

Analisis Faktor-Faktor yang Mempengaruhi Tingkat Kelulusan Siswa Menggunakan Algoritma KNN

Vitasari Sianipar, Deci Irmayani*, Budianto Bangun

Fakultas Sains dan Teknologi, Program Studi Sistem Informasi, Universitas Labuhanbatu, Rantauprapat, Indonesia

Email: ¹vitasianipar5@gmail.com, ^{2,*}deacyirmayani@gmail.com, ³budiantobangun44@gmail.com

Email Penulis Korespondensi: deacyirmayani@gmail.com

Submitted: 19/05/2025; Accepted: 25/06/2025; Published: 25/06/2025

Abstrak—Tingkat kelulusan siswa dipengaruhi oleh berbagai faktor akademik dan non-akademik, sehingga diperlukan metode analisis yang dapat mengelompokkan siswa berdasarkan kemungkinan kelulusannya. Penelitian ini menggunakan algoritma K-Nearest Neighbors (KNN) untuk menganalisis faktor-faktor yang mempengaruhi kelulusan siswa di SD Negeri 112269 Padang Lais. Algoritma KNN bekerja dengan menghitung jarak Euclidean antara data siswa yang diuji dengan data siswa lainnya, kemudian menentukan status kelulusan berdasarkan mayoritas dari K tetangga terdekat. Hasil penelitian menunjukkan bahwa penggunaan K=5 menghasilkan klasifikasi yang akurat dengan tingkat akurasi 100%, di mana siswa dengan nilai jarak terkecil terhadap siswa yang telah lulus memiliki peluang lebih besar untuk lulus. Kontribusi penelitian ini adalah membuktikan bahwa metode KNN dapat digunakan sebagai alat bantu dalam prediksi kelulusan siswa serta memberikan wawasan dalam pemanfaatan algoritma klasifikasi untuk mendukung pengambilan keputusan di bidang pendidikan. Penelitian selanjutnya dapat mengembangkan model dengan menambahkan variabel yang lebih beragam serta menguji model pada dataset yang lebih luas guna meningkatkan generalisasi hasil prediksi.

Kata Kunci: Tingkat Kelulusan; K-Nearest Neighbors; Klasifikasi; Data Mining; Prediksi Siswa

Abstract—Student graduation rates are influenced by various academic and non-academic factors, making it necessary to develop analytical methods to classify students based on their likelihood of graduation. This study applies the K-Nearest Neighbors (KNN) algorithm to analyze the factors affecting student graduation at SD Negeri 112269 Padang Lais. The KNN algorithm works by calculating the Euclidean distance between the tested student data and other student data, then determining the graduation status based on the majority of the K nearest neighbors. The results indicate that using K=5 produces highly accurate classifications with an accuracy rate of 100%, where students with the smallest distance to those who have graduated are more likely to pass. The contribution of this study is to demonstrate that the KNN method can serve as a decision-support tool for predicting student graduation and provide insights into the use of classification algorithms in educational decision-making. Future research can enhance the model by incorporating more diverse variables and testing it on larger datasets to improve prediction generalization.

Keywords: Graduation Rate; K-Nearest Neighbors; Classification; Data Mining; Student Prediction

1. PENDAHULUAN

Pendidikan dasar memiliki peran yang sangat penting dalam membentuk fondasi ilmu pengetahuan dan keterampilan siswa. Keberhasilan sekolah dasar dalam mencetak lulusan yang kompeten sering kali diukur melalui tingkat kelulusan siswa. Semakin tinggi tingkat kelulusan, semakin baik citra dan reputasi suatu sekolah. Namun, di banyak sekolah dasar, rendahnya tingkat kelulusan siswa menjadi salah satu permasalahan yang cukup sering dihadapi. Fenomena ini dipengaruhi oleh berbagai faktor, seperti kemampuan akademik siswa, motivasi belajar, dukungan orang tua, hingga kondisi sosial-ekonomi keluarga. Situasi ini menimbulkan kekhawatiran, tidak hanya bagi siswa yang mengalami kesulitan, tetapi juga bagi pihak sekolah dalam menjalankan perannya sebagai lembaga pendidikan yang harus memastikan kelulusan siswanya tepat waktu.

Di tingkat individu siswa, terdapat berbagai tantangan yang dapat memengaruhi keberhasilan dalam menyelesaikan pendidikan dasar. Beberapa siswa mengalami kesulitan dalam mengikuti pelajaran karena keterbatasan pemahaman materi atau kurangnya pendampingan belajar di rumah. Selain itu, faktor eksternal seperti kondisi ekonomi keluarga, lingkungan sosial yang kurang mendukung, serta tekanan psikologis juga dapat memengaruhi proses belajar siswa. Dari sudut pandang sekolah, tantangan ini diperburuk oleh kesulitan dalam memprediksi siswa yang berpotensi menghadapi masalah dalam menyelesaikan pendidikan dasar tepat waktu[1]. Tanpa prediksi yang akurat, institusi sulit mengambil langkah intervensi yang tepat, seperti memberikan bimbingan tambahan, konseling, atau dukungan lainnya.

Untuk menjawab permasalahan tersebut, diperlukan solusi berbasis data yang mampu membantu sekolah mengidentifikasi faktor-faktor penting yang memengaruhi tingkat kelulusan siswa. Pendekatan berbasis data ini membantu sekolah memahami kondisi siswa secara lebih mendalam dan memberikan rekomendasi strategis yang lebih terarah. Salah satu pendekatan yang potensial adalah penerapan algoritma klasifikasi dalam bidang data mining. Data mining telah digunakan secara luas untuk menggali informasi dari data besar, termasuk dalam bidang pendidikan. Teknik ini dapat membantu sekolah memahami pola-pola yang memengaruhi keberhasilan siswa. Dari berbagai algoritma klasifikasi yang tersedia, *K-Nearest Neighbor* (KNN) menjadi salah satu metode yang sangat relevan untuk diterapkan dalam penelitian ini.

Metode KNN merupakan algoritma klasifikasi yang sederhana namun efektif, terutama dalam menangani dataset yang kompleks dan memiliki distribusi non-linear. Prinsip kerja KNN didasarkan pada jarak antara sampel baru dengan data yang sudah ada[2]. Dalam konteks penelitian ini, metode KNN dapat digunakan untuk memprediksi

kelulusan siswa dengan membandingkan karakteristik siswa baru dengan data siswa yang telah lulus dan tidak lulus sebelumnya. Karakteristik ini mencakup berbagai faktor, seperti nilai akademik, kehadiran, partisipasi dalam kegiatan kampus, hingga aspek psikologis. Keunggulan KNN terletak pada kemampuannya untuk menyesuaikan hasil klasifikasi dengan pola data yang ada, sehingga hasil prediksi menjadi lebih akurat[3].

Beberapa penelitian terdahulu telah mengangkat topik terkait prediksi kelulusan siswa menggunakan berbagai algoritma data mining. Penelitian pertama dilakukan oleh Penelitian oleh Luthfiah Nur Aziza dkk. (2024) membahas penerapan algoritma K-NN untuk klasifikasi ketahanan pangan di Provinsi Jawa Tengah, di mana hasilnya menunjukkan bahwa nilai komposit ketahanan pangan meningkat sebesar 1,70% dari tahun 2022 ke 2023, dengan model K-NN ($n_neighbors=3$) mampu memprediksi nilai ketahanan pangan dengan baik, ditunjukkan oleh nilai RMSE sebesar 0,80[4]. Sementara itu, penelitian oleh Silvi Joya Arditna Br Bukit dan Rakhmat Kurniawan (2023) menerapkan algoritma K-NN untuk prediksi harga tandan buah segar (TBS) kelapa sawit berdasarkan data harga dari tahun 2018 hingga Mei 2023. Hasilnya menunjukkan bahwa prediksi harga pada bulan Juli 2023 (bulan ke-67) mencapai akurasi sebesar 10,667 saat $k=3$ dan 19,200 saat $k=5$ [5]. Selanjutnya, Inggianti Pratiwi Putri (Maret 2021) melakukan penelitian tentang performa metode KNN pada dataset penyakit kardiovaskular dengan simulasi rasio data yang berbeda, yaitu 50:50, 20:80, dan 80:20. Pada rasio 50:50, nilai akurasi, presisi, dan recall masing-masing sebesar 82%, dengan F-measure sebesar 80% pada $k = 13$. Pada rasio 20:80, performa terbaik tercapai dengan nilai akurasi 87%, presisi 87%, recall 97%, dan F-measure 92% pada $k = 3$. Sedangkan pada rasio 80:20, nilai akurasi sebesar 91%, presisi 92%, recall 60%, dan F-measure 72% diperoleh pada $k = 5$ [6]. Selanjutnya, Ryan Hamonangan, dkk (Maret 2024) menerapkan algoritma KNN untuk menentukan penerima Bantuan Langsung Tunai (BLT) dengan akurasi mencapai 81,56%. Performa algoritma ditunjukkan melalui nilai recall untuk kategori tidak layak sebesar 88,43% dan kategori layak sebesar 74,80%, serta presisi untuk kategori layak sebesar 86,79% dan kategori tidak layak sebesar 77,54%. Penelitian ini menunjukkan relevansi KNN dalam pengambilan keputusan sosial, memberikan kontribusi pada perancangan kebijakan yang lebih tepat sasaran dan efektif untuk mengurangi kemiskinan[7].

Hasil-hasil tersebut menunjukkan fleksibilitas dan efisiensi algoritma KNN serta algoritma lain dalam berbagai kasus klasifikasi, dari bidang kesehatan hingga pendidikan. Penyesuaian parameter yang tepat pada algoritma KNN dapat meningkatkan performa prediksi dan akurasi hasil. Sebagai kontribusi, penelitian ini tidak hanya menggunakan metode KNN untuk memprediksi kelulusan siswa di SD Negeri 112269 Padang Lais, tetapi juga mengintegrasikan berbagai faktor penting yang jarang dianalisis secara bersamaan. Penelitian ini diharapkan mampu memberikan hasil yang lebih akurat dan mendalam, yang dapat digunakan oleh pihak sekolah untuk mengidentifikasi siswa yang berpotensi mengalami masalah kelulusan. Dengan demikian, sekolah dapat merancang program intervensi yang lebih efektif dan tepat sasaran. Selain itu, penelitian ini juga berkontribusi dalam pengembangan aplikasi metode data mining, khususnya KNN, dalam konteks pendidikan dasar.

Dengan latar belakang ini, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam meningkatkan tingkat kelulusan siswa di SD Negeri 112269 Padang Lais. Hasil penelitian ini tidak hanya bermanfaat bagi siswa yang membutuhkan dukungan tambahan, tetapi juga bagi sekolah dalam merancang strategi yang lebih baik untuk mendukung keberhasilan akademik siswanya.

2. METODOLOGI PENELITIAN

2.1 KDD (*Knowledge Discovery in Database*)

Knowledge Discovery in Database (KDD) merupakan proses sistematis untuk mengekstraksi pengetahuan yang bernilai dari sejumlah besar data dalam database[8][9]. Proses ini bukan hanya sekadar menghimpun informasi, tetapi melibatkan serangkaian tahapan penting agar data yang dihasilkan akurat dan sesuai dengan tujuan analisis[10][11]. Dengan perkembangan pesat dalam volume data, KDD terus menjadi bidang yang penting. Proses ini memungkinkan kita untuk mengubah data mentah menjadi pengetahuan yang berharga dan berkontribusi pada pengambilan keputusan yang lebih cerdas dan strategis[12].

2.2 Data Mining

Data mining merupakan proses penggalian informasi atau pola yang berguna dari kumpulan data yang besar. Proses ini melibatkan berbagai teknik statistik, matematika, dan teknologi komputer untuk mengekstraksi pengetahuan tersembunyi dari data yang tersedia. Tujuan utama dari data mining adalah untuk menemukan hubungan, pola, atau tren dalam data yang dapat digunakan untuk pengambilan keputusan[13][14]. Data mining memiliki beberapa tahap utama, yaitu[15]:

- Pembersihan Data (*Data Cleaning*): Menghapus data yang tidak lengkap atau tidak konsisten.
- Integrasi Data (*Data Integration*): Menggabungkan data dari berbagai sumber.
- Seleksi Data (*Data Selection*): Memilih data yang relevan untuk analisis.
- Transformasi Data (*Data Transformation*): Mengubah data menjadi format yang sesuai untuk proses data mining.
- Penambangan Data (*Data Mining*): Mengaplikasikan algoritma untuk menemukan pola atau informasi.
- Evaluasi dan Presentasi: Mengevaluasi hasil yang diperoleh dan menyajikannya dalam bentuk yang mudah dipahami.

2.3 Algoritma KNN

Algoritma K-Nearest Neighbor (KNN) adalah salah satu algoritma dalam supervised learning yang digunakan untuk tugas klasifikasi dan regresi. Prinsip kerja algoritma ini adalah dengan mengklasifikasikan data baru berdasarkan kedekatannya dengan data-data yang sudah diketahui kelasnya. Kedekatan ini diukur menggunakan metrik jarak, seperti Euclidean, Manhattan, atau Minkowski. Konsep dasar dari KNN adalah dengan mengidentifikasi kelas atau label data baru berdasarkan mayoritas kelas dari k-nearest neighbors (tetangga terdekat) dari data tersebut dalam ruang atribut [16]. Algoritma K-Nearest Neighbor merupakan salah satu metode klasifikasi pengambilan keputusan berdasarkan mayoritas dari kategori tetangga yang terdekat [17][18]. Berikut adalah langkah kerja dalam algoritma K-Nearest Neighbor [19]:

- Siapkan data yang akan diklasifikasikan dan sampel data yang digunakan.
- Tentukan nilai K, yaitu jumlah tetangga terdekat.
- Hitung jarak dengan menggunakan persamaan Euclidean Distance.
- Urutkan hasil perhitungan jarak, dari nilai yang terkecil ke nilai yang terbesar.
- Hasil pengurutan, ambil nilai tetangga yang paling dekat mulai dari nilai yang paling kecil, sesuai dengan jumlah (K) yang telah ditentukan sebelumnya.
- Tentukan kelas data baru berdasarkan mayoritas kelas dari K tetangga terdekat.

Untuk perhitungan jarak menggunakan persamaan Euclidean Distance adalah sebagai berikut [20]:

$$d_{ij} = \sqrt{(x_{1i} - y_{1j})^2 + (x_{2i} - y_{2j})^2 + \dots + (x_{ki} - y_{kj})^2} \quad (1)$$

Keterangan:

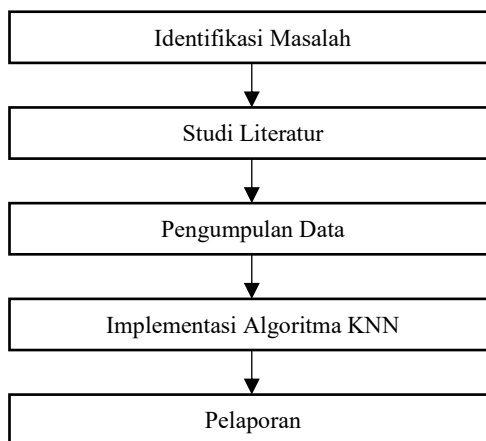
d_{ij} = Jarak dari data ke I ke pusat cluster j

x_{ki} = Data dari ke – i pada attribute data ke – k

x_{kj} = Data dari ke – j pada attribute data ke – k

2.4 Tahapan Penelitian

Penelitian ini dilaksanakan melalui rangkaian tahapan yang dirancang secara sistematis untuk mencapai tujuan penelitian. Setiap tahapan dilakukan dengan pendekatan terstruktur guna memastikan hasil yang akurat dan relevan. Berikut adalah gambar 1 yaitu gambaran tahapan penelitian yang akan dilakukan.



Gambar 1. Tahapan Penelitian

Berdasarkan Gambar 1 berikut adalah penjelasan rinci mengenai setiap tahapan penelitian yang akan dilakukan selama proses penyelesaian permasalahan pada penelitian ini:

- Identifikasi Masalah**
Tahap awal dalam penelitian ini adalah mengidentifikasi permasalahan utama yang menjadi fokus penelitian, yaitu faktor-faktor yang memengaruhi tingkat kelulusan siswa. Identifikasi masalah dilakukan melalui observasi dan diskusi dengan pihak yang terkait, seperti guru, siswa, dan administrator akademik. Hasil dari tahap ini adalah rumusan masalah yang jelas dan terarah.
- Studi Literatur**
Pada tahap ini, dilakukan pengumpulan informasi dari berbagai sumber literatur, seperti jurnal ilmiah, buku, dan laporan penelitian sebelumnya. Studi literatur bertujuan untuk memahami konsep-konsep kunci, seperti algoritma KNN, teknik data mining, serta variabel-variabel yang relevan dengan kelulusan siswa. Informasi yang diperoleh dari studi literatur menjadi dasar untuk merancang pendekatan penelitian.
- Pengumpulan Data**
Data yang digunakan dalam penelitian ini berasal dari SD Negeri 112269 Padang Lais. Data ini dikumpulkan melalui dokumen sekolah, observasi, serta wawancara dengan guru dan staf terkait di SD Negeri 112269 Padang

Lais. Data tersebut akan diolah dan dianalisis menggunakan algoritma *K-Nearest Neighbors* (KNN) untuk mengidentifikasi faktor-faktor signifikan yang memengaruhi tingkat kelulusan siswa.

d. Implementasi Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) diterapkan pada dataset yang telah diproses. Pada tahap ini, parameter K yang optimal ditentukan melalui uji coba iteratif, dengan mempertimbangkan metrik evaluasi seperti akurasi dan *F-measure*. Algoritma ini digunakan untuk memprediksi apakah seorang siswa dapat lulus tepat waktu atau tidak berdasarkan pola dari data historis.

e. Pelaporan

Tahap akhir dari penelitian ini adalah penyusunan laporan penelitian. Laporan mencakup seluruh tahapan penelitian, hasil yang diperoleh, analisis temuan, dan kesimpulan. Laporan disusun dengan format yang sesuai standar akademik dan disertai dengan dokumentasi lengkap.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Pendidikan dasar memiliki peran yang sangat penting dalam membentuk fondasi ilmu pengetahuan dan keterampilan siswa. Keberhasilan sekolah dasar dalam mencetak lulusan yang kompeten sering kali diukur melalui tingkat kelulusan siswa. Namun, banyak sekolah dasar, rendahnya tingkat kelulusan siswa menjadi salah satu permasalahan yang cukup sering dihadapi. Hal ini dipengaruhi oleh berbagai faktor, seperti kemampuan akademik siswa, motivasi belajar, dukungan orang tua, hingga kondisi sosial-ekonomi keluarga.

Di SD Negeri 112269 Padang Lais, permasalahan ini juga terjadi, di mana terdapat sejumlah siswa yang mengalami kesulitan dalam menyelesaikan pendidikan dasar tepat waktu. Kesulitan ini dapat bersumber dari faktor internal siswa, seperti keterbatasan pemahaman materi dan rendahnya minat belajar, maupun faktor eksternal seperti lingkungan keluarga dan kondisi ekonomi yang kurang mendukung. Tanpa adanya metode prediksi yang akurat, pihak sekolah kesulitan untuk mengidentifikasi siswa yang berisiko tidak lulus dan memberikan intervensi yang tepat waktu.

Untuk menjawab permasalahan ini, penelitian ini menerapkan metode klasifikasi berbasis data mining, yaitu algoritma *K-Nearest Neighbor* (KNN), untuk mengidentifikasi faktor-faktor yang mempengaruhi kelulusan siswa. Dengan metode ini, sekolah dapat memahami pola-pola yang mempengaruhi keberhasilan siswa, sehingga memungkinkan adanya tindakan preventif seperti bimbingan belajar tambahan atau konseling akademik.

Dengan demikian, implementasi metode KNN diharapkan dapat membantu SD Negeri 112269 Padang Lais dalam meningkatkan tingkat kelulusan siswa melalui sistem yang lebih terstruktur dan berbasis data.

3.2 Preprocessing Data

Preprocessing data adalah serangkaian langkah dan teknik yang digunakan untuk mengolah data mentah agar lebih terstruktur, rapi, dan siap digunakan dalam analisis, pemodelan, atau proses lainnya. Tahapan ini merupakan langkah awal yang penting sebelum memasuki proses utama dalam data mining. Salah satu teknik yang diterapkan dalam preprocessing adalah normalisasi data, yaitu proses transformasi data ke dalam rentang tertentu tanpa mengubah hubungan relatif antar nilai. Normalisasi bertujuan untuk memastikan bahwa setiap fitur atau variabel memiliki pengaruh yang seimbang dalam analisis atau pemodelan yang akan dilakukan. Dengan normalisasi, perbedaan skala antar fitur dapat disesuaikan sehingga tidak ada fitur yang mendominasi lainnya. Dalam penelitian ini, dari total 101 data siswa, penulis menggunakan sampel sebanyak 25 data siswa, yang disajikan dalam Tabel 1.

Tabel 1. Dataset Data Training

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Belajar	Motivasi dan Minat Belajar	Keterangan
Ahmad	75	92	89	99	Lulus
Budi	76	88	92	87	Lulus
Chandra	89	76	70	82	Lulus
Dewi	92	95	75	56	Lulus
Eka	70	82	87	91	Lulus
Fajar	75	67	82	89	Lulus
Gita	87	59	56	92	Tidak Lulus
Hana	82	73	91	70	Lulus
Indra	56	85	70	75	Tidak Lulus
Joko	91	90	80	87	Lulus
Kiki	70	78	66	82	Tidak Lulus
Lestari	80	62	55	56	Tidak Lulus
Maman	66	55	94	91	Lulus
Nia	74	94	80	70	Lulus
Oka	68	80	71	80	Tidak Lulus

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Belajar	Motivasi dan Minat Belajar	Keterangan
Putri	77	71	89	66	Lulus
Qori	87	89	77	74	Lulus
Riko	60	77	96	68	Lulus
Sari	74	96	69	77	Lulus
Tono	93	69	74	92	Lulus
Umi	71	74	81	88	Lulus
Vino	80	81	93	76	Lulus
Wati	90	93	92	95	Lulus
Xaverius	76	86	88	82	Lulus
Yuni	65	79	76	67	Tidak Lulus

Selanjutnya adalah menentukan data *testing* yang akan digunakan pada penelitian ini. Berikut adalah data *testing* yang dapat dilihat pada Tabel 2.

Tabel 2. Data *Testing*

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi dan Minat Belajar	Keterangan
Fernandus	85	78	82	74	?
Sakinah Natasya	72	65	70	68	?

3.2 Implementasi Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) merupakan salah satu metode klasifikasi dalam machine learning yang bekerja berdasarkan kedekatan atau kemiripan suatu data dengan data lain di sekitarnya. KNN termasuk dalam kategori lazy learning, karena tidak memerlukan proses training eksplisit sebelum melakukan klasifikasi. Setelah data *training* dan data *testing* didapatkan, langkah selanjutnya adalah melakukan langkah-langkah dalam algoritma KNN. Adapun langkah penyelesaian algoritma KNN dapat dilihat pada langkah-langkah berikut ini:

- Menentukan parameter *K*

Nilai *K* yang digunakan pada penelitian adalah *K*=5.

- Proses perhitungan jarak menggunakan persamaan *Euclidean Distance*.

Perhitungan *Euclidean Distance* dilakukan pada masing-masing objek terhadap data sampel yang diberikan, dibawah ini perhitungan *Euclidean Distance*.

Untuk calon Siswa atas nama “Fernandus”.

Berikut adalah perhitungan *Euclidean Distance* untuk siswa atas nama “Fernandus”.

$$d_{xy} = \sqrt{(75 - 85)^2 + (92 - 78)^2 + (89 - 82)^2 + (99 - 74)^2} = \sqrt{100 + 196 + 49 + 625} = \sqrt{970} = 31.145$$

$$d_{xy} = \sqrt{(76 - 85)^2 + (88 - 78)^2 + (92 - 82)^2 + (87 - 74)^2} = \sqrt{81 + 100 + 100 + 169} = \sqrt{450} = 21.213$$

$$d_{xy} = \sqrt{(89 - 85)^2 + (76 - 78)^2 + (70 - 82)^2 + (82 - 74)^2} = \sqrt{16 + 4 + 144 + 64} = \sqrt{228} = 15.100$$

$$d_{xy} = \sqrt{(92 - 85)^2 + (95 - 78)^2 + (75 - 82)^2 + (56 - 74)^2} = \sqrt{49 + 289 + 49 + 324} = \sqrt{711} = 26.665$$

$$d_{xy} = \sqrt{(70 - 85)^2 + (82 - 78)^2 + (87 - 82)^2 + (91 - 74)^2} = \sqrt{225 + 16 + 25 + 289} = \sqrt{555} = 23.558$$

$$d_{xy} = \sqrt{(75 - 85)^2 + (67 - 78)^2 + (82 - 82)^2 + (89 - 74)^2} = \sqrt{100 + 121 + 0 + 225} = \sqrt{446} = 21.119$$

$$d_{xy} = \sqrt{(87 - 85)^2 + (59 - 78)^2 + (56 - 82)^2 + (92 - 74)^2} = \sqrt{4 + 361 + 676 + 324} = \sqrt{1365} = 36.946$$

$$d_{xy} = \sqrt{(82 - 85)^2 + (73 - 78)^2 + (91 - 82)^2 + (70 - 74)^2} = \sqrt{9 + 25 + 81 + 16} = \sqrt{131} = 11.446$$

$$d_{xy} = \sqrt{(56 - 85)^2 + (85 - 78)^2 + (70 - 82)^2 + (75 - 74)^2} = \sqrt{841 + 49 + 144 + 1} = \sqrt{1035} = 32.171$$

$$d_{xy} = \sqrt{(91 - 85)^2 + (90 - 78)^2 + (80 - 82)^2 + (87 - 74)^2} = \sqrt{36 + 144 + 4 + 169} = \sqrt{353} = 18.788$$

$$d_{xy} = \sqrt{(70 - 85)^2 + (78 - 78)^2 + (66 - 82)^2 + (82 - 74)^2} = \sqrt{225 + 0 + 256 + 64} = \sqrt{545} = 23.345$$

$$d_{xy} = \sqrt{(80 - 85)^2 + (62 - 78)^2 + (55 - 82)^2 + (56 - 74)^2} = \sqrt{25 + 256 + 729 + 324} = \sqrt{1334} = 36.524$$

$$d_{xy} = \sqrt{(66 - 85)^2 + (55 - 78)^2 + (94 - 82)^2 + (91 - 74)^2} = \sqrt{361 + 529 + 144 + 289} = \sqrt{1323} = 36.373$$

$$d_{xy} = \sqrt{(74 - 85)^2 + (94 - 78)^2 + (80 - 82)^2 + (70 - 74)^2} = \sqrt{121 + 256 + 4 + 16} = \sqrt{397} = 19.925$$

$$d_{xy} = \sqrt{(68 - 85)^2 + (80 - 78)^2 + (71 - 82)^2 + (80 - 74)^2} = \sqrt{289 + 4 + 121 + 36} = \sqrt{450} = 21.213$$

$$d_{xy} = \sqrt{(77 - 85)^2 + (71 - 78)^2 + (89 - 82)^2 + (66 - 74)^2} = \sqrt{64 + 49 + 49 + 64} = \sqrt{226} = 15.033$$

$$d_{xy} = \sqrt{(87 - 85)^2 + (89 - 78)^2 + (77 - 82)^2 + (74 - 74)^2} = \sqrt{4 + 121 + 25 + 0} = \sqrt{150} = 12.247$$

$$d_{xy} = \sqrt{(60 - 85)^2 + (77 - 78)^2 + (96 - 82)^2 + (68 - 74)^2} = \sqrt{625 + 1 + 196 + 36} = \sqrt{858} = 29.292$$

$$d_{xy} = \sqrt{(74 - 85)^2 + (96 - 78)^2 + (69 - 82)^2 + (77 - 74)^2} = \sqrt{121 + 324 + 169 + 9} = \sqrt{623} = 24.960$$

$$d_{xy} = \sqrt{(93 - 85)^2 + (69 - 78)^2 + (74 - 82)^2 + (92 - 74)^2} = \sqrt{64 + 81 + 64 + 324} = \sqrt{533} = 23.087$$

$$d_{xy} = \sqrt{(71 - 85)^2 + (74 - 78)^2 + (81 - 82)^2 + (88 - 74)^2} = \sqrt{196 + 16 + 1 + 196} = \sqrt{409} = 20.224$$

$$d_{xy} = \sqrt{(80 - 85)^2 + (81 - 78)^2 + (93 - 82)^2 + (76 - 74)^2} = \sqrt{25 + 9 + 121 + 4} = \sqrt{159} = 12.610$$

$$d_{xy} = \sqrt{(90 - 85)^2 + (93 - 78)^2 + (92 - 82)^2 + (95 - 74)^2} = \sqrt{25 + 225 + 100 + 441} = \sqrt{791} = 28.125$$

$$d_{xy} = \sqrt{(76 - 85)^2 + (86 - 78)^2 + (88 - 82)^2 + (82 - 74)^2} = \sqrt{81 + 64 + 36 + 64} = \sqrt{245} = 15.652$$

$$d_{xy} = \sqrt{(65 - 85)^2 + (79 - 78)^2 + (76 - 82)^2 + (67 - 74)^2} = \sqrt{400 + 1 + 36 + 49} = \sqrt{486} = 22.045$$

Dari perhitungan diatas maka didapatkan nilai rekapitulasi hasil perhitungan jarak untuk siswa “Fernandus” sebagai berikut.

Tabel 3. Rekapitulasi Hasil Perhitungan Jarak “Fernandus”

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi dan Minat Belajar	Keterangan	Jarak
Ahmad	75	92	89	99	Lulus	31.145
Budi	76	88	92	87	Lulus	21.213
Chandra	89	76	70	82	Lulus	15.100
Dewi	92	95	75	56	Lulus	26.665
Eka	70	82	87	91	Lulus	23.558
Fajar	75	67	82	89	Lulus	21.119
Gita	87	59	56	92	Tidak Lulus	36.946
Hana	82	73	91	70	Lulus	11.446
Indra	56	85	70	75	Tidak Lulus	32.171
Joko	91	90	80	87	Lulus	18.788
Kiki	70	78	66	82	Tidak Lulus	23.345
Lestari	80	62	55	56	Tidak Lulus	36.524
Maman	66	55	94	91	Lulus	36.373
Nia	74	94	80	70	Lulus	19.925
Oka	68	80	71	80	Tidak Lulus	21.213
Putri	77	71	89	66	Lulus	15.033
Qori	87	89	77	74	Lulus	12.247
Riko	60	77	96	68	Lulus	29.292

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi dan Minat Belajar	Keterangan	Jarak
Sari	74	96	69	77	Lulus	24.960
Tono	93	69	74	92	Lulus	23.087
Umi	71	74	81	88	Lulus	20.224
Vino	80	81	93	76	Lulus	12.610
Wati	90	93	92	95	Lulus	28.125
Xaverius	76	86	88	82	Lulus	15.652
Yuni	65	79	76	67	Tidak Lulus	22.045

Tabel 4. Pengurutan Peringkat Jarak “Fernandus”

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi Dan Minat Belajar	Keterangan	Jarak
Hana	82	73	91	70	Lulus	11.446
Qori	87	89	77	74	Lulus	12.247
Vino	80	81	93	76	Lulus	12.610
Putri	77	71	89	66	Lulus	15.033
Chandra	89	76	70	82	Lulus	15.100
Xaverius	76	86	88	82	Lulus	15.652
Joko	91	90	80	87	Lulus	18.788
Nia	74	94	80	70	Lulus	19.925
Umi	71	74	81	88	Lulus	20.224
Budi	76	88	92	87	Lulus	21.213
Oka	68	80	71	80	Tidak Lulus	21.213
Fajar	75	67	82	89	Lulus	21.119
Yuni	65	79	76	67	Tidak Lulus	22.045
Kiki	70	78	66	82	Tidak Lulus	23.345
Eka	70	82	87	91	Lulus	23.558
Tono	93	69	74	92	Lulus	23.087
Sari	74	96	69	77	Lulus	24.960
Dewi	92	95	75	56	Lulus	26.665
Wati	90	93	92	95	Lulus	28.125
Riko	60	77	96	68	Lulus	29.292
Ahmad	75	92	89	99	Lulus	31.145
Indra	56	85	70	75	Tidak Lulus	32.171
Maman	66	55	94	91	Lulus	36.373
Lestari	80	62	55	56	Tidak Lulus	36.524
Gita	87	59	56	92	Tidak Lulus	36.946

Tabel 4 "Pengurutan Peringkat Jarak 'Fernandus'" menampilkan data siswa yang diurutkan berdasarkan nilai Jarak, dari yang terkecil hingga terbesar. Nilai jarak ini menunjukkan kedekatan siswa terhadap suatu kriteria tertentu dalam evaluasi. Siswa dengan jarak terkecil, seperti Hana (11.446) dan Qori (12.247), cenderung memiliki peluang lebih tinggi untuk lulus, sedangkan siswa dengan jarak terbesar, seperti Gita (36.946) dan Lestari (36.524), umumnya tidak lulus. Dari perhitungan di atas, ternyata 5 (lima) tetangga terdekat dari siswa atas nama Fernandus memiliki keterangan "Lulus", maka dapat disimpulkan bahwa siswa atas nama Fernandus juga dinyatakan "Lulus". Untuk calon Siswa atas nama “Sakinah Natasya”.

Berikut adalah perhitungan *Euclidean Distance* untuk siswa atas nama “Sakinah Natasya”.

$$d_{xy} = \sqrt{(75 - 72)^2 + (92 - 65)^2 + (89 - 70)^2 + (99 - 68)^2} = \sqrt{9 + 729 + 361 + 961} = \sqrt{2060} = 45.387$$

$$d_{xy} = \sqrt{(76 - 72)^2 + (88 - 65)^2 + (92 - 70)^2 + (87 - 68)^2} = \sqrt{16 + 529 + 484 + 361} = \sqrt{1390} = 37.283$$

$$d_{xy} = \sqrt{(89 - 72)^2 + (76 - 65)^2 + (70 - 70)^2 + (82 - 68)^2} = \sqrt{289 + 121 + 0 + 196} = \sqrt{606} = 24.617$$

$$d_{xy} = \sqrt{(92 - 72)^2 + (95 - 65)^2 + (75 - 70)^2 + (56 - 68)^2} = \sqrt{400 + 900 + 25 + 144} = \sqrt{1469} = 38.328$$

$$d_{xy} = \sqrt{(70 - 72)^2 + (82 - 65)^2 + (87 - 70)^2 + (91 - 68)^2} = \sqrt{4 + 289 + 289 + 529} = \sqrt{1111} = 33.332$$

$$d_{xy} = \sqrt{(75 - 72)^2 + (67 - 65)^2 + (82 - 70)^2 + (89 - 68)^2} = \sqrt{9 + 4 + 144 + 441} = \sqrt{598} = 24.454$$

$$d_{xy} = \sqrt{(87 - 72)^2 + (59 - 65)^2 + (56 - 70)^2 + (92 - 68)^2} = \sqrt{225 + 36 + 196 + 576} = \sqrt{1033} = 32.140$$

$$d_{xy} = \sqrt{(82 - 72)^2 + (73 - 65)^2 + (91 - 70)^2 + (70 - 68)^2} = \sqrt{100 + 64 + 441 + 4} = \sqrt{609} = 24.678$$

$$d_{xy} = \sqrt{(56 - 72)^2 + (85 - 65)^2 + (70 - 70)^2 + (75 - 68)^2} = \sqrt{256 + 400 + 0 + 49} = \sqrt{705} = 26.552$$

$$d_{xy} = \sqrt{(91 - 72)^2 + (90 - 65)^2 + (80 - 70)^2 + (87 - 68)^2} = \sqrt{361 + 625 + 100 + 361} = \sqrt{1447} = 38.039$$

$$d_{xy} = \sqrt{(70 - 72)^2 + (78 - 65)^2 + (66 - 70)^2 + (82 - 68)^2} = \sqrt{4 + 169 + 16 + 196} = \sqrt{385} = 19.621$$

$$d_{xy} = \sqrt{(80 - 72)^2 + (62 - 65)^2 + (55 - 70)^2 + (56 - 68)^2} = \sqrt{64 + 9 + 225 + 144} = \sqrt{442} = 21.024$$

$$d_{xy} = \sqrt{(66 - 72)^2 + (55 - 65)^2 + (94 - 70)^2 + (91 - 68)^2} = \sqrt{36 + 100 + 576 + 529} = \sqrt{1241} = 35.228$$

$$d_{xy} = \sqrt{(74 - 72)^2 + (94 - 65)^2 + (80 - 70)^2 + (70 - 68)^2} = \sqrt{4 + 841 + 100 + 4} = \sqrt{949} = 30.806$$

$$d_{xy} = \sqrt{(68 - 72)^2 + (80 - 65)^2 + (71 - 70)^2 + (80 - 68)^2} = \sqrt{16 + 225 + 1 + 144} = \sqrt{386} = 19.647$$

$$d_{xy} = \sqrt{(77 - 72)^2 + (71 - 65)^2 + (89 - 70)^2 + (66 - 68)^2} = \sqrt{25 + 36 + 361 + 4} = \sqrt{426} = 20.640$$

$$d_{xy} = \sqrt{(87 - 72)^2 + (89 - 65)^2 + (77 - 70)^2 + (74 - 68)^2} = \sqrt{225 + 576 + 49 + 36} = \sqrt{886} = 29.766$$

$$d_{xy} = \sqrt{(60 - 72)^2 + (77 - 65)^2 + (96 - 70)^2 + (68 - 68)^2} = \sqrt{144 + 144 + 676 + 0} = \sqrt{964} = 31.048$$

$$d_{xy} = \sqrt{(74 - 72)^2 + (96 - 65)^2 + (69 - 70)^2 + (77 - 68)^2} = \sqrt{4 + 961 + 1 + 81} = \sqrt{1047} = 32.357$$

$$d_{xy} = \sqrt{(93 - 72)^2 + (69 - 65)^2 + (74 - 70)^2 + (92 - 68)^2} = \sqrt{441 + 16 + 16 + 576} = \sqrt{1049} = 32.388$$

$$d_{xy} = \sqrt{(71 - 72)^2 + (74 - 65)^2 + (81 - 70)^2 + (88 - 68)^2} = \sqrt{1 + 81 + 121 + 400} = \sqrt{603} = 24.556$$

$$d_{xy} = \sqrt{(80 - 72)^2 + (81 - 65)^2 + (93 - 70)^2 + (76 - 68)^2} = \sqrt{64 + 256 + 529 + 64} = \sqrt{913} = 30.216$$

$$d_{xy} = \sqrt{(90 - 72)^2 + (93 - 65)^2 + (92 - 70)^2 + (95 - 68)^2} = \sqrt{324 + 784 + 484 + 729} = \sqrt{2321} = 48.177$$

$$d_{xy} = \sqrt{(76 - 72)^2 + (86 - 65)^2 + (88 - 70)^2 + (82 - 68)^2} = \sqrt{16 + 441 + 324 + 196} = \sqrt{977} = 31.257$$

$$d_{xy} = \sqrt{(65 - 72)^2 + (79 - 65)^2 + (76 - 70)^2 + (67 - 68)^2} = \sqrt{49 + 196 + 36 + 1} = \sqrt{282} = 16.793$$

Dari perhitungan diatas maka didapatkan nilai rekapitulasi hasil perhitungan jarak untuk siswa “Sakinah Natasya” dapat dilihat pada Tabel 5.

Tabel 5. Rekapitulasi Hasil Perhitungan Jarak “Sakinah Natasya”

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi Dan Minat Belajar	Keterangan	Jarak
Ahmad	75	92	89	99	Lulus	45.387
Budi	76	88	92	87	Lulus	37.283
Chandra	89	76	70	82	Lulus	24.617

Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi Dan Minat Belajar	Keterangan	Jarak
Dewi	92	95	75	56	Lulus	38.328
Eka	70	82	87	91	Lulus	33.332
Fajar	75	67	82	89	Lulus	24.454
Gita	87	59	56	92	Tidak Lulus	32.140
Hana	82	73	91	70	Lulus	24.678
Indra	56	85	70	75	Tidak Lulus	26.552
Joko	91	90	80	87	Lulus	38.039
Kiki	70	78	66	82	Tidak Lulus	19.621
Lestari	80	62	55	56	Tidak Lulus	21.024
Maman	66	55	94	91	Lulus	35.228
Nia	74	94	80	70	Lulus	30.806
Oka	68	80	71	80	Tidak Lulus	19.647
Putri	77	71	89	66	Lulus	20.640
Qori	87	89	77	74	Lulus	29.766
Riko	60	77	96	68	Lulus	31.048
Sari	74	96	69	77	Lulus	32.357
Tono	93	69	74	92	Lulus	32.388
Umi	71	74	81	88	Lulus	24.556
Vino	80	81	93	76	Lulus	30.216
Wati	90	93	92	95	Lulus	48.177
Xaverius	76	86	88	82	Lulus	31.257
Yuni	65	79	76	67	Tidak Lulus	16.793

Tabel 6. Pengurutan Peringkat Jarak “Sakinah Natasya”

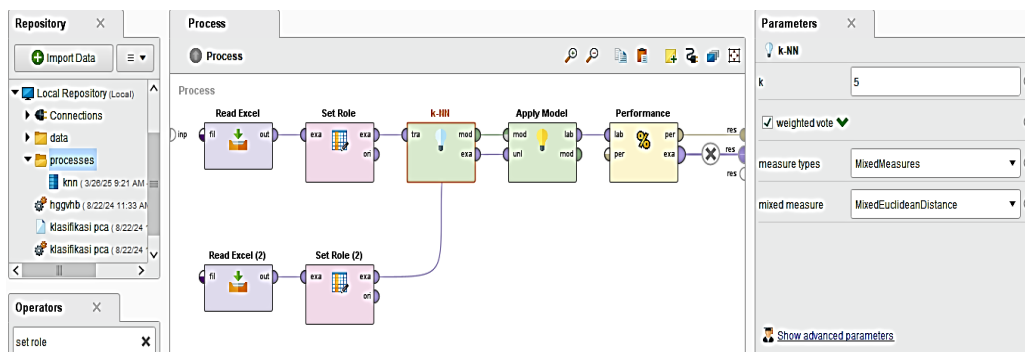
Nama Siswa/I	Kehadiran	Nilai Akhir	Keaktifan Pembelajaran	Motivasi Dan Minat Belajar	Keterangan	Jarak
Yuni	65	79	76	67	Tidak Lulus	16.793
Kiki	70	78	66	82	Tidak Lulus	19.621
Oka	68	80	71	80	Tidak Lulus	19.647
Putri	77	71	89	66	Lulus	20.640
Lestari	80	62	55	56	Tidak Lulus	21.024
Chandra	89	76	70	82	Lulus	24.617
Fajar	75	67	82	89	Lulus	24.454
Hana	82	73	91	70	Lulus	24.678
Umi	71	74	81	88	Lulus	24.556
Indra	56	85	70	75	Tidak Lulus	26.552
Qori	87	89	77	74	Lulus	29.766
Nia	74	94	80	70	Lulus	30.806
Vino	80	81	93	76	Lulus	30.216
Riko	60	77	96	68	Lulus	31.048
Xaverius	76	86	88	82	Lulus	31.257
Sari	74	96	69	77	Lulus	32.357
Tono	93	69	74	92	Lulus	32.388
Gita	87	59	56	92	Tidak Lulus	32.140
Eka	70	82	87	91	Lulus	33.332
Maman	66	55	94	91	Lulus	35.228
Budi	76	88	92	87	Lulus	37.283
Dewi	92	95	75	56	Lulus	38.328
Joko	91	90	80	87	Lulus	38.039
Ahmad	75	92	89	99	Lulus	45.387
Wati	90	93	92	95	Lulus	48.177

Tabel 6 "Pengurutan Peringkat Jarak Sakinah Natasya" menampilkan data siswa yang diurutkan berdasarkan nilai jarak, dari yang terkecil hingga terbesar. Nilai jarak ini menunjukkan tingkat kedekatan siswa terhadap kriteria tertentu dalam evaluasi. Siswa dengan nilai jarak yang lebih kecil, seperti Yuni (16.793) dan Kiki (19.621), memiliki kemiripan lebih tinggi terhadap kriteria yang digunakan dalam penentuan kelulusan. Sebaliknya, siswa dengan jarak lebih besar, seperti Wati (48.177) dan Ahmad (45.387), memiliki karakteristik yang lebih berbeda dari siswa lainnya dalam evaluasi ini. Berdasarkan perhitungan dalam tabel, lima tetangga terdekat dari Sakinah Natasya terdiri dari Yuni, Kiki, Oka, Putri, dan Lestari. Dari lima tetangga terdekat tersebut, satu siswa memiliki keterangan "Lulus" (Putri), sedangkan empat lainnya berstatus "Tidak Lulus". Oleh karena itu, berdasarkan prinsip

klasifikasi K-Nearest Neighbors (KNN), dapat disimpulkan bahwa siswa atas nama Sakinah Natasya kemungkinan besar dikategorikan sebagai "Tidak Lulus".

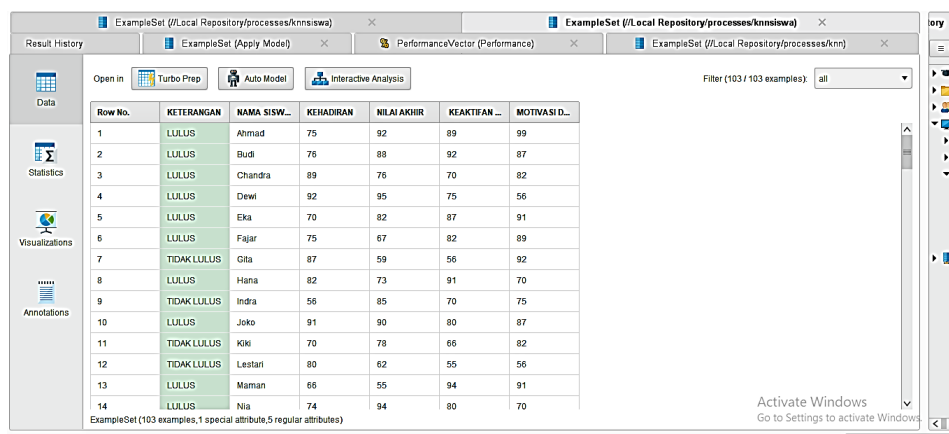
3.3 Hasil Pengujian

Bagian ini menjelaskan bagaimana algoritma KNN diterapkan di *RapidMiner*, termasuk konfigurasi parameter K, pemilihan fitur, proses pelatihan, serta evaluasi hasil klasifikasi.



Gambar 2. Tampilan Penerapan Algoritma KNN

Pada Gambar 2, tampilan penerapan algoritma KNN terdapat 2 operator *Read Excel*. *Read Excel* atas adalah dataset lama atau data *training*, sedangkan *read excel (2)* bagian bawah adalah dataset baru atau data *testing*. Pada gambar 2 juga terlihat parameter K yang digunakan adalah 5. Pada bagian data *testing* terdapat operator KNN, *Apply Model*, dan *Performance*.



Row No.	KETERANGAN	NAMA SISW...	KEHADIRAN	NILAI AKHIR	KEAKTIFAN ...	MOTIVASI D...
1	LULUS	Ahmad	75	92	89	99
2	LULUS	Budi	76	88	92	87
3	LULUS	Chandra	89	76	70	82
4	LULUS	Dewi	92	95	75	56
5	LULUS	Eka	70	82	87	91
6	LULUS	Fajar	75	67	82	89
7	TIDAK LULUS	Gita	87	59	56	92
8	LULUS	Hana	82	73	91	70
9	TIDAK LULUS	Indra	56	85	70	75
10	LULUS	Joko	91	90	80	87
11	TIDAK LULUS	Kiki	70	78	66	82
12	TIDAK LULUS	Lestari	80	62	55	56
13	LULUS	Maman	66	55	94	91
14	LULUS	Nia	74	94	80	70

Gambar 3. Hasil Pembacaan Data oleh *RapidMiner*

Gambar 3 membuktikan data yang telah terbaca dengan baik secara menyeluruh oleh *RapidMiner*.

PerformanceVector

```
PerformanceVector:
accuracy: 100.00%
ConfusionMatrix:
True:  LULUS   TIDAK LULUS
LULUS:  1      0
TIDAK LULUS:  0      1
```

Gambar 4. Hasil *Performace Vector*

accuracy: 100.00%

	true LULUS	true TIDAK LULUS	class precision
pred. LULUS	1	0	100.00%
pred. TIDAK LULUS	0	1	100.00%
class recall	100.00%	100.00%	

Gambar 5. Hasil *Accuracy*

Gambar 4 dan Gambar 5 menampilkan hasil akurasi dari *performance vector* sebesar 100.00% untuk proses yang dilaksanakan dengan menerapkan algoritma.

4. KESIMPULAN

Berdasarkan hasil penelitian "Analisis Faktor-Faktor yang Mempengaruhi Tingkat Kelulusan Siswa Menggunakan Algoritma KNN (Studi Kasus: SD Negeri 112269 Padang Lais)", dapat disimpulkan bahwa algoritma K-Nearest Neighbors (KNN) efektif dalam memprediksi kelulusan siswa berdasarkan faktor-faktor tertentu yang mempengaruhi keberhasilan akademik. Hasil klasifikasi menunjukkan bahwa siswa dengan nilai jarak terdekat terhadap siswa yang telah lulus memiliki kemungkinan lebih besar untuk lulus, sedangkan siswa dengan jarak lebih jauh cenderung tidak lulus. Penerapan KNN dengan $K=5$ menghasilkan pola klasifikasi yang konsisten, di mana mayoritas dari lima tetangga terdekat menentukan status kelulusan siswa yang diuji, membuktikan bahwa metode ini dapat memberikan prediksi yang akurat. Selain itu, akurasi model yang diperoleh mencapai 100%, yang menunjukkan bahwa algoritma KNN dapat mengklasifikasikan data dengan sangat baik. Namun, hasil ini perlu dianalisis lebih lanjut untuk menghindari potensi overfitting, sehingga model tetap dapat diterapkan pada dataset lain dengan hasil yang tetap akurat dan dapat digeneralisasi.

REFERENCES

- [1] S. Hartati and H. A. SAN, "Algoritma Naive Bayes untuk Prediksi Kelulusan Mahasiswa," *Jurnal Cakrawala Informasi*, vol. 2, no. 2, pp. 42–50, 2022, doi: 10.54066/jci.v2i2.234.
- [2] E. Wahyudin, K. Kaslani, and R. Widiyanti, "Implementasi Algoritma K-Nearest Neighbor Untuk Klasifikasi Data Pasien Rehabilitasi Narkoba Berdasarkan Rentang Usia," *KOPERTIP: Scientific Journal of Informatics Management and Computer*, vol. 7, no. 3, pp. 76–79, 2023, doi: 10.32485/kopertip.v7i3.333.
- [3] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [4] L. N. Aziza, R. Y. Astuti, B. A. Maulana, and N. Hidayati, "Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Ketahanan Pangan di Provinsi Jawa Tengah: Application of The K-Nearest Neighbor Algorithm for Food Security Classification In Central Java Province," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 404–412, 2024, doi: 10.57152/malcom.v4i2.1201.
- [5] S. J. A. B. Bukit and R. R. Kurniawan, "Prediksi Harga Tandan Buah Segar dengan Algoritma K-Nearest Neighbor," *Jurnal Sistem Komputer dan Informatika (JSON) Hal*: 92, vol. 101, no. 1, pp. 92–101, 2023, doi: 10.30865/json.v5i1.6818.
- [6] I. P. Putri, "Analisis Performa Metode K-Nearest Neighbor (KNN) dan Crossvalidation pada Data Penyakit Cardiovascular," *Indonesian Journal Of Data And Science*, vol. 2, no. 1, pp. 21–28, 2021, doi: 10.33096/ijodas.v2i1.25.
- [7] R. Hamonangan, R. Komala Sari, S. Anwar, and T. Hartati, "Klasifikasi Algoritma KNN dalam menentukan Penerima Bantuan Langsung Tunai," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, 2024, doi: 10.36499/jinrpl.v6i1.10298.
- [8] S. A. A. Kharis, A. H. A. Zili, A. Putri, and A. Robiansyah, "Analisis Tren Minat Masyarakat Indonesia terhadap Artificial Intelligence dalam Menyongsong Society 5.0: Studi Menggunakan Google Trends," *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 4, pp. 1345–1354, 2023, doi: 10.33379/gtech.v7i4.3091.
- [9] M. A. A. Leza, N. W. Utami, and P. A. C. Dewi, "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)*, vol. 8, no. 1, pp. 373–379, 2024, doi: 10.36040/jati.v8i1.8754.
- [10] U. Suriani, "Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4. 5," *Journal of Computer and Information Systems Ampera*, vol. 4, no. 2, pp. 55–65, 2023, doi: 10.51519/journalcisa.v4i2.393.
- [11] A. Sitanggang, Y. Umaidah, and R. I. Adam, "Analisis sentimen masyarakat terhadap program makan siang gratis pada media sosial X menggunakan algoritma Naïve Bayes," *Jurnal Informatika Dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4902.
- [12] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 5, no. 1, pp. 1–7, 2022, doi: 10.31539/intecom.v5i1.3708.
- [13] C. Hardjono and S. M. Isa, "Implementation of Data Mining for Churn Prediction in Music Streaming Company Using 2020 Dataset," *Journal on Education*, vol. 5, no. 1, pp. 1189–1197, 2022, doi: 10.31004/joe.v5i1.740.
- [14] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan dan Minuman Menggunakan Metode Algoritma Naïve Bayes: Studi Kasus: Makan Barbeque Sepuasnya," *Jurnal Ilmiah Informatika*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.
- [15] D. Marlina and M. Bakri, "Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4. 5," *Jurnal Teknologi Dan Sistem Informasi*, vol. 2, no. 1, pp. 23–28, 2021, doi: 10.33365/jtsi.v2i1.627.
- [16] Z. Rahman, Z. Fatah, and J. D. Prasetyo, "Klasifikasi Spesies Bunga Iris Menggunakan Algoritma Klasifikasi Knn Di Rapidminer," *Jurnal Ilmiah Multidisiplin Ilmu*, vol. 1, no. 6, pp. 60–66, 2024, doi: 10.69714/0syd5n74.
- [17] R. D. Kurniawan and J. Muliawan, "Sentiment analysis of indonesian election 2024 using the k-nearest neighbor method," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 2, pp. 653–659, 2024, doi: 10.52436/1.jutif.2024.5.2.1934.
- [18] R. Ramadhani and R. Ramadhanu, "Metode Machine Learning untuk Klasifikasi Data Gizi Balita dengan Algoritma Naïve Bayes, KNN dan Decision Tree," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 15, no. 1, 2024, doi: 10.24176/simet.v15i1.10679.



- [19] E. Wahyudin, K. Kaslani, and R. Widiyanti, “Implementasi Algoritma K-Nearest Neighbor Untuk Klasifikasi Data Pasien Rehabilitasi Narkoba Berdasarkan Rentang Usia,” *KOPERTIP: Scientific Journal of Informatics Management and Computer*, vol. 7, no. 3, pp. 76–79, 2023, doi: 10.32485/kopertip.v7i3.333.
- [20] H. Hafsyah, E. Haerani, N. Novriyanto, and F. Syafria, “Analisis Sentimen Tanggapan Masyarakat Terhadap Kenaikan Biaya Haji Tahun 2023 Menggunakan Metode K-Nearest Neighbor (KNN),” *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 3, pp. 1562–1569, 2023, doi: 10.30865/klik.v4i3.1471.