

Klasifikasi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma *K-Nearest Neighbor* (K-NN) Pada Data Akademik Perguruan Tinggi

¹Davina Rizky Efendi, ²Deci Irmayani, ³Volvo Sihombing

^{1,2,3}Manajemen Informatika, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

Email : 1davinarizkyharahap21@gmail.com, 2deacyirmayani@gmail.com,
3volvolumbantoran@gmail.com

Corresponding Author : davinarizkyharahap21@gmail.com

Abstract

Higher education is an important factor in scoring quality human resources, where one indicator of success is the graduation rate of students on time. This study aims to classify the graduation rate of students using the algorithm K-Nearest Neighbor (K-NN) based on academic data which includes GPA, number of credits, frequency of repetition of courses, and attendance. The results of the classification showed that 30% of students successfully graduated on time, while the rest had delays. With the k-NN approach, it is expected that this model can help universities in predicting student graduation more accurately and optimizing academic interventions to improve graduation efficiency.

Keywords : *K-Nearest Neighbor (K-NN), graduation rate, academic Data, Student classification, Higher Education.*

1. Pendahuluan

Pendidikan tinggi memegang peranan penting dalam mencetak sumber daya manusia yang berkualitas. Salah satu indikator keberhasilan pendidikan di perguruan tinggi adalah tingkat kelulusan mahasiswa dalam waktu yang ditentukan, yang mencerminkan kualitas proses akademik dan efektivitas layanan akademik. Dalam era digital, pemanfaatan teknologi informasi dan analisis data menjadi kunci dalam pengambilan keputusan di bidang pendidikan. Seiring dengan berkembangnya volume data akademik, penting bagi institusi untuk memanfaatkannya untuk meningkatkan manajemen akademik, salah satunya dalam memprediksi tingkat kelulusan mahasiswa. Klasifikasi tingkat kelulusan mahasiswa menjadi sangat relevan dalam upaya evaluasi dan intervensi akademik. Salah satu metode yang dapat digunakan untuk ini adalah algoritma *K-Nearest Neighbor* (K-NN), yang bekerja dengan mengklasifikasikan data berdasarkan kedekatannya dengan data yang sudah diketahui kelasnya. Algoritma ini efektif dalam memodelkan hubungan antara variabel akademik dan prediksi kelulusan mahasiswa. Namun, banyak perguruan tinggi yang belum maksimal dalam memanfaatkan data akademik untuk menganalisis tingkat kelulusan, sehingga intervensi yang tepat sering terlambat dilakukan. Penelitian ini bertujuan untuk mengisi celah tersebut dengan menerapkan algoritma K-NN pada data akademik mahasiswa untuk mengklasifikasikan tingkat kelulusan tepat waktu. Dengan menggunakan data seperti IPK semester, jumlah SKS, frekuensi pengulangan mata kuliah, dan indeks kehadiran, penelitian ini bertujuan menghasilkan model yang dapat memberikan prediksi akurat tentang kelulusan mahasiswa. Hasil klasifikasi ini diharapkan membantu institusi

pendidikan dalam meningkatkan tingkat kelulusan dan menciptakan budaya akademik yang lebih responsif berbasis data.

2. Landasan Teori

Tingkat Kelulusan

Tingkat kelulusan merupakan salah satu indikator utama dalam menilai keberhasilan proses pendidikan di perguruan tinggi. Tingkat kelulusan mengacu pada jumlah mahasiswa yang berhasil menyelesaikan studi dalam batas waktu yang telah ditetapkan, biasanya delapan semester untuk jenjang sarjana. Faktor-faktor yang memengaruhi kelulusan meliputi kemampuan akademik, motivasi belajar, manajemen waktu, serta dukungan dari lingkungan kampus. Perguruan tinggi secara umum menargetkan persentase kelulusan yang tinggi karena hal ini tidak hanya mencerminkan mutu penyelenggaraan pendidikan, tetapi juga menjadi acuan dalam evaluasi akreditasi institusi. Ketepatan waktu kelulusan menjadi perhatian karena berdampak pada efisiensi sumber daya, keberlanjutan studi, dan kesiapan lulusan dalam memasuki dunia kerja.

Mahasiswa

Mahasiswa merupakan individu yang sedang menempuh pendidikan tinggi di perguruan tinggi atau universitas, berada pada fase transisi dari remaja menuju dewasa yang ditandai dengan peningkatan tanggung jawab, kemandirian, serta pencarian jati diri. Mereka dihadapkan pada berbagai tuntutan akademik yang kompleks dan persaingan yang ketat, baik secara internal maupun eksternal. Selain menyelesaikan beban studi, mahasiswa juga dituntut untuk aktif dalam kegiatan organisasi, pengembangan diri, serta membangun jejaring sosial yang mendukung masa depan mereka. Kemampuan mengelola waktu, adaptasi terhadap sistem perkuliahan, serta konsistensi belajar menjadi faktor penentu keberhasilan seorang mahasiswa dalam menyelesaikan pendidikan tinggi tepat waktu.

Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) merupakan proses sistematis dalam mengekstraksi pengetahuan yang berguna dan tersembunyi dari kumpulan data yang besar. Proses ini terdiri dari beberapa tahapan utama, yaitu seleksi data, praproses data, transformasi data, data mining, dan evaluasi atau interpretasi hasil. Tujuan utama dari KDD adalah menemukan pola atau informasi baru yang sebelumnya tidak diketahui, namun memiliki nilai dalam pengambilan keputusan. Data mining menjadi inti dari proses KDD karena pada tahap inilah algoritma statistik atau machine learning diterapkan untuk menemukan pola, klasifikasi, atau asosiasi dari data yang telah dipersiapkan. Dengan perkembangan teknologi dan bertambahnya volume data yang tersedia, KDD menjadi pendekatan yang sangat penting dalam analisis data modern.

KDD sangat relevan dalam analisis data akademik mahasiswa, terutama dalam menggali informasi dari data yang tampak biasa namun menyimpan pola tertentu terkait tingkat kelulusan. Melalui proses KDD, data akademik seperti IPK, SKS, dan jumlah mata kuliah yang diulang dapat diproses dan ditransformasikan untuk kemudian diklasifikasikan menggunakan algoritma K-Nearest Neighbor (K-NN). Dengan pendekatan ini, institusi pendidikan dapat memperoleh pengetahuan yang lebih dalam tentang faktor-faktor yang memengaruhi kelulusan mahasiswa serta membuat prediksi yang lebih akurat terkait performa akademik mereka. Hasil akhir dari KDD tidak hanya

mendesripsikan kondisi saat ini, tetapi juga memberikan wawasan prediktif yang mendukung perbaikan strategi pembelajaran dan manajemen akademik.

Data Mining

Data mining merupakan proses penting dalam menemukan pola tersembunyi, hubungan, atau tren yang bermakna dari kumpulan data yang besar dan kompleks. Proses ini melibatkan serangkaian tahapan seperti pembersihan data, integrasi, seleksi, transformasi, dan penerapan teknik analitik seperti klasifikasi, klustering, regresi, atau asosiasi. Tujuan utama data mining adalah untuk mengubah data mentah menjadi informasi yang berguna dalam mendukung pengambilan keputusan yang lebih tepat dan strategis. Dalam implementasinya, data mining banyak digunakan di berbagai bidang seperti bisnis, kesehatan, keuangan, hingga pendidikan, karena kemampuannya dalam mengolah data historis menjadi pengetahuan baru yang bersifat prediktif maupun deskriptif.

Metode *K-Nearest Neighbor (KNN)*

Algoritma *K-Nearest Neighbor (K-NN)* merupakan salah satu metode klasifikasi dalam machine learning yang bersifat non-parametrik dan termasuk ke dalam kategori supervised learning. Prinsip kerja K-NN cukup sederhana namun efektif, yaitu dengan mengklasifikasikan suatu data berdasarkan mayoritas label dari sejumlah tetangga terdekat (nearest neighbors) yang dihitung menggunakan jarak, seperti Euclidean distance. Nilai *K* dalam K-NN merujuk pada jumlah tetangga yang akan dijadikan acuan untuk menentukan kelas suatu data baru. Semakin kecil nilai *K*, maka model akan lebih sensitif terhadap data sekitar dan berisiko overfitting; sebaliknya, nilai *K* yang terlalu besar dapat mengaburkan batas antar kelas. K-NN tidak memerlukan proses pelatihan model secara eksplisit karena klasifikasi dilakukan langsung saat proses prediksi, sehingga metode ini dikenal sebagai *lazy learner*.

Model Klasifikasi

Model klasifikasi adalah salah satu metode dalam teknik data mining yang digunakan untuk memetakan data ke dalam kelas-kelas tertentu berdasarkan atribut atau variabel yang dimiliki. Proses klasifikasi melibatkan pembelajaran dari data latih (training data) yang telah diberi label, untuk kemudian membentuk suatu model yang mampu memprediksi kelas dari data baru yang belum dikenal. Model ini umumnya digunakan dalam pengambilan keputusan otomatis, sistem rekomendasi, analisis perilaku pelanggan, diagnosis medis, hingga prediksi risiko. Metode klasifikasi dapat dibangun menggunakan berbagai algoritma, seperti Decision Tree, Naive Bayes, *k-Nearest Neighbor*, dan Support Vector Machine, tergantung pada jenis data dan tujuan analisis.

3. Metode Penelitian

Metode Penelitian

Metode penelitian yang digunakan dalam studi ini adalah pendekatan kuantitatif dengan teknik klasifikasi menggunakan algoritma *K-Nearest Neighbor (K-NN)*. Data yang dianalisis merupakan data akademik mahasiswa yang terdiri dari beberapa variabel seperti Indeks Prestasi Kumulatif (IPK), jumlah SKS yang telah ditempuh, frekuensi pengulangan mata kuliah, dan lama studi. Data ini kemudian diproses dan dibagi ke dalam dua kategori kelas, yakni "lulus tepat waktu" dan "tidak lulus tepat waktu". Penelitian ini

melalui tahapan mulai dari pengumpulan data, pembersihan data (*data cleaning*), normalisasi, pembagian data latih dan data uji, serta penerapan algoritma K-NN untuk klasifikasi.

Untuk proses klasifikasi, digunakan teknik validasi silang (*cross-validation*) dan pengukuran performa model dengan metrik evaluasi seperti akurasi, presisi, recall, dan *F1-score*. Pemilihan nilai k pada algoritma K-NN juga diuji secara bertahap untuk mendapatkan performa terbaik. Seluruh proses pengolahan data dan implementasi algoritma dilakukan menggunakan perangkat lunak pendukung seperti Python atau software data mining lainnya yang relevan. Dengan pendekatan ini, diharapkan hasil penelitian mampu memberikan gambaran sejauh mana K-NN efektif dalam mengklasifikasikan tingkat kelulusan mahasiswa berdasarkan data akademik yang tersedia.

Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui penyebaran kuesioner kepada mahasiswa Universitas Labuhanbatu (ULB) sebagai responden. Kuesioner disusun secara terstruktur dan berisi pertanyaan-pertanyaan yang berkaitan dengan data akademik mahasiswa seperti Indeks Prestasi Kumulatif (IPK), jumlah SKS yang telah ditempuh, jumlah mata kuliah yang diulang, kehadiran, serta lama studi. Kuesioner ini dirancang untuk menggali informasi yang relevan dan berkaitan langsung dengan variabel-variabel yang memengaruhi tingkat kelulusan mahasiswa. Penyebaran kuesioner dilakukan secara langsung maupun melalui media online agar menjangkau lebih banyak responden secara efisien.

Data yang diperoleh dari kuesioner kemudian dikompilasi dan divalidasi untuk memastikan kelengkapan dan keakuratannya sebelum dilakukan proses analisis. Validasi dilakukan dengan cara mencocokkan isian responden dengan data akademik yang tersedia di bagian administrasi kampus, apabila memungkinkan, guna menghindari kesalahan input atau data yang tidak konsisten. Setelah data dinyatakan valid, data tersebut kemudian digunakan dalam proses klasifikasi menggunakan algoritma *K-Nearest Neighbor* (K-NN) untuk mengidentifikasi pola-pola yang dapat mengindikasikan kelulusan mahasiswa, baik yang lulus tepat waktu maupun tidak. Teknik ini dinilai efektif untuk mendapatkan data primer yang spesifik sesuai kebutuhan penelitian.

Teknik Analisis Data

Teknik analisis data dalam penelitian ini menggunakan aplikasi *Orange Data Mining*, sebuah perangkat lunak open-source yang berbasis visual programming dan sangat cocok digunakan dalam pemodelan klasifikasi, regresi, dan clustering. Orange menyediakan antarmuka yang intuitif, memungkinkan pengguna untuk melakukan proses data mining tanpa harus menulis kode pemrograman secara manual. Dalam penelitian ini, data yang telah dikumpulkan dari hasil kuesioner diimpor ke dalam Orange, kemudian melalui tahapan preprocessing seperti normalisasi data, penghapusan atribut yang tidak relevan, dan pemilihan fitur yang berpengaruh terhadap tingkat kelulusan mahasiswa. Setelah data siap, dilakukan proses pelatihan dan pengujian model dengan metode *K-Nearest Neighbor* (K-NN).

4. Hasil Dan Pembahasan

Pengumpulan Data

Pengumpulan data pada tahap awal penelitian ini dilakukan dengan merancang instrumen berupa kuesioner tertutup yang disusun berdasarkan indikator-indikator yang relevan terhadap kelulusan mahasiswa di lingkungan perguruan tinggi. Kuesioner ini mencakup sejumlah variabel utama seperti jumlah SKS yang telah ditempuh, rata-rata nilai akademik, jumlah kehadiran dalam perkuliahan, lama studi, dan IPK mahasiswa. Setiap pertanyaan disajikan dalam bentuk pilihan jawaban terstruktur agar memudahkan pengkodean dan pengolahan data secara sistematis. Sebelum penyebaran, kuesioner diuji coba dalam skala kecil untuk memastikan bahwa isi pertanyaan mudah dipahami dan sesuai dengan konteks penelitian. Setelahnya, kuesioner disebarluaskan secara daring melalui platform media sosial dan grup komunikasi mahasiswa Universitas Labuhanbatu, serta secara langsung kepada mahasiswa tingkat akhir. Pendekatan ini dipilih untuk memperoleh data yang beragam dan akurat sebagai dasar dalam membangun model klasifikasi tingkat kelulusan menggunakan algoritma *K-Nearest Neighbor*.

Tabel 1. Data Penelitian

Nama	Jumlah SKS	Rata-Rata Nilai	Jumlah Kehadiran (%)	Lama Studi (Semester)	IPK
M1	123	2.95	88	8	3.91
M2	137	3.41	100	11	3.3
....
M100	142	2.52	71	9	3.3

Pada tabel di atas merupakan data yang digunakan dalam penelitian ini yang terdiri dari 100 entri responden yang mencerminkan data akademik mahasiswa Universitas Labuhanbatu. Setiap entri mencakup enam atribut utama, yaitu Nama Mahasiswa, Jumlah SKS, Rata-Rata Nilai, Jumlah Kehadiran, Lama Studi, dan IPK. Atribut-atribut ini dipilih karena dianggap sebagai indikator penting yang memengaruhi tingkat kelulusan mahasiswa dalam proses studi di perguruan tinggi. Data ini kemudian diklasifikasikan ke dalam label “Lulus” atau “Tidak Lulus” berdasarkan kriteria akademik yang telah ditentukan, seperti nilai ambang batas IPK dan ketepatan waktu studi. Kombinasi nilai dari masing-masing atribut digunakan dalam proses klasifikasi dengan algoritma *K-Nearest Neighbor* (K-NN), sehingga memungkinkan terbentuknya pola yang sistematis dalam memprediksi status kelulusan mahasiswa secara lebih objektif dan terukur.

Pembersihan Data

Pembersihan data pada penelitian ini merupakan tahap krusial yang dilakukan untuk memastikan bahwa data akademik yang digunakan benar-benar layak dan berkualitas sebelum masuk ke tahap analisis. Proses ini mencakup peninjauan ulang terhadap seluruh entri mahasiswa untuk mendeteksi dan menghapus data duplikat, seperti nama mahasiswa yang tercatat lebih dari satu kali atau entri yang memiliki atribut identik secara mencurigakan. Selain itu, data yang tidak lengkap—seperti kolom IPK kosong, jumlah SKS yang tidak masuk akal, atau kehadiran melebihi batas maksimal—juga dieliminasi guna menjaga integritas dataset. Pembersihan ini bertujuan agar hanya data yang valid, akurat, dan konsisten yang digunakan dalam proses klasifikasi menggunakan algoritma *K-Nearest Neighbor* (K-NN).

Pembagian Data

Pada penelitian ini, pembagian data dilakukan ke dalam dua subset utama, yaitu data training dan data testing. Dari total 100 data mahasiswa yang tersedia, sebanyak 20 data digunakan sebagai data training untuk membangun dan melatih model klasifikasi menggunakan algoritma *K-Nearest Neighbor* (K-NN). Tujuan dari pelatihan ini adalah agar model dapat mengenali pola dan hubungan antaratribut seperti jumlah SKS, rata-rata nilai, jumlah kehadiran, lama studi, dan IPK dalam menentukan tingkat kelulusan mahasiswa. Sementara itu, 80 data sisanya digunakan sebagai data testing untuk menguji sejauh mana model yang telah dilatih mampu mengklasifikasikan data baru secara akurat. Strategi pembagian ini sangat penting untuk menghindari *overfitting* dan memastikan bahwa model yang dikembangkan memiliki performa yang baik dalam menggeneralisasi data akademik mahasiswa lainnya.

Tabel 2. Data Training

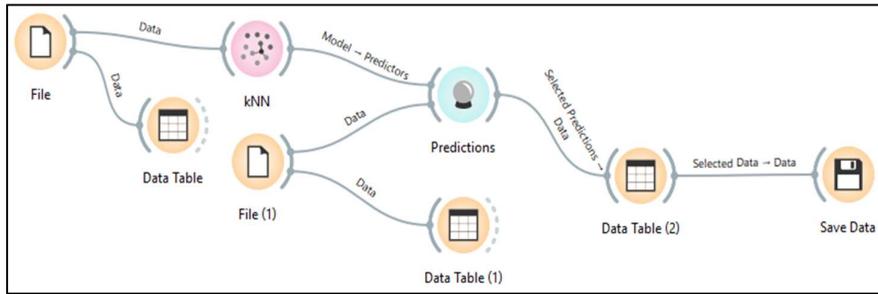
Nama	Jumlah SKS	Rata-Rata Nilai	Jumlah Kehadiran (%)	Lama Studi (Semester)	IPK	Keterangan
M1	124	2.66	95	8	3.64	Tepat Waktu
M2	135	3.25	70	9	3.24	Tidak Tepat Waktu
M3	139	3.01	77	11	3.05	Tidak Tepat Waktu
...
M20	142	2.52	71	9	3.3	Tidak Tepat Waktu

Tabel 3. Data Testing

Nama	Jumlah SKS	Rata-Rata Nilai	Jumlah Kehadiran (%)	Lama Studi (Semester)	IPK
M1	123	2.95	88	8	3.91
M2	137	3.41	100	11	3.3
M3	122	2.56	74	7	2.84
...
M80	138	3.75	87	11	2.51

Perancangan Model Klasifikasi

Perancangan model klasifikasi pada penelitian ini bertujuan untuk mengelompokkan data ke dalam kategori "Lulus" dan "Tidak Lulus" berdasarkan atribut-atribut akademik mahasiswa, yaitu jumlah SKS, rata-rata nilai, jumlah kehadiran, lama studi, dan IPK. Proses ini dimulai dengan menetapkan label target sebagai kategori kelulusan, lalu memilih fitur-fitur relevan yang memengaruhi keberhasilan studi mahasiswa. Algoritma klasifikasi *K-Nearest Neighbor* (K-NN) digunakan dalam proses ini karena kemampuannya dalam mengidentifikasi pola berdasarkan kedekatan antar data. Model kemudian dibentuk menggunakan data training agar mampu mengenali kecenderungan atau pola kelulusan berdasarkan kombinasi nilai dari atribut-atribut tersebut, sehingga dapat digunakan untuk memprediksi kelulusan mahasiswa baru secara akurat.



Gambar 1. Perancangan Model Klasifikasi

Gambar di atas merupakan representasi visual dari alur kerja dalam aplikasi Orange yang digunakan untuk merancang model klasifikasi menggunakan metode *K-Nearest Neighbor* (K-NN). Alur dimulai dari widget "File", yang berfungsi untuk memuat data training dari file eksternal, yang berisi data mahasiswa beserta atribut-atribut seperti jumlah SKS, rata-rata nilai, jumlah kehadiran, lama studi, dan IPK. Data ini kemudian diteruskan ke widget "Data Table" untuk ditampilkan dan ditinjau secara manual, memastikan data telah dimuat dengan benar dan siap untuk diproses lebih lanjut.

Setelah itu, data diteruskan ke widget "Select Columns" untuk menetapkan atribut-atribut sebagai fitur serta menentukan label target klasifikasi, yaitu kategori kelulusan ("Lulus" atau "Tidak Lulus"). Widget "kNN" digunakan untuk melatih model klasifikasi berdasarkan data training yang telah disiapkan. Selanjutnya, widget "File (1)" digunakan untuk memuat data testing yang akan diprediksi. Data ini dikoneksikan ke widget "Predictions", yang menggabungkan data testing dengan model yang sudah dilatih untuk menghasilkan prediksi kelulusan. Hasil prediksi ditampilkan dalam widget "Data Table (2)", yang menyajikan output secara rinci. Dengan alur ini, seluruh proses dari pemuatan data hingga prediksi dapat dilacak dan dianalisis secara sistematis untuk mendukung hasil penelitian.

Tabel 3. Hasil Klasifikasi

Nama	Jumlah SKS	Rata-Rata Nilai	Jumlah Kehadiran (%)	Lama Studi (Semester)	IPK	Keterangan
M1	123	2.95	88	8	3.91	Tidak Tepat Waktu
M2	137	3.41	100	11	3.3	Tidak Tepat Waktu
M3	122	2.56	74	7	2.84	Tidak Tepat Waktu
M4	138	2.81	97	10	3.33	Tidak Tepat Waktu
M5	120	2.7	81	11	3.92	Tidak Tepat Waktu
M6	132	3.36	94	9	2.7	Tidak Tepat Waktu
M7	134	2.63	77	7	3.96	Tidak Tepat Waktu
M8	143	3.46	84	7	3.12	Tidak Tepat Waktu
M9	126	3.18	82	10	3.92	Tidak Tepat Waktu
M10	123	3.77	73	7	3.86	Tidak Tepat Waktu
M11	122	2.74	87	10	2.98	Tidak Tepat Waktu
M12	136	3.23	79	8	3.79	Tidak Tepat Waktu
M13	129	2.85	100	10	3.93	Tidak Tepat Waktu
M14	121	3.72	100	8	3.99	Tepat Waktu

M15	139	3.72	84	10	3.96	Tidak Tepat Waktu
M16	133	3.42	71	8	3.39	Tidak Tepat Waktu
M17	127	3.47	99	7	3.77	Tepat Waktu
M18	141	3.13	70	9	2.93	Tidak Tepat Waktu
M19	132	2.72	84	11	2.85	Tidak Tepat Waktu
M20	142	3.09	88	12	3.36	Tepat Waktu
M21	144	2.99	79	10	2.62	Tidak Tepat Waktu
M22	142	2.84	98	7	2.88	Tepat Waktu
M23	123	2.56	73	11	3.55	Tidak Tepat Waktu
M24	130	3.71	80	10	2.9	Tidak Tepat Waktu
M25	121	3.41	95	9	3.26	Tepat Waktu
M26	128	3.63	83	12	3.56	Tidak Tepat Waktu
M27	139	2.57	95	8	2.56	Tepat Waktu
M28	123	3.43	85	9	3.19	Tidak Tepat Waktu
M29	121	2.94	78	7	3.6	Tidak Tepat Waktu
M30	142	3.66	82	11	3.36	Tidak Tepat Waktu
M31	132	3.22	98	8	3.76	Tidak Tepat Waktu
M32	142	2.72	87	9	3.66	Tepat Waktu
M33	121	2.65	86	10	2.83	Tidak Tepat Waktu
M34	137	3.34	99	10	3.23	Tidak Tepat Waktu
M35	142	3.76	84	12	3.83	Tidak Tepat Waktu
M36	133	3.05	91	9	3.09	Tidak Tepat Waktu
M37	120	2.56	93	7	2.67	Tepat Waktu
M38	140	2.72	79	9	2.82	Tidak Tepat Waktu
M39	139	3.64	73	10	3.16	Tidak Tepat Waktu
M40	126	3.19	70	12	3.79	Tidak Tepat Waktu
M41	127	2.88	77	8	2.87	Tidak Tepat Waktu
M42	139	2.62	92	12	2.86	Tepat Waktu
M43	144	3.32	99	7	2.92	Tepat Waktu
M44	126	3.42	82	7	2.97	Tidak Tepat Waktu
M45	135	3.55	78	12	2.96	Tidak Tepat Waktu
M46	121	3.47	82	11	3.42	Tidak Tepat Waktu
M47	136	3.56	75	8	3.81	Tidak Tepat Waktu
M48	131	3.25	75	12	3.05	Tidak Tepat Waktu
M49	124	3.43	90	10	3.2	Tidak Tepat Waktu
M50	132	2.59	74	8	3.51	Tidak Tepat Waktu
M51	131	2.75	73	10	3.95	Tidak Tepat Waktu
M52	131	3.57	83	8	3.13	Tidak Tepat Waktu
M53	132	3.23	91	10	3.58	Tidak Tepat Waktu
M54	126	3.04	92	12	3.27	Tepat Waktu
M55	142	3.03	83	7	3.66	Tidak Tepat Waktu

M56	133	2.65	84	7	2.82	Tidak Tepat Waktu
M57	130	3.26	81	10	3.29	Tidak Tepat Waktu
M58	120	3.3	76	7	3.41	Tidak Tepat Waktu
M59	135	3.15	94	8	3.18	Tepat Waktu
M60	124	2.86	79	12	3.19	Tidak Tepat Waktu
M61	131	2.94	77	8	3.8	Tidak Tepat Waktu
M62	133	3.76	84	9	3.48	Tidak Tepat Waktu
M63	140	2.65	98	7	3.43	Tepat Waktu
M64	134	3.7	92	10	2.67	Tidak Tepat Waktu
M65	142	2.93	96	9	3.67	Tepat Waktu
M66	128	3.03	94	9	3.22	Tepat Waktu
M67	130	3.6	98	10	3.45	Tidak Tepat Waktu
M68	135	3.19	75	10	2.62	Tidak Tepat Waktu
M69	136	3.37	96	7	3.67	Tepat Waktu
M70	120	3.03	95	10	3.02	Tepat Waktu
M71	144	3.15	79	9	3.99	Tidak Tepat Waktu
M72	127	3.65	98	7	2.51	Tepat Waktu
M73	129	2.78	83	10	2.63	Tidak Tepat Waktu
M74	132	2.72	94	9	3.67	Tidak Tepat Waktu
M75	141	3.77	93	10	3.03	Tepat Waktu
M76	127	2.53	81	12	3.77	Tidak Tepat Waktu
M77	123	3.7	70	10	2.94	Tidak Tepat Waktu
M78	144	2.79	72	8	3.83	Tidak Tepat Waktu
M79	144	3.11	89	11	3.5	Tepat Waktu
M80	138	3.75	87	11	2.51	Tepat Waktu

Berdasarkan hasil klasifikasi terhadap 80 data mahasiswa yang telah dianalisis, diperoleh informasi bahwa sebanyak 24 mahasiswa (30%) berhasil menyelesaikan studinya tepat waktu, sementara sisanya sebanyak 56 mahasiswa (70%) tidak dapat menyelesaikan studi sesuai dengan waktu ideal yang telah ditentukan, yaitu 8 semester atau 4 tahun masa studi. Jumlah ini mengindikasikan bahwa mayoritas mahasiswa mengalami keterlambatan dalam proses penyelesaian studi mereka. Tingginya proporsi mahasiswa yang tidak lulus tepat waktu menunjukkan bahwa masih terdapat berbagai kendala yang memengaruhi keberhasilan mahasiswa dalam menyelesaikan pendidikan tepat waktu, baik dari aspek akademik maupun non-akademik. Fakta ini menjadi penting untuk ditelusuri lebih lanjut guna mengetahui pola-pola yang dapat menjadi acuan dalam menyusun strategi pembinaan akademik yang lebih efektif ke depannya.

Apabila ditinjau lebih dalam dari variabel-variabel yang memengaruhi klasifikasi tersebut, seperti jumlah SKS, rata-rata nilai mata kuliah, jumlah kehadiran perkuliahan, lama studi, serta IPK akhir, dapat dilihat bahwa mahasiswa yang lulus tepat waktu cenderung memiliki kecenderungan akademik yang lebih stabil. Mahasiswa dalam kategori ini rata-rata memiliki jumlah SKS yang konsisten dengan kurikulum standar, nilai rata-rata mata kuliah di atas 70, tingkat kehadiran perkuliahan yang tinggi mendekati atau mencapai 100%, serta IPK yang berada pada kisaran 3.0 ke atas. Selain itu, mereka juga menuntaskan seluruh kewajiban akademik, termasuk tugas akhir dan skripsi, tepat

pada waktunya tanpa pengulangan mata kuliah yang signifikan. Pola ini menandakan bahwa kedisiplinan dalam mengikuti kegiatan akademik, kemampuan mengelola waktu, serta komitmen belajar yang tinggi merupakan faktor utama yang mendukung ketepatan waktu kelulusan mahasiswa.

Sebaliknya, mahasiswa yang tidak lulus tepat waktu umumnya menunjukkan kecenderungan yang berbeda dalam hal perilaku akademik. Mereka lebih sering memiliki jumlah SKS yang cenderung lebih tinggi karena harus mengulang beberapa mata kuliah, rata-rata nilai yang berada di bawah 70, serta jumlah kehadiran yang bervariasi dan bahkan beberapa di antaranya kurang dari 75%. Selain itu, lama studi mahasiswa dalam kelompok ini rata-rata melebihi delapan semester, bahkan ada yang mencapai hingga sepuluh semester. Meskipun terdapat beberapa mahasiswa dalam kategori ini yang memiliki IPK tinggi, hal tersebut tidak serta-merta menjamin kelulusan tepat waktu. Ini mengindikasikan bahwa nilai IPK bukanlah satu-satunya indikator keberhasilan studi, melainkan harus dilihat secara menyeluruh dari aspek lain seperti efisiensi pengambilan mata kuliah, kesiapan menyusun tugas akhir, serta kemampuan menyelesaikan masalah pribadi atau administratif yang mungkin muncul selama masa studi.

Temuan ini memberikan pemahaman bahwa dalam konteks akademik, keberhasilan mahasiswa dalam menyelesaikan studi secara tepat waktu sangat dipengaruhi oleh kombinasi berbagai faktor. Disiplin dalam menghadiri perkuliahan, strategi belajar yang baik, konsistensi dalam pengambilan beban SKS, serta manajemen waktu yang efektif menjadi kunci penting. Oleh karena itu, untuk menekan angka keterlambatan kelulusan, pihak institusi pendidikan tinggi, khususnya program studi, perlu meningkatkan peran pembimbing akademik dan sistem monitoring studi mahasiswa. Bimbingan akademik yang dilakukan secara berkala dan tepat sasaran mampu membantu mahasiswa dalam merancang strategi akademik yang optimal sesuai dengan kemampuan dan kondisi masing-masing. Dengan demikian, hasil klasifikasi ini tidak hanya menggambarkan kondisi realitas akademik mahasiswa saat ini, tetapi juga menjadi dasar penting dalam merumuskan kebijakan akademik yang lebih adaptif dan solutif di masa yang akan datang.

5. Kesimpulan

Kesimpulan

Berdasarkan hasil klasifikasi terhadap 80 data mahasiswa, ditemukan bahwa hanya 24 orang atau sebesar 30% mahasiswa yang berhasil menyelesaikan studinya tepat waktu, sementara 56 orang atau 70% lainnya mengalami keterlambatan dalam menyelesaikan studi. Hasil ini menunjukkan bahwa mayoritas mahasiswa masih menghadapi berbagai hambatan yang menyebabkan studi mereka melewati batas waktu ideal. Analisis lebih lanjut terhadap variabel-variabel seperti jumlah SKS, rata-rata nilai mata kuliah, tingkat kehadiran, lama studi, dan IPK menunjukkan bahwa mahasiswa yang lulus tepat waktu cenderung memiliki performa akademik yang lebih stabil dan disiplin dalam menjalani proses perkuliahan. Di sisi lain, mahasiswa yang tidak tepat waktu cenderung memiliki nilai akademik yang fluktuatif, kehadiran yang kurang maksimal, serta pengambilan beban studi yang tidak efisien. Oleh karena itu, hasil klasifikasi ini menggambarkan bahwa ketepatan waktu kelulusan tidak hanya dipengaruhi oleh kecerdasan akademik semata, melainkan juga sangat bergantung pada faktor kedisiplinan, manajemen waktu, dan kemampuan menyusun strategi belajar yang tepat sejak awal masa studi.

Saran

Sebagai langkah strategis untuk meningkatkan jumlah mahasiswa yang lulus tepat waktu, pihak institusi pendidikan tinggi, khususnya program studi dan dosen pembimbing akademik, disarankan untuk lebih intensif dalam melakukan pendampingan akademik secara berkelanjutan. Pemantauan dan evaluasi kemajuan studi mahasiswa perlu dilakukan secara berkala agar hambatan-hambatan yang muncul dapat segera diidentifikasi dan ditangani. Selain itu, penting bagi mahasiswa untuk dibekali dengan pelatihan manajemen waktu, strategi belajar efektif, serta motivasi internal yang kuat agar mereka dapat menyelesaikan setiap tahapan akademik secara sistematis dan efisien. Diharapkan, dengan adanya pendekatan pembinaan yang lebih terstruktur dan dukungan yang berkelanjutan dari lingkungan kampus, angka keterlambatan kelulusan dapat ditekan secara signifikan dan menciptakan lulusan yang tidak hanya berkualitas, tetapi juga tepat waktu dalam menyelesaikan pendidikannya.

6. Daftar Pustaka

- Agarwal, R., Ardy Garini, D., Dhamayanti, D., Yulianti, E., Fahmi Kamal, Widi Winarso, Lia Mardiani, Dewi, A. P., Sudarwati, W., Sidaria, S., Murni, D., Khairina, I., Nelwati, N., Manaransyah, G., Rahman, A., Rachmawaty, I. K., Tamam. Sanusi, U., Andriani, D. S., Saputra, A., ... Edi Gunawan, Andika Triansyah, F. P. H. (2022). Pengaruh Fasilitas, Harga Dan Lokasi Terhadap Kepuasan Siswa Pada Sekolah Menengah Kejuruan Muhammadiyah 1 Kalianda. *Jurnal Ilmiah Akuntansi Dan Manajemen*, 5(1), 1. <https://doi.org/10.52157/me.v12i2.205>
- Aldi, F., Nozomi, I., Sentosa, R. B., & Junaidi, A. (2023). Machine Learning to Identify Monkey Pox Disease. *Sinkron*, 8(3), 1335–1347. <https://doi.org/10.33395/sinkron.v8i3.12524>
- Amalia, R. (2020). Penerapan Data Mining untuk Memprediksi Hasil Kelulusan Siswa Menggunakan Metode Naïve Bayes. *Juisi*, 06(01), 33–42.
- Andrianto, R., & Irawan, F. (2023). Implementasi Metode Regresi Linear Berganda Pada Sistem Prediksi Jumlah Tonase Kelapa Sawit di PT . Paluta Inti Sawit. *Jurnal Pendidikan Tambusai*, 7(1), 2926–2934.
- Atalya Angelus Leza, M., Widya Utami, N., & Anugrah Cahya Dewi, P. (2024). Prediksi Prestasi Siswa Smas Katolik Santo Yoseph Denpasar Berdasarkan Kedisiplinan Dan Tingkat Ekonomi Orang Tua Menggunakan Metode Knowledge Discovery in Database Dan Algoritma Regresi Linier Berganda. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 373–379. <https://doi.org/10.36040/jati.v8i1.8754>
- Badaru, B. S. (2022). Journal of Lex Theory (JLT). *Journal of Lex Theory (JLT)*, 1(2), 122–135.
- Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Hartati, T., Nurdiawan, O., & Wiyandi, E. (2021). Analisis Dan Penerapan Algoritma K-Means Dalam Strategi Promosi Kampus Akademi Maritim Suaka Bahari. *Jurnal Sains Teknologi Transportasi Maritim*, 3(1), 1–7. <https://doi.org/10.51578/j.sitektransmar.v3i1.30>
- Hssina, B., Grota, A., & Erritali, M. (2021). Recommendation system using the k-nearest

- neighbors and singular value decomposition algorithms. *International Journal of Electrical and Computer Engineering*, 11(6), 5541–5548. <https://doi.org/10.11591/ijece.v11i6.pp5541-5548>
- Indah Lestari, Y., & Defit, S. (2021). *Jurnal Informatika Ekonomi Bisnis Prediksi Tingkat Kepuasan Pelayanan Online Menggunakan Metode Algoritma C.45*. 3, 148–154. <https://doi.org/10.37034/infv3i3.104>
- Maruli Tua Silaen. (2023). Klasifikasi Karakteristik Kepribadian Siswa Berdasarkan the Big Five Personality Dengan Menggunakan Metode K- Nearest Neighbor (Knn). *Jurnal Informatika Dan Rekayasa Elektronik*, 6(1), 121–129. <https://doi.org/10.36595/jire.v6i1.860>
- Muhadat, I. S. (2021). Kasgot Sebagai Alternatif Pupuk Organik Padat Pada Tanaman Sawit (*Brassica juncea* L) Dengan Metode Vertikular. *Industry and Higher Education*, 3(1), 1689–1699.
- Muhammad, L. J., Islam, M. M., Usman, S. S., & Ayon, S. I. (2020). Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery. *SN Computer Science*, 1(4), 1–7. <https://doi.org/10.1007/s42979-020-00216-w>
- Palacios, C. A., Reyes-Suárez, J. A., Bearzotti, L. A., Leiva, V., & Marchant, C. (2021). Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in chile. *Entropy*, 23(4), 1–23. <https://doi.org/10.3390/e23040485>
- Punkastyo, D. A., Septian, F., & Syaripudin, A. (2024). Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Prediksi Kelulusan Siswa. *Journal of System and Computer Engineering (JSCE)*, 5(1), 24–35. <https://doi.org/10.61628/jsce.v5i1.1073>
- Purwati, N., Pedliyansah, Y., Kurniawan, H., Karnila, S., & Herwanto, R. (2023). Komparasi Metode Apriori dan FP-Growth Data Mining Untuk Mengetahui Pola Penjualan. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(2), 155–161. <https://doi.org/10.30591/jpit.v8i2.4876>
- Rachman, R., & Hunaifi, N. (2020). Penerapan Metode Algoritma Apriori dan FP-Tree Pada Penentuan Pola Pembelian Obat. *Paradigma - Jurnal Komputer Dan Informatika*, 22(2), 175–182. <https://doi.org/10.31294/p.v22i2.8258>
- Samudi, S., Widodo, S., & Brawijaya, H. (2020). The K-Medoids Clustering Method for Learning Applications during the COVID-19 Pandemic. *Sinkron*, 5(1), 116. <https://doi.org/10.33395/sinkron.v5i1.10649>
- Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine. *Sinkron*, 8(3), 1366–1378. <https://doi.org/10.33395/sinkron.v8i3.12447>
- Sarimole, F. M., & Rosiana, A. (2022). Classification of Maturity Levels in Areca Fruit Based on Hsv Image Using the Knn Method. *Journal of Applied Engineering and Technological Science*, 4(1), 64–73. <https://doi.org/10.37385/jaets.v4i1.951>
- Setiawan, A., Rabi, A., & Gumilang, Y. S. A. (2024). Pengolahan Citra untuk Sortir Buah Stroberi Berdasarkan Kematangan Menggunakan Algoritma K-Nearst Neighbors (KNN). *Blend Sains Jurnal Teknik*, 2(4), 322–328. <https://doi.org/10.56211/blendsains.v2i4.551>
- Sriwinarti, N. K., & Juniarti, P. (2021). *Analisis Metode K-Nearest Neighbors (K-NN) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa (Analysis of K-Nearest Neighbors (K-NN) and Naive Bayes Methods in Predicting Student*

Graduation). 3(2), 106–112.

- Wahyudi, A., Ovelia Tampubolon, S., Afrilia Putri, N., Ghassa, A., Rasywir, E., & Kisbianty, D. (2022). Penerapan Data Mining Algoritma Naive Bayes Classifier Untuk Mengetahui Minat Beli Pelanggan Terhadap INDIHOME. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(2), 240–247. <https://doi.org/10.33998/jakakom.2022.2.2.111>
- Xu, Y., Zheng, X., Li, Y., Ye, X., Cheng, H., Wang, H., & Lyu, J. (2023). Exploring patient medication adherence and data mining methods in clinical big data: A contemporary review. In *Journal of Evidence-Based Medicine* (Vol. 16, Issue 3). <https://doi.org/10.1111/jebm.12548>