Studi Teknik Informatika menunjukkan pola fluktuatif dari tahun ke tahun. Peringkat akreditasi program studi dan perguruan tinggi dapat dicabut sebelum masa berlakunya jika ditemukan bahwa program tersebut tidak lagi memenuhi persyaratan akreditasi [4].

Untuk mengatasi masalah kelulusan tepat waktu, diperlukan sistem yang dapat memprediksi kemungkinan seorang mahasiswa menyelesaikan studinya tepat waktu. Saat ini, belum ada sistem yang efektif untuk memberikan gambaran tersebut, sehingga program studi tidak dapat melakukan penanganan dini dan pencegahan. Tujuan dari penelitian ini adalah untuk menggunakan metode *Support Vector Machine* (SVM) untuk merancang dan membangun sebuah *website* yang memprediksi kelulusan mahasiswa Teknik Informatika di Institut Teknologi dan Bisnis Asia Malang. Bidang ilmu komputer lain yang dikenal sebagai *machine learning* adalah pengembangan algoritma yang memungkinkan komputer untuk belajar melalui data, yang sering disebut sebagai "*learn from data*". [5]. Dengan adanya teknologi berbasis *machine*

learning, institusi pendidikan dapat melakukan analisis prediksi untuk menentukan mahasiswa yang berisiko tidak lulus tepat waktu. Hal ini juga mendukung pembuatan sistem cerdas dan tidak hanya sistem pencatatan.

Beberapa penelitian menunjukkan bahwa metode SVM memiliki keunggulan dalam prediksi dibandingkan dengan metode lainnya. Junaidi, Anggela, dan Kariman (2024) Dalam penelitian mereka, mereka menemukan bahwa SVM lebih baik daripada algoritma *Naive Bayes* dan *Random Forest* dalam memprediksi kelulusan mahasiswa tepat waktu [6]. Selain itu, Sadat et al. (2023) melakukan perbandingan antara SVM dan algoritma pohon keputusan, dan hasilnya menunjukkan bahwa SVM lebih unggul dalam menentukan penerimaan mahasiswa di universitas negeri [7]. Selanjutnya, Hidayatunnisa'i et al. (2023) melakukan perbandingan antara algoritma *Naive Bayes* dan *Support Vector Machine* (SVM), dan hasilnya menunjukkan bahwa SVM lebih unggul dalam hal akurasi dan waktu eksekusi saat menganalisis kualitas butir soal [8]. Dan yang terakhir, Firdaus et al. (2024) melakukan analisis sentimen terhadap komentar masyarakat mengenai penetapan Hari Raya Idul Adha 2023 pada video YouTube, dan hasilnya menunjukkan bahwa algoritma *Support Vector Machine* (SVM) lebih unggul dengan akurasi 96% dibandingkan *Random Forest* yang memperoleh akurasi 94% [9]. Temuan-temuan ini menegaskan bahwa SVM merupakan pilihan yang lebih baik untuk berbagai aplikasi prediktif dalam konteks pendidikan.

Tujuan penelitian ini adalah merancang dan membangun website prediksi kelulusan mahasiswa Teknik Informatika menggunakan metode SVM. Faktor internal dan eksternal setidaknya mempengaruhi keterlambatan kelulusan mahasiswa [10]. Penelitian ini akan membatasi ruang lingkup pada prediksi kelulusan mahasiswa Teknik Informatika di Institut Teknologi dan Bisnis Asia Malang, dengan data yang digunakan terbatas pada variabel Indeks Prestasi Semester (IPS) semester 1-4, status pekerjaan, jenis kelamin, akumulasi nilai, dan riwayat cuti kuliah. Diharapkan bahwa sistem ini akan membantu lembaga dalam mengevaluasi dan memprediksi kemungkinan mahasiswa akan menyelesaikan studi mereka pada waktu yang telah ditentukan. Selain itu, sistem ini dapat memberikan informasi yang lebih akurat tentang faktor-faktor yang mempengaruhi kelulusan siswa, sehingga pihak kampus dapat mengambil tindakan yang lebih bijaksana untuk meningkatkan jumlah mahasiswa yang lulus tepat waktu.

Penelitian ini diharapkan akan membuat aplikasi cerdas yang membantu mahasiswa dan institusi mengantisipasi masalah akademik tentang faktor kelulusan tepat waktu yang berdampak pada banyak hal. Aplikasi cerdas ini akan bermanfaat bagi pihak institusi dalam memberikan intervensi penanganan dini bagi mahasiswa yang berisiko tidak lulus tepat waktu mulai semester 4. Dengan adanya teknologi berbasis

machine learning, sistem ini dapat memberikan rekomendasi yang lebih objektif dan berbasis data, sehingga pengambilan keputusan oleh akademik dapat dilakukan dengan lebih efektif. Selain itu, website prediksi ini diharapkan dapat digunakan sebagai alat pemantauan yang membantu dosen pembimbing akademik memberikan arahan dan bimbingan kepada mahasiswa mereka tentang cara mengoptimalkan waktu studi mereka.

2. Metode Penelitian

2.1 Perancangan

Perancangan sistem prediksi kelulusan tepat waktu mahasiswa Teknik Informatika di ITB Asia Malang dimulai dengan identifikasi tujuan dan kebutuhan sistem. Tujuan penelitian ini adalah untuk membuat alat yang menggunakan teknologi *machine learning* untuk memprediksi mahasiswa akan lulus tepat waktu atau tidak, khususnya metode SVM. Untuk mengetahui apakah mayoritas mahasiswa dapat menyelesaikan pendidikan tepat waktu atau justru mengalami keterlambatan hingga tidak dapat menyelesaikan studinya, diperlukan penelitian yang dapat mengidentifikasi faktor-faktor penyebab kelulusan mahasiswa [11]. Peneliti melakukan analisis menyeluruh terhadap komponen yang mempengaruhi kelulusan dalam perancangan ini, seperti IPS semester 1-4, riwayat cuti, akumulasi nilai mata kuliah, dan jenis kelamin mahasiswa. Dengan memahami faktor-faktor ini, sistem dapat dirancang untuk memberikan prediksi yang lebih akurat dan relevan.

Data yang digunakan dalam penelitian ini terdiri dari data historis mahasiswa yang dikumpulkan dari sistem informasi akademik Institut Teknologi dan Bisnis ASIA Malang. Data ini meliputi informasi akademik, non-akademik, dan faktor-faktor pendukung lainnya yang dapat mempengaruhi kelulusan. Peneliti juga melakukan wawancara dengan dosen dan pihak program studi untuk mendapatkan wawasan tambahan mengenai faktor-faktor yang mungkin tidak tercakup dalam data yang ada. Karena pola komunikasi dosen dan mahasiswa sangat berpengaruh bagi mahasiswa [12].

Berhasil atau tidaknya penelitian ilmiah pendidikan bergantung pada teknik pengumpulan data dan instrumen penelitian [13]. Teknik pengumpulan data dilakukan melalui beberapa tahap, termasuk pengumpulan data kuantitatif dari sistem informasi akademik dan pengumpulan data kualitatif melalui wawancara. Data kuantitatif yang diperoleh kemudian diolah melalui tahap data *preprocessing*, yang mencakup pembersihan data, normalisasi, dan transformasi variabel. Teknik pra-pemrosesan dalam proses penambangan data telah terbukti memiliki dampak signifikan terhadap kinerja keseluruhan dalam tugas penambangan data [14]. Langkah ini krusial untuk menjamin bahwa data yang digunakan dalam melatih model SVM memiliki kualitas tinggi dan siap

untuk dianalisis. Dengan perancangan yang matang, sistem diharapkan dapat memberikan hasil prediksi yang akurat dan bermanfaat bagi pihak program studi.

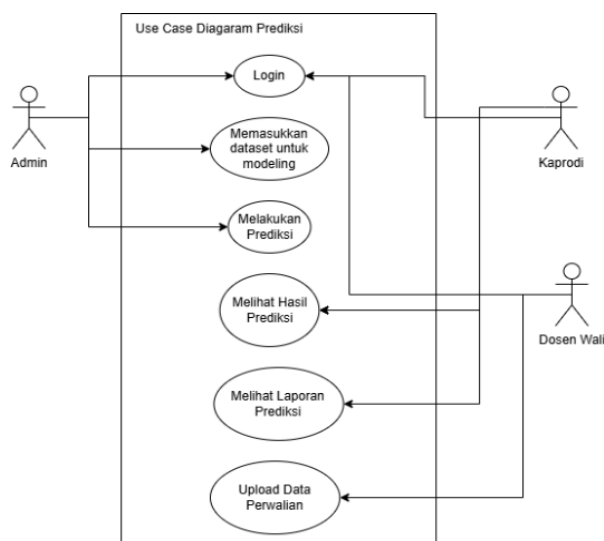
2.2 Gambaran Alur Sistem

Implementasi sistem prediksi kelulusan tepat waktu dilakukan dengan mengembangkan aplikasi berbasis website yang dapat diakses oleh pengguna, termasuk mahasiswa dan dosen wali. Aplikasi ini dibuat dengan antarmuka yang ramah pengguna, sehingga memudahkan *stakeholder* dalam memasukkan data yang dibutuhkan, menjelaskan berbagai fungsionalitas *stakeholder* dalam sistem. Kaprodi dapat mengakses laporan kelulusan mahasiswa berdasarkan data yang tersedia, sementara Admin bertanggung jawab untuk mengunggah dataset guna mendukung analisis lebih lanjut. Selain itu, Dosen Wali memiliki peran dalam mengisi serta memperbarui data perwalian mahasiswa yang dibimbingnya agar tetap akurat dan terkini. Fungsionalitas *stakeholder* ditunjukkan pada Tabel 1.

Table 1. *Table* Fungsionalitas Stakeholder

No	Use Case Name	Aktor	Deskripsi
1	Melihat Laporan Kelulusan	Kaprodi	Kaprodi dapat melihat kelulusan mahasiswa secara tepat waktu berdasarkan data yang tersedia.
2	Memasukkan Dataset	Admin	Admin mengunggah dataset untuk kebutuhan modelling guna mendukung analisis lebih lanjut.
3	Upload Data Perwalian	Dosen Wali	Dosen Wali mengisi dan memperbarui data perwalian mahasiswa yang dia bimbing.

Adapun *use case* diagram bisa dilihat pada Gambar 1 dibawah ini.



Gambar 1. *Use Case* Diagram Alur Sistem Prediksi

Setelah seluruh data berhasil dihimpun, langkah selanjutnya adalah mengolahnya dengan pendekatan *Support Vector Machine* (SVM). Metode ini dipilih karena kemampuannya dalam mengenali pola tersembunyi di dalam data yang kompleks. Model SVM yang telah melalui proses pelatihan sebelumnya kemudian digunakan untuk menganalisis dan mengelompokkan data berdasarkan karakteristik tertentu. Dalam hal ini, data mahasiswa akan dipisahkan ke dalam dua kategori utama, yakni mereka yang berpotensi lulus tepat waktu dan mereka yang diprediksi mengalami keterlambatan kelulusan.

Proses klasifikasi dilakukan dengan mencari garis batas (*margin*) yang paling optimal agar pemisahan antar kelas menjadi lebih jelas, meskipun data yang digunakan bersifat non-linear atau memiliki banyak dimensi. Dengan pendekatan ini, diharapkan prediksi yang dihasilkan menjadi lebih akurat dan dapat dijadikan dasar dalam pengambilan keputusan. Selain itu, penerapan model ini juga mempertimbangkan keunikan dari masing-masing data mahasiswa, sehingga hasil analisis tidak hanya bergantung pada angka semata, tetapi juga mencerminkan kondisi yang lebih mendalam dan kontekstual.

Hasil analisis dari sistem prediksi kemudian divisualisasikan untuk memudahkan pengguna dalam memahami informasi yang disajikan. Visualisasi data berperan penting dalam mendukung analisis data secara efektif [15]. Visualisasi ini tidak hanya membantu pengguna dalam menginterpretasikan hasil prediksi. Selain itu, visualisasi data juga membantu dalam memberikan pemahaman yang lebih jelas tentang faktor-faktor yang memengaruhi kelulusan. Dengan implementasi yang baik, sistem ini diharapkan dapat memberikan dukungan yang signifikan bagi pihak program studi dalam mengambil keputusan terkait intervensi akademik bagi mahasiswa yang berisiko tidak lulus tepat waktu.

2.3 Pengujian

Uji coba sistem prediksi kelulusan tepat waktu dilakukan untuk memastikan bahwa aplikasi bekerja dengan optimal dan menghasilkan output yang akurat. Pengujian dilakukan agar aplikasi dapat beradaptasi dengan kebutuhan [16]. Proses pengujian dimulai dengan pengujian fungsionalitas, Setiap fitur dalam aplikasi diuji untuk memastikan bahwa seluruh fungsionalitas beroperasi sesuai dengan ekspektasi. Pengujian kedua, pengujian terhadap antarmuka pengguna untuk memastikan bahwa aplikasi mudah digunakan dan tidak membingungkan bagi pengguna. Umpan balik dari pengguna awal, termasuk mahasiswa dan dosen, sangat penting dalam tahap ini untuk melakukan perbaikan yang diperlukan.

Setelah tahapan pengujian antarmuka selesai dilakukan, fokus selanjutnya diarahkan pada evaluasi akurasi dari model *Support Vector Machine* (SVM) yang telah dibangun. Dalam hal ini, peneliti memanfaatkan data historis yang sebelumnya telah

dikumpulkan untuk menguji seberapa akurat model tersebut dalam memprediksi kelulusan mahasiswa. Pengujian ini menjadi krusial karena akan menunjukkan apakah model mampu mengenali pola yang relevan dan membuat klasifikasi yang tepat berdasarkan data yang tersedia.

Untuk menilai kinerja model secara menyeluruh, digunakan beberapa metrik evaluasi seperti akurasi, presisi, dan *recall*. Ketiga ukuran ini memberikan perspektif yang saling melengkapi tentang seberapa baik model bekerja dalam konteks prediksi kelulusan tepat waktu. Hasil yang diperoleh tidak hanya menggambarkan efektivitas sistem, tetapi juga menjadi acuan penting bagi peneliti dalam mengevaluasi kembali parameter model atau strategi pelatihan jika ditemukan ketidaksesuaian antara prediksi dan kenyataan di lapangan.

Pengujian sistem terakhir mencakup pengujian terhadap keandalan dan keamanan aplikasi. Peneliti memastikan bahwa data dalam basis data terlindungi dan tidak mudah diakses oleh pihak yang tidak berwenang. Keamanan database memiliki peran krusial dalam menjaga informasi sensitif yang tersimpan di dalamnya [17]. Dengan melakukan pengujian menyeluruh, peneliti dapat memastikan bahwa sistem prediksi kelulusan tepat waktu tidak hanya efektif dalam memberikan hasil yang akurat, tetapi juga aman dan dapat diandalkan untuk digunakan oleh pihak program studi dalam mendukung pengembangan akademik mahasiswa.

3. Hasil dan Pembahasan

3.1 Pengolahan Data Awal

Pengolahan data awal merupakan langkah krusial dalam proses analisis data mahasiswa yang akan digunakan dalam sistem prediksi kelulusan tepat waktu. Data yang diperoleh mencakup berbagai informasi penting, seperti umur, jenis kelamin, status pekerjaan, riwayat cuti, serta IPS untuk semester 1 hingga 4, dan status kelulusan. Namun, untuk memastikan bahwa data yang digunakan dalam pemrosesan klasifikasi adalah relevan dan berkualitas, beberapa langkah perlu dilakukan. Pertama, melakukan seleksi terhadap data yang tidak diperlukan untuk pembentukan model. Dalam hal ini, informasi seperti nama mahasiswa dihapus dari dataset, karena tidak memberikan kontribusi signifikan terhadap analisis prediksi kelulusan atau hanya sebagai id. Dataset untuk pembentukan model bisa dilihat pada Tabel 2.

Table 2. Contoh Dataset Untuk Pembentukan Model

NAM	U	JEN	STA	PE	I	I	I	I	STAT
A	M	IS	TUS	RN	P	P	P	P	US
U	KE	BE	AH	S	S	S	S	S	KEL
R	LA	KE	CU	1	2	3	4	4	ULUS
		MIN	RJA	TI					AN
ANIK									
WIDA		Pere			2,			3,	
YANT		mpu			7	2,	3,	1	Terla
I	20	an	Ya	Iya	6	8	2	7	mbat

DWI									
HES									
YNA									
PRIH		Pere							
ASTA		mpu							
NTY	20	an	Tida	Tid	3,	3,	3,	3,	Terla
MUR			k	ak	3	3	4	4	mbat
YA									
ARIEF		Pere							
BASU		mpu							
KI	23	an	Ya	Iya	3,	3,	3,	2	Terla
NANI					5	3	7	9	mbat
K		Pere							
SUSA		mpu							
NTI	19	an	Tida		3,	3,	3,	3,	Terla
RIFK			k	Iya	1	4	6	3	mbat
A					7	1	1	6	
ISTIQ		Pere							
FARI		mpu							
NA	19	an	Ya	Iya	2,	8	3,	8	Terla
SUHA					9	9	3	5	mbat
RYON		Laki							
O	20	-laki	Ya	Iya	2,	2,	3,		
FARI					9	8	0	3,	Terla
KHAT					5	2	9	1	mbat
UN		Pere							
NAZL		mpu							
I	21	an	Tida	Tid	2,	3,		2,	
			k	ak	7	1	2,	9	
					6	4	6	5	Tepat

Selanjutnya, transformasi data untuk mengubah format isian dari huruf atau string menjadi angka atau integer. Proses ini penting karena algoritma klasifikasi, seperti SVM, memerlukan data numerik untuk melakukan analisis. Proses ini disebut sebagai proses encoding. Proses encoding, yang juga dikenal sebagai penyandian, dapat dilakukan satu kali atau berulang kali sesuai kebutuhan [18]. Misalnya, status pekerjaan dapat dikodekan dengan angka, di mana 0 menunjukkan tidak bekerja dan 1 menunjukkan bekerja. Demikian pula, variabel lain seperti jenis kelamin dan riwayat cuti juga diubah menjadi format numerik. Melalui tahap *preprocessing*, dataset dipersiapkan agar siap digunakan dalam proses klasifikasi, sehingga dapat meningkatkan akurasi serta efektivitas model prediksi yang dikembangkan. Dataset untuk pembentukan model setelah dilakukan *preprocessing* ditunjukkan pada Tabel 3.

Table 3. Contoh Dataset Setelah Dibersihkan

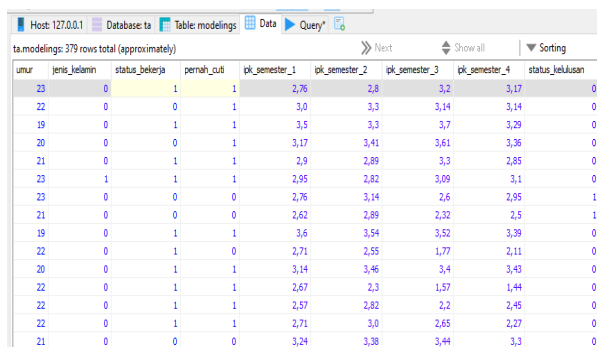
UM	JENI	STAT	PER	IP	IP	IP	IP	STATU
UR	S	US	NAH	S	S	S	S	S
	KEL	BEK	CUT	1	2	3	4	KELUL
	AMIN	ERJA	I					USAN
				2.	2.	3.	3.	
23	0	1	1	76	8	2	17	0
					3.	3.	3.	
22	0	0	1	3	3	14	14	0
					3.	3.	3.	
19	0	1	1	5	3	7	29	0
					3.	3.	3.	
20	0	0	1	17	41	61	36	0
					2.	2.	3.	
21	0	1	1	9	89	3	85	0
					2.	2.	3.	
23	1	1	1	95	82	09	1	0
					2.	3.	2.	
23	0	0	0	76	14	6	95	1

3.2 Eksperimen dan Pengujian Model

3.2.1 Dataset dimasukkan ke database mysql

Pada pengolahan data awal sudah selesai dilakukan, dataset dimasukkan ke dalam database MySQL menggunakan *plugin* Laravel Excel Import. Plugin ini

memungkinkan pengelolaan impor data dari *file Excel* secara efisien, sehingga data dapat langsung terintegrasi ke dalam *database* tanpa proses manual yang memakan waktu. Dengan menyiapkan *file Excel* yang berisi *dataset* mahasiswa, langkah selanjutnya adalah membuat kelas *import* yang mendefinisikan pemetaan data dari file ke kolom-kolom dalam tabel *database*. Setelah proses impor selesai, data mahasiswa tersimpan dengan rapi dalam *database*, siap digunakan untuk analisis lebih lanjut dan pemrosesan klasifikasi dalam sistem prediksi kelulusan tepat waktu. Gambaran *dataset* yang dimasukkan ditunjukkan pada Gambar 2.



umur	jenis_kelamin	status_bekerja	pernah_cuti	ipk_semester_1	ipk_semester_2	ipk_semester_3	ipk_semester_4	status_kelulusan
23	0	1	1	2,76	2,8	3,2	3,17	0
22	0	0	1	3,0	3,3	3,14	3,14	0
19	0	1	1	3,5	3,3	3,7	3,29	0
20	0	0	1	3,17	3,41	3,61	3,36	0
21	0	1	1	2,9	2,89	3,3	2,85	0
23	1	1	1	2,95	2,82	3,09	3,1	0
23	0	0	0	2,76	3,14	2,6	2,95	1
21	0	0	0	2,62	2,89	2,32	2,5	1
19	0	1	1	3,6	3,54	3,52	3,39	0
22	0	1	0	2,71	2,55	1,77	2,11	0
20	0	1	1	3,14	3,46	3,4	3,43	0
22	0	1	1	2,67	2,3	1,57	1,44	0
22	0	1	1	2,57	2,82	2,2	2,45	0
22	0	1	1	2,71	3,0	2,65	2,27	0
21	0	0	0	3,24	3,38	3,44	3,3	0

Gambar 2. *Dataset* Yang Sudah Dimasukkan Ke Database MySQL

3.2.2 Melakukan pemrosesan *dataset* dengan menggunakan metode SVM di Laravel

Proses pengolahan *dataset* dengan metode SVM di dalam *framework* Laravel dimulai dengan menarik data yang sebelumnya telah tersimpan dalam basis data. Data ini mencakup sejumlah fitur penting yang diyakini berpengaruh terhadap prediksi kelulusan mahasiswa, seperti usia, jenis kelamin, status pekerjaan, riwayat cuti, serta IPS dari semester pertama hingga keempat. Selain itu, status kelulusan mahasiswa juga turut disertakan sebagai bagian penting dari proses klasifikasi.

Setelah data berhasil diperoleh, setiap fitur dipersiapkan dalam bentuk *array* agar mudah diolah oleh model SVM. Sementara itu, label yang menunjukkan apakah seorang mahasiswa lulus tepat waktu atau tidak dijadikan acuan utama dalam proses pelatihan model. Dengan pendekatan ini, sistem dapat memanfaatkan informasi yang tersedia secara maksimal untuk mengidentifikasi pola-pola tertentu yang relevan. Hal ini bertujuan agar hasil analisis yang dihasilkan tidak hanya akurat, tetapi juga dapat digunakan sebagai dasar pengambilan keputusan di masa mendatang.

Setelah menyiapkan *dataset*, langkah berikutnya adalah melakukan normalisasi fitur dengan menghitung *Z-Score*. Normalisasi berperan penting dalam menyamakan skala semua fitur, sehingga tidak ada fitur yang memiliki pengaruh lebih besar dalam proses pelatihan model. Selanjutnya, *dataset* dipisahkan berdasarkan label kelulusan, di mana fitur dari

mahasiswa yang lulus dan tidak lulus dikelompokkan. Untuk mengatasi masalah ketidakseimbangan kelas, *oversampling* dilakukan dengan menduplikasi data dari kelas minoritas, sehingga jumlah data di setiap kelas menjadi seimbang. Setelah itu, *dataset* diacak untuk memastikan bahwa data tercampur dengan baik sebelum dibagi menjadi data pelatihan dan pengujian.

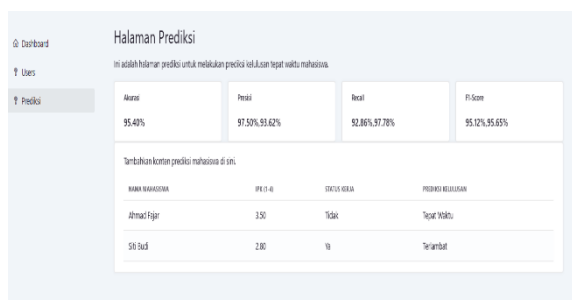
Setelah data pelatihan disiapkan, langkah selanjutnya adalah membangun dan melatih model SVM dengan data tersebut. Proses pelatihan ini memungkinkan model untuk mengenali pola-pola dari berbagai fitur yang telah ditentukan sebelumnya. Setelah pelatihan selesai, model kemudian digunakan untuk melakukan prediksi terhadap data pengujian guna mengevaluasi seberapa baik model mampu mengklasifikasikan data baru yang belum pernah dilihat sebelumnya.

Untuk mengukur kinerja model secara objektif, hasil prediksi dibandingkan langsung dengan label sebenarnya. Dari perbandingan ini, dihitung metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-Score*. Selain itu, digunakan juga *confusion matrix* yang memberikan gambaran lebih mendalam mengenai distribusi prediksi benar dan salah pada tiap kelas. Seluruh hasil evaluasi ini disajikan dalam format JSON yang terstruktur, sehingga memudahkan proses analisis lanjutan maupun dokumentasi. Pendekatan ini membantu peneliti untuk menilai apakah model sudah cukup andal atau masih perlu disempurnakan.

3.2.3 Menampilkan akurasi, *precision*, *recall*, dan *F1-Score* di tampilan *website*

Data berhasil diambil secara otomatis dari endpoint yang telah ditentukan sebelumnya. Setelah proses pemanggilan data selesai, konten utama dari halaman ditampilkan dalam bentuk visual yang informatif. Salah satu bagian utama dari antarmuka ini adalah deretan kartu yang menampilkan metrik kinerja model, seperti akurasi, presisi, *recall*, dan *F1-Score*. Masing-masing metrik ditampilkan dalam format persentase hingga dua angka di belakang koma, sehingga memudahkan pengguna dalam memahami seberapa baik model bekerja berdasarkan data yang diuji.

Selain kartu metrik, halaman ini juga menyajikan sebuah tabel yang memuat rincian hasil prediksi terhadap mahasiswa. Informasi yang ditampilkan meliputi nama mahasiswa, IPK terkini, status pekerjaan, serta hasil prediksi kelulusan. Tabel ini berfungsi sebagai pelengkap yang memberikan konteks lebih dalam terhadap performa model dalam skenario nyata. Dengan tampilan yang terstruktur, pengguna dapat menelusuri hasil prediksi secara lebih rinci dan memahami bagaimana model *Support Vector Machine* diaplikasikan dalam konteks akademik. Visualisasi antarmuka halaman ini ditampilkan pada Gambar 3 sebagai referensi.



Gambar 3. Tampilan Website Untuk Menunjukkan Hasil Prediksi

3.2.4 Evaluasi kinerja model menggunakan cross-validation

Hasil evaluasi kinerja model menggunakan *cross-validation* menunjukkan performa yang sangat memuaskan. *Cross validation* berpengaruh terhadap hasil akurasi yang didapatkan [19]. Rata-rata akurasi model mencapai 97,21%, yang mengindikasikan kemampuan model dalam memprediksi kelulusan mahasiswa dengan tingkat keakuratan yang tinggi. Selain itu, rata-rata *precision* untuk kedua kelas adalah 97,27% dan 97,14%, sedangkan rata-rata *recall* masing-masing kelas adalah 97,25% dan 97,13%. Metrik *F1-Score* juga menunjukkan hasil yang baik, dengan nilai rata-rata 97,25% untuk kelas pertama dan 97,12% untuk kelas kedua. Temuan ini menegaskan bahwa model tidak hanya akurat, tetapi juga efektif dalam mengidentifikasi mahasiswa yang berpotensi lulus tepat waktu.

Selanjutnya, analisis terhadap *confusion matrix* yang dihasilkan dari setiap *fold* dalam proses *cross-validation* memberikan gambaran yang lebih mendalam mengenai performa model secara menyeluruh. Setiap matriks mencatat jumlah prediksi yang tepat maupun yang meleset untuk masing-masing kelas, sehingga memudahkan dalam mengidentifikasi pola kesalahan. Dari hasil yang diperoleh, sebagian besar prediksi model berada pada kategori benar, terutama dalam mengklasifikasikan mahasiswa yang diperkirakan lulus tepat waktu. Hal ini menunjukkan bahwa model bekerja secara konsisten dalam berbagai pengujian lintas data.

Sebagai contoh, pada salah satu *fold*, model berhasil mengklasifikasikan 41 mahasiswa yang lulus dengan akurat, sementara hanya terjadi satu kesalahan dalam mengidentifikasi mahasiswa yang seharusnya tidak lulus. Capaian ini menegaskan bahwa model SVM yang dikembangkan tidak hanya memiliki tingkat akurasi yang tinggi secara numerik, tetapi juga dapat diandalkan dalam situasi nyata. Kemampuan model untuk mempertahankan performa stabil dari satu *fold* ke *fold* lainnya menambah kepercayaan terhadap validitas prediksi yang dihasilkan. Rincian hasil evaluasi menggunakan *cross-validation* ini divisualisasikan pada Gambar 4 sebagai referensi tambahan.



Gambar 4. Hasil Evaluasi Menggunakan Cross-Validation

3.2.5 Pengujian menu utama aplikasi menggunakan black box testing

Pengujian pada menu utama aplikasi dilakukan menggunakan metode *black box testing*, yang berfokus pada pengujian fungsionalitas tanpa melihat struktur internal kode program. Pengujian dilakukan terhadap tiga menu utama, yaitu *dashboard*, *users*, dan *Prediksi*, dengan tujuan untuk memastikan bahwa setiap fitur pada masing-masing menu dapat berjalan sesuai dengan kebutuhan pengguna.

Pada menu *dashboard*, pengujian difokuskan pada validitas data yang ditampilkan, seperti jumlah total mahasiswa dan hasil prediksi kelulusan, apakah sesuai dengan data aktual yang tersimpan di sistem. Pada menu *users*, dilakukan pengujian terhadap fungsi-fungsi manajemen akun seperti tambah, ubah, dan hapus pengguna untuk memastikan sistem merespons input dengan benar dan melakukan perubahan sesuai perintah. Sedangkan pada menu *Prediksi*, pengujian dilakukan untuk memastikan data hasil prediksi yang telah dilakukan ditampilkan secara utuh dan akurat. Setiap pengujian mencakup skenario input yang valid maupun tidak valid untuk memastikan sistem dapat menangani keduanya dengan baik.

Table 4. Pengujian Black Box Testing Menu Utama Aplikasi

N O	SKENARIO PENGUJIAN	TEST CASE	HASIL YANG DIHARAPKAN	HASIL PENGUJIAN	STATUS

1	Akses Menu Dashboard	Mengklik Tombol Menu Dashboard	Tampil Halaman Dashboard	Sesuai	Valid
2	Akses Menu Users	Mengklik Tombol Menu Users	Tampil Halaman Untuk Melakukan Manajemen User	Sesuai	Valid
3	Akses Menu Prediksi	Mengklik Tombol Menu Prediksi	Tampil Halaman Hasil Prediksi Kelulusan Mahasiswa	Sesuai	Valid

4. Kesimpulan

Hasil penelitian ini menunjukkan bahwa model prediksi kelulusan tepat waktu mahasiswa dengan metode SVM memiliki kinerja yang sangat optimal. Dengan rata-rata akurasi mencapai 97,21%, serta nilai *precision*, *recall*, dan F1-Score yang tinggi, model ini terbukti efektif dalam mengidentifikasi mahasiswa yang berpotensi lulus tepat waktu. Hasil evaluasi melalui *cross-validation* juga menunjukkan konsistensi dalam performa model, yang dapat diandalkan untuk pengambilan keputusan dalam konteks akademik. Temuan ini memberikan kontribusi signifikan terhadap upaya peningkatan kualitas pendidikan dan pengelolaan akademik di institusi pendidikan tinggi.

Untuk pengembangan lebih lanjut, disarankan agar penelitian ini mempertimbangkan penggunaan lebih banyak variabel yang dapat mempengaruhi kelulusan mahasiswa, seperti faktor sosial dan ekonomi, serta keterlibatan dalam kegiatan ekstrakurikuler. Selain itu, penerapan teknik *machine learning* lainnya, seperti Random Forest atau Neural Networks, dapat dieksplorasi untuk membandingkan kinerja dengan model SVM yang telah digunakan. Implementasi sistem prediksi ini dalam bentuk aplikasi berbasis web juga dapat membantu pihak akademik dalam memantau dan memberikan intervensi yang tepat bagi mahasiswa yang berisiko tidak lulus tepat waktu. Dengan langkah-langkah ini, diharapkan dapat tercipta lingkungan akademik yang lebih mendukung bagi mahasiswa dalam mencapai tujuan pendidikan mereka.

Ucapan Terimakasih

Terima kasih kepada dosen pembimbing dan Institut Teknologi dan Bisnis ASIA Malang

Daftar Rujukan

[1] Asril, "Membangun SDM Indonesia Yang Berkualitas Melalui Pendidikan Tinggi Dan Pendidikan Vokasi," Kemenko pmk. Accessed: Jan. 04, 2025. [Online]. Available: <https://www.kemendiknas.go.id/membangun-sdm-indonesia-yang-berkualitas-melalui-pendidikan-tinggi-dan-pendidikan-vokasi>

[2] I. Riadi, R. Umar, and R. Anggara, "Prediksi Kelulusan Tepat Waktu Berdasarkan Riwayat Akademik Menggunakan Metode K-Nearest Neighbor," *Jurnal Teknologi Informasi dan Ilmu*

Komputer, vol. 11, no. 2, pp. 249–256, Apr. 2024, doi: 10.25126/jtiik.20241127330.

[3] batvokasi, "Permendikbud No.53 Tahun 2023: Mahasiswa Tahun Pertama Maksimum Ambil 20 SKS," Universitas Diponegoro. Accessed: Jan. 12, 2025. [Online]. Available: https://bat.vokasi.undip.ac.id/2024/01/12/permendikbud-no-53-tahun-2023-mahasiswa-tahun-pertama-maksimum-ambil-20-sks/?utm_source=chatgpt.com

[4] S. Rachardian, E. Sedyono, F. "Prediksi kelulusan tepat waktu mahasiswa untuk pemantauan program studi menggunakan metode data mining," *AITI: Jurnal Teknologi Informasi*, vol. 21, no. 2, pp. 168–182, 2024. doi: 10.24246/aiti.v21i2.168-182

[5] A. Darmawan Sidik, A. Ansawarman, K. Kunci, J. Kendaraan Bermotor, M. Regresi, and F. Jalan, "Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine Learning," *Formosa Journal of Multidisciplinary Research (FJMR)*, vol. 1, no. 3, pp. 559–568, 2022, doi: 10.55927.

[6] Satrio Junaidi, R. Valicia Anggela, and D. Kariman, "Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Network (ANN)," *Journal of Applied Computer Science and Technology*, vol. 5, no. 1, pp. 109–119, Jun. 2024, doi: 10.52158/jacost.v5i1.489.

[7] M. A. Sadat, Pujiono, A. Pambudi, and S. Ibad, "COMPARISON OF ALGORITHM BETWEEN CLASSIFICATION & REGRESSION TREES AND SUPPORT VECTOR MACHINE IN DETERMINING STUDENT ACCEPTANCE IN STATE UNIVERSITIES," *JUTIF*, vol. 4, no. 6, pp. 1589–1604, Jan. 2024.

[8] Hidayatunnisa'i, Kusriani, dan Kusnawi, "Perbandingan Kinerja Metode Naïve Bayes dan Support Machine (SVM) dalam Analisis Kualitas Butir Soal," *Jurnal FASILKOM*, vol. 13, no. 2, pp. 165–172, Aug. 2023, doi: 10.37859/jf.v13i02.5087.

[9] R. Firdaus, R. A. Hariri, and H. F. Amran, "Sentimen Analisis Masyarakat Tentang Penetapan Hari Raya Idul Adha Tahun 2023 Pada Video Youtube Menggunakan Algoritma Random Forest dan Support Vector Machine," *FASILKOM*, vol. 14, no. 1, pp. 278–285, Apr. 2024, doi: [10.37859/jf.v14i1.7012](https://doi.org/10.37859/jf.v14i1.7012).

[10] Y. Nursyafti and W. Purwanto, "Faktor-Faktor Penghambat Kelulusan Tepat Waktu Mahasiswa D3 Jurusan Teknik Otomotif Fakultas Teknik Universitas Negeri Padang Tahun Masuk 2016 dan 2017."

[11] A. Fatkhudin, M. Y. Febrianto, F. A. Artanto, M. Waffa, N. Hadinata, and R. Fahlevi, "ALGORITMA DECISION TREE C.45 DALAM ANALISA KELULUSAN MAHASISWA PROGRAM STUDI MANAJEMEN INFORMATIKA UMP," vol. 8, no. 2, 2022, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>

[12] Y. Damayanti, "Pola Komunikasi Dosen Pembimbing dengan Mahasiswa dalam Penyelesaian Tugas Akhir di Masa Pandemi (Studi di Program Studi Ilmu Komunikasi Universitas Peradaban)," *jkp*, vol. 1, no. 1, pp. 19–28, Feb. 2023.

[13] Ardiansyah, Risnita, & Jailani, M. S. (2023). "Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif". *IHSAN : Jurnal Pendidikan Islam*, 1(2), 1–9. Doi: 10.61104/ihsan.v1i2.57

[14] H. M. Marin-Castro and E. Tello-Leal, "Event log preprocessing for process mining: A review," Nov. 01, 2021, MDPI. doi: 10.3390/app112210556.

[15] R. Al Ghivary, N. Wulandari, N. Srikandi, and A. M. Nazilatul F, "PERAN VISUALISASI DATA UNTUK MENUNJANG ANALISA DATA KEPENDUDUKAN DI INDONESIA," 2023.

[16] M. Dzulhi Raihan, N. Agitha, A. Yudo Husodo, F. Bimantoro, and B. Rabbani, "PENGUJIAN MULTIPLATFORM PADA APLIKASI NTB MALL (Multiplatform Testing on the NTB Mall Application)." [Online]. Available: <http://begawe.unram.ac.id/index.php/JBTI/>

[17] A. M. Ujung, M. Irwan, and P. Nasution, "Pentingnya Sistem Keamanan Database untuk melindungi data pribadi," *JISKA: Jurnal Sistem Informasi Dan Informatika*, vol. 1, no. 2, p. 44, 2023, [Online]. Available: <http://jurnal.unidha.ac.id/index.php/jteksis>

- [18]L. San and S. Gischa, "Pengertian Endcoding dan Decoding dalam Komunikasi," *kompas.com*. Accessed: Jan. 20, 2025. [Online]. Available: <https://www.kompas.com/skola/read/2022/09/27/143000269/pengertian-endcoding-dan-decoding-dalam-komunikasi>
- [19]R. N. Irawan, K. M. Hindrayani, and M. Idhom, "Penerapan Cross Validation sebagai Analisis Sentimen Pelayanan Publik

Kereta Api Lokal Daop 8 Menggunakan Metode Multinomial Naïve Bayes," *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 2, pp. 954–963, Apr. 2024, doi: 10.33379/gtech.v8i2.4117.