KLASIFIKASI SISWA BERPRESTASI MENGGUNAKAN METODE *K-NEAREST NEIGHBOR* PADA SDN 06 PAGI PASAR REBO

Abdul Fatah Khoiron[1], Abdurahman[2], Acep[3]
Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer
Universitas Indraprasta PGRI
Jalan Raya Tengah No 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur
khoir957@gmail.com[1], abdjur05@gmail.com[2], acepdpk@gmail.com[3]

Abstrak

Proses penentuan siswa berprestasi di SDN Baru 06 Pagi selama ini dilakukan secara manual memerlukan waktu yang relatif lama serta berpotensi menghasilkan penilaian yang kurang objektif. Penelitian ini bertujuan untuk mempermudah pihak sekolah dalam mengklasifikasikan siswa berprestasi dan tidak berprestasi di SDN Baru 06 Pagi. Oleh karena itu, penelitian ini menerapkan metode *K-Nearest Neighbor* (KNN) sebagai metode klasifikasi dengan membandingkan data uji terhadap data latih yang telah diberi label berprestasi dan tidak berprestasi berdasarkan nilai kedekatan. Hasil penelitian menunjukkan bahwa aplikasi yang dikembangkan mampu membantu pihak sekolah dalam menentukan status prestasi siswa secara lebih efektif, cepat, dan objektif. Evaluasi kinerja model menghasilkan nilai akurasi sebesar 85,48%, presisi 82,69%, *recall* 100%, dan *F1-Score* 90,53%. Berdasarkan hasil evaluasi tersebut, metode *K-Nearest Neighbor* memiliki kinerja yang baik dan layak diterapkan sebagai pendukung pengambilan keputusan dalam klasifikasi siswa berprestasi.

Kata Kunci: *Data Mining*, *K-Nearest Neighbor*, Klasifikasi Siswa Berprestasi, SDN Baru 06 Pagi

Abstract

The process of determining outstanding students at SDN Baru 06 Pagi *has so far been done manually, requiring a relatively long time and potentially resulting in less objective assessments. This research aims to facilitate the school in classifying students as high-achievers and low-achievers at* SDN Baru 06 Pagi*. Therefore, this research applies the K-Nearest Neighbor (KNN) method as a classification method by comparing test data against training data that has been labeled as high-achieving and low-achieving based on proximity values. The research results indicate that the developed application is capable of assisting the school in determining students' achievement status more effectively, quickly, and objectively. The model's performance evaluation resulted in an accuracy of 85.48%, a precision of 82.69%, a recall of 100%, and an F1-score of 90.53%. Based on these evaluation results, the K-Nearest Neighbor method has excellent performance and is suitable to be applied as a decision support tool in the classification of high-achieving students.*

Keywords: *Data Mining, K-Nearest Neighbor, Classification of High-Achieving Students,* SDN Baru 06 Pagi

PENDAHULUAN

Sekolah sebagai lembaga pendidikan menghasilkan berbagai jenis data, seperti data nilai siswa, absensi, sikap, dan kegiatan ekstrakurikuler yang terus bertambah setiap semester. Jika data tersebut dikelola dengan baik, maka dapat menjadi informasi yang bermanfaat dalam mendukung proses evaluasi dan penentuan kebijakan pendidikan. Menurut Barus et al. (2025) menjelaskan bahwa teknologi informasi berperan penting dalam mendukung proses pengambilan keputusan di lembaga pendidikan karena mampu menyediakan informasi yang lebih cepat, akurat, dan terintegrasi.

Dalam bidang pendidikan, penerapan data mining dapat dimanfaatkan untuk berbagai kebutuhan, salah satunya adalah klasifikasi siswa berprestasi berdasarkan data akademik maupun non-akademik yang dimiliki siswa. Data mining dapat digunakan untuk mengolah data pendidikan menjadi informasi yang berguna dalam mendukung proses pengambilan keputusan dan evaluasi prestasi siswa (Rahayu et al., 2024).

Penerapan data mining pada lingkungan sekolah dapat membantu pihak sekolah dalam melakukan analisis data siswa secara lebih cepat, tepat, dan objektif. Dengan adanya sistem klasifikasi siswa berprestasi, sekolah dapat mengetahui siswa yang memiliki potensi prestasi berdasarkan kriteria tertentu, seperti nilai mata pelajaran, tingkat kehadiran, sikap, dan kegiatan ekstrakurikuler. Selain itu, hasil klasifikasi juga dapat digunakan sebagai bahan evaluasi bagi guru dan pihak sekolah dalam

meningkatkan kualitas pembelajaran. Menurut Husein & Hutauruk (2022) menjelaskan bahwa teknik data mining dapat dimanfaatkan untuk proses klasifikasi siswa berprestasi sehingga membantu pihak sekolah dalam menentukan siswa yang memenuhi kriteria tertentu secara lebih sistematis. Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Noviyanto et al., 2021). Selain itu, data mining didefinisikan juga sebagai proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengidentifikasi informasi yang bermanfaat dari kumpulan data yang besar (Ramadhani & Sensuse, 2023).

Namun, pada kenyataannya masih banyak sekolah yang belum menerapkan teknologi pengolahan data secara optimal, khususnya dalam proses penentuan siswa berprestasi. Proses penilaian umumnya masih dilakukan secara manual dengan melihat data satu per satu sehingga membutuhkan waktu yang cukup lama dan hasil penilaian menjadi kurang objektif. Selain itu, banyaknya data siswa yang harus dianalisis juga menjadi kendala tersendiri bagi pihak sekolah dalam menentukan siswa berprestasi secara efektif dan efisien.

Salah satunya SDN Baru 06 Pagi Pasar Rebo. Proses penentuan siswa berprestasi di SDN Baru 06 Pagi Pasar Rebo masih dilakukan secara manual sehingga membutuhkan waktu yang cukup lama dalam pengolahan dan penilaian data siswa. Selain itu banyaknya data siswa seperti nilai akademik, absensi, sikap, dan kegiatan ekstrakurikuler belum dimanfaatkan secara optimal sebagai informasi dalam pengambilan keputusan untuk menentukan siswa berprestasi. Kondisi ini dapat menyebabkan hasil keputusan kurang objektif. Metode penelitian dalam penelitian ini menggunakan metode K-Nearest Neighbor (KNN). Menurut Nofriansyah & Nurcahyo (2021) menjelaskan bahwa KNN adalah algoritma yang mengklasifikasikan objek baru berdasarkan atribut dan data latih yang telah ada sebelumnya, di mana kelas yang ditetapkan merupakan kelas yang paling sering muncul di antara k tetangga terdekat dari objek tersebut. Selain itu, KNN memiliki keunggulan dalam hal kemudahan implementasi dan kemampuannya menghasilkan akurasi yang cukup tinggi meskipun dengan proses yang sederhana (Wahyuni & Wahyudi, 2019).

Hasil dari penelitian ini adalah merancang sebuah sistem klasifikasi siswa berprestasi yang dapat membantu pihak sekolah agar mendapatkan hasil yang berjalan secara cepat, efektif dan hasil yang objektif.

PENELITIAN RELEVAN

Berdasarkan penelitian sebelumnya oleh (Ardiaz Fahrezi Maulana Rusli, 2025) dengan judul Penerapan Data Mining untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. Hasil penelitian adalah penerapan algoritma KNN mampu membantu pihak toko atau perusahaan dalam mengetahui produk yang memiliki potensi penjualan tertinggi secara lebih cepat dan efektif dibandingkan proses analisis manual.

Penelitian lainnya (Nur Mauliza, 2025) dengan judul Sistem Klasifikasi Data Penjualan Barang Di Zahra Market Menggunakan Metode Naive Bayes dan K-Nearest Neighbor (KNN). Hasil Penelitian ini Hasil pengujian menunjukkan bahwa KNN memperoleh akurasi tertinggi sebesar 89,58%, sehingga dinilai lebih efektif dibandingkan Naïve Bayes yang memperoleh akurasi 83,33%.

Perbedaan penelitian tersebut dengan penelitian terdahulu yang dijadikan acuan adalah penelitian terdahulu menggunakan studi kasus pada produk terlaris, sedangkan penelitian yang dilakukan menggunakan studi kasus klasifikasi untuk siswa berprestasi.

METODE PENELITIAN

Waktu yang digunakan dalam pelaksanaan penelitian ini selama lima bulan, yang dimulai dari bulan April tahun 2026 sampai dengan bulan Agustus tahun 2026. Tahapan penelitian terdiri dari :

1. Tahap Pengumpulan data

   Tahap pengumpulan data dilakukan untuk mendapatkan data yang dibutuhkan dalam proses penelitian. Data yang dikumpulkan adalah data siswa yang meliputi nilai akademik, tingkat kehadiran (absensi), penilaian sikap, dan keikutsertaan dalam kegiatan ekstrakurikuler. Teknik pengumpulan data yang digunakan dalam penelitian yaitu wawancara secara langsung dengan kepala sekolah SDN Baru 06 Pagi, yaitu Ibu Asmaul Husnah, M.Si., M.Pd., serta Guru Koordinator penelitian, yaitu Ibu Prima

Jayensih, S.Pd. untuk mendapatkan informasi mengenai proses penilaian siswa berprestasi yang selama ini diterapkan di sekolah serta kriteria yang digunakan.

2. Tahap *Pre-processing* Data
   Selanjutnya dilakukan proses *pre-processing* data untuk mempersiapkan data sebelum diolah menggunakan algoritma klasifikasi. Tahap *pre-processing* meliputi pemeriksaan kelengkapan data, pengubahan data kategorik ke bentuk numerik (*encoding*), serta normalisasi data agar seluruh atribut memiliki skala yang sama dan dapat diproses secara optimal oleh algoritma.
3. Tahap Implementasi Algoritma
   Pada tahap ini, algoritma KNN diimplementasikan ke dalam sistem klasifikasi siswa berprestasi berbasis web. Implementasi mencakup seluruh proses mulai dari pembagian data latih dan data uji, perhitungan jarak *Euclidean Distance*, penentuan kelas berdasarkan voting K tetangga terdekat, hingga pengembangan antarmuka sistem menggunakan bahasa pemrograman PHP dan basis data MySQL.
4. Tahap Evaluasi Sistem
   Setelah implementasi selesai, dilakukan evaluasi untuk mengukur kinerja sistem yang telah dibangun. Evaluasi dilakukan menggunakan metode confusion matrix dengan menghitung nilai akurasi, presisi, recall, dan F1-Score.
5. Tahap Penarikan Simpulan
   Tahap terakhir adalah penarikan simpulan berdasarkan hasil evaluasi dan pembahasan yang telah dilakukan.

Secara umum, cara kerja algoritma KNN dalam proses klasifikasi adalah sebagai berikut (Nofriansyah & Nurcahyo, 2021):

1. Menentukan nilai k sebagai jumlah tetangga terdekat yang akan dijadikan acuan dalam proses klasifikasi.
2. Menghitung jarak antara data uji dengan seluruh data latih menggunakan metode perhitungan jarak tertentu, seperti *Euclidean Distance*.
3. Mengurutkan hasil perhitungan jarak dari yang terkecil hingga terbesar.
4. Mengambil k data latih yang memiliki jarak terdekat dengan data uji.
5. Menentukan kelas dari data uji berdasarkan kelas yang paling dominan (mayoritas) di antara k tetangga terdekat tersebut.

Dalam algoritma KNN, perhitungan jarak antar data umumnya menggunakan rumus *Euclidean Distance*. Menurut Wahyuni & Wahyudi (2019) menjelaskan bahwa *Euclidean Distance* merupakan metode perhitungan jarak yang paling umum digunakan dalam algoritma KNN karena mudah dipahami dan diterapkan. Rumus *Euclidean Distance* adalah sebagai berikut:

$$\mathbf{d(x, y) = \sqrt{\Sigma\ (x_i - y_i)^2}}$$

Keterangan:

1. d(x, y) : jarak *Euclidean* antara data x dan data y
2. x_i : nilai atribut ke-i dari data uji
3. y_i : nilai atribut ke-i dari data latih
4. i : indeks atribut (i = 1, 2, 3, ..., n)

Semakin kecil nilai jarak yang dihasilkan, maka semakin mirip atau semakin dekat kedua data tersebut.

Pemilihan nilai *k* merupakan salah satu aspek penting dalam algoritma KNN yang dapat mempengaruhi hasil klasifikasi secara signifikan. Menurut Kurniawan et al. (2022), tidak ada aturan baku dalam menentukan nilai *k* yang optimal, namun terdapat beberapa pertimbangan yang perlu diperhatikan:

1. Nilai k yang terlalu kecil (misalnya k = 1) dapat menyebabkan model terlalu sensitif terhadap data pencilan (*outlier*) sehingga rentan mengalami *overfitting*.
2. Nilai k yang terlalu besar dapat menyebabkan batas keputusan menjadi terlalu *smooth* sehingga kurang mampu menangkap pola yang lebih detail (*underfitting*).

3. Secara umum, nilai k yang disarankan adalah bilangan ganjil untuk menghindari hasil seri dalam penentuan kelas mayoritas.
4. Nilai k yang optimal dapat ditentukan melalui eksperimen dengan mencoba beberapa nilai k dan memilih yang menghasilkan akurasi tertinggi.

HASIL DAN PEMBAHASAN

Dalam penelitian ini, data latih diperoleh dari data penilaian siswa SDN Baru 06 Pagi Pasar Rebo yang telah melalui proses seleksi sesuai kebutuhan penelitian. Total data latih yang digunakan berjumlah 244 data siswa yang telah memiliki label kelas Berprestasi dan Tidak Berprestasi. Setiap data siswa terdiri atas atribut nilai, absensi, sikap, dan ekstrakulikuler yang digunakan sebagai dasar penentuan klasifikasi. Atribut tersebut dipilih berdasarkan kriteria yang digunakan oleh pihak sekolah dalam menentukan siswa berprestasi. Rincian data latih ditampilkan pada table 4.1 berikut.

Tabel 1. Data Latih

No	Nama	Nilai	Absensi	Sikap	Ekstrakulikuler	Label
1	Adam Faiz Al Arkhan	77	91	D	Aktif	Tidak Berprestasi
2	Adam Nur Pratama	88	98	C	Sangat Aktif	Tidak Berprestasi
3	Adara Fitri Ramadhani	89	94	B	Aktif	Berprestasi
4	Adelia Putri	81	94	B	Aktif	Tidak Berprestasi
5	Adiba Syakila Az Zahra	81	99	C	Aktif	Tidak Berprestasi
..	…….	…	…	…	…	……...
..	…….	…	…	…	…	……...
..	…….	…	…	…	…	……...
242	Zein Sykar Pinontoan	87	94	C	Aktif	Tidak Berprestasi
243	Ziandra Karnavian Arifin	88	98	B	Aktif	Berprestasi
244	Zivana Aqilla Putri Rizdha	83	98	C	Aktif	Tidak Berprestasi

Data uji memiliki atribut yang sama dengan data latih, yaitu nilai akademik, absensi, sikap, dan kegiatan ekstrakurikuler. Perbedaannya terletak pada fungsi data tersebut. Data latih digunakan sebagai referensi dalam proses klasifikasi, sedangkan data uji digunakan untuk mengetahui hasil prediksi yang dihasilkan oleh algoritma K-Nearest Neighbor.

Tabel 2. Data Uji

No	Nama	Nilai	Absensi	Sikap	Ekstrakulikuler
1	Adelia Kirana	90	100	B	Sangat Aktif
2	Adinda Firadita	86	100	B	Aktif
3	Adinda Nayla Qodriyah	89	100	A	Sangat Aktif
..	…….	…	…	…	…
60	Widya Ash Shafar	89	100	B	Sangat Aktif
61	Zaidan Nathan Atallah	82	99	C	Aktif
62	Zaidan Yusuf Ruhimat	84	97	A	Aktif

Normalisasi data bertujuan untuk menyamakan skala nilai dari seluruh atribut sehingga tidak ada atribut yang mendominasi proses perhitungan jarak pada algoritma KNN. Proses normalisasi dilakukan pada empat parameter: nilai, absensi, sikap, dan ektrakulikuler.
Contoh 10 siswa uji:

Tabel 3. Contoh 10 Sampel Siswa Uji yang Akan Dihitung

No	Siswa Uji	Nilai	Absensi	Sikap	Ekstrakulikuler
1	Adelia Kirana	90	100	B (3)	Sangat Aktif (2)
2	Adinda Firadita	86	100	B (3)	Aktif (1)

3	Adinda Nayla Qodriyah	89	100	A (4)	Sangat Aktif (2)
…	……	….	…..	….	…
8	Ardina Ludy Kurnia Putri	86	99	B (3)	Sangat Aktif (2)
9	Arganta Deandre Wijaya	91	97	B (3)	Aktif (1)
10	Arsyad Ananta Amrullah	84	100	B (3)	Aktif (1)

Dari seluruh data latih, didapat hasil normalisasi untuk data uji:
Nilai *min* : 71 (Putri Adibah)
Nilai *max* : 93 (Banyak siswa)
Normalisasi nilai:

1. Adelia Kirana : X_normalisasi = (90-71)/(93-71) = 0.8636
2. Adinda Firadita : X_normalisasi = (86-71)/(93-71) = 0.6818
3. Adinda Nayla Qodriyah : X_normalisasi = (89-71)/(93-71) = 0.8182

…… : ……

9. Arganta Deandre Wijaya : X_normalisasi = (91-71)/(93-71) = 0.9091
10. Arsyad Ananta Amrullah : X_normalisasi = (84-71)/(93-71) = 0.5909

Dilakukan untuk semua data uji, sehingga didapat hasil normalisasi seperti pada tabel 4.

Tabel 4. Hasil Normalisasi 10 Data Uji

No	Nama Siswa	Nilai	Absensi	Sikap	Ekstrakurikuler
1	Adelia Kirana	0,8636	1,0000	0,6667	1,0000
2	Adinda Firadita	0,6818	1,0000	0,6667	0,0000
3	Adinda Nayla Qodriyah	0,8182	1,0000	1,0000	1,0000
…	…….	……..	……..	…….	……..
9	Arganta Deandre Wijaya	0,9091	0,9400	0,6667	0,0000
10	Arsyad Ananta Amrullah	0,5909	1,0000	0,6667	0,0000

Perhitungan *Euclidean Distance*
Setelah data dinormalisasi, selanjut nya menghitung jarak antara setiap data uji dan seluruh data latih menggunakan rumus *Euclidean.*
Contoh perhitungan yang diambil dari data uji yang bernama Adelia Kirana dengan perhitungan normalisasi yang dihasilkan: nilai = 0,8636; absensi = 1,0000; sikap = 0,6667; Ekstrakulikuler = 1,0000.
Normalisasi siswa latih 1-5 dan 240-244 dari data latih.

Tabel 5. Normalisasi Data Latih

No	Nama Siswa	Nilai	Absensi	Sikap	Ekstrakurikuler
1	Adam Faiz Al Arkhan	77	91	D (1)	Aktif (1)
2	Adam Nur Pratama	88	98	C (2)	Sangat Aktif (2)
3	Adara Fitri Ramadhani	89	94	B (3)	Aktif (1)
…	…	…	…	…	…
…	…	…	…	…	…

No	Nama Siswa	Nilai	Absensi	Sikap	Ekstrakurikuler
243	Ziandra Karnavian Arifin	88	98	B (3)	Aktif (1)
244	Zivana Aqilla Putri Rizdha	83	98	C (2)	Aktif (1)

Nilai *min* : 71
Nilai *max* : 93
Normalisasi nilai:

4. Adam Faiz Al Arkhan : X_normalisasi = (77-71)/(93-71) = 0.2727
5. Adam Nur Pratama : X_normalisasi = (88-71)/(93-71) = 0.7727
6. Adara Fitri Ramadhani : X normalisasi = (89-71)/(93-71) = 0.8182
……. : ……
243. Ziandra Karnavian Arifin : X_normalisasi = (88-71)/(93-71) = 0.7727
244. Zivana Aqilla Putri Rizdha : X_normalisasi = (83-71)/(93-71) = 0.5455

Dengan perhitungan yang sama didapat tabel 6.

Tabel 6. Hasil Normalisasi Data Latih

No	Nama Siswa	Nilai	Absensi	Sikap	Ekstrakurikuler
1	Adam Faiz Al Arkhan	0,2727	0,8200	0,0000	0,0000
2	Adam Nur Pratama	0,7727	0,9600	0,3333	1,0000
3	Adara Fitri Ramadhani	0,8182	0,8800	0,6667	0,0000
…	…	…	…	…	…
…	…	…	…	…	…
243	Ziandra Karnavian Arifin	0,7727	0,9800	0,6667	0,0000
244	Zivana Aqilla Putri Rizdha	0,5455	0,9600	0,3333	0,0000

Dilanjutkan perhitungan normalisasi nya hingga seluruh 244 data latih. Maka *euclidean distance* yang didapat antara Adelia Kirana dengan 244 data latih adalah sebagai berikut:

1) Adam Faiz Al Arkhan (0,2727; 0,8200; 0,0000; 0,0000)

$$d = \sqrt{(0.8636 - 0{,}2727)^2 + (1 - 0.8200)^2 + (0.6667 - 0)^2 + (1 - 0)^2}$$

$$= \sqrt{0.3492 + 0.0324 + 0.4445 + 1}$$

$$= \sqrt{1.8261} = 1.3513$$

…. …….. ……

240) Zivana Aqilla Putri Rizdha(0,5455; 0,9600; 0,3333; 0,0000)

$$d = \sqrt{(0.8636 - 0.5455)^2 + (1 - 0.9600)^2 + (0.6667 - 0.3333)^2 + (1 - 0)^2}$$
$$= \sqrt{0.1012 + 0.0016 + 0.1112 + 1}$$
$$= \sqrt{1.2140} = 1.1018$$

Proses perhitungan jarak *Euclidean* yang sama dilakukan terhadap seluruh 244 data latih untuk masing-masing 62 data uji.

Selanjutnya adalah tahap Perankingan dan Voting. Setelah seluruh nilai *Euclidean Distance* antara data uji dan data latih dihitung, selanjutnya mengurutkan nilai jarak tersebut dari yang terkecil hingga terbesar.
Pada penelitian ini digunakan nilai k = 5, sehingga lima data latih dengan nilai jarak terkecil dipilih sebagai tetangga terdekat (*nearest neighbor*) yang akan digunakan dalam proses penentuan kelas.

Dengan data uji Adelia Kirana, maka di dapat perankingan sesuai tabel 7.

Tabel 7. 5 Tetangga Terdekat untuk Adelia Kirana

Ranking	Nama Siswa	Jarak	Label
1	Dhafin Athallah Aldian	0,0497	Berprestasi
2	Safinatunnajah Zain Al-Alawi	0,0497	Berprestasi
3	Yudanta Rais Arkarna	0,0497	Berprestasi
4	Muhammad Habibi	0,0605	Berprestasi
5	Bilqis Humairah Nurmakin	0,0993	Berprestasi

Dengan data uji lain yang bernama Adinda Firadita dan perhitungan KNN seperti data uji Adelia Kirana, maka di dapat tetangga terdekat Adinda Firadita sesuai tabel 8.

Tabel 8. 5 Tetangga Terdekat untuk Adinda Firadita

Ranking	Nama Siswa	Jarak	Label
1	Muhamad Zulfiqar Abdul Wahab	0,0000	Berprestasi
2	Sabda Aryo Jati	0,0000	Berprestasi
3	Arfan Fathurohman	0,0400	Berprestasi
4	Muhammad Mujiburrohman	0,0400	Berprestasi

Berdasarkan nilai TP, TN, FP, dan FN yang telah diperoleh dari *Confusion Matrix*, selanjutnya dilakukan perhitungan metrik evaluasi model yang meliputi *Accuracy, Precision, Recall, dan F1-Score.*

Tabel 9. Evaluasi Model KNN

Metrik	Rumus	Hasil Perhitungan
Accuracy	(TP + TN) / (TP + TN + FP + FN)	(43 + 10) / (43 + 10 + 9 + 0) = 53 / 62 = 0,8548 = 85,48%
Precision	*TP / (TP + FP)*	43 / (43 + 9) = 43 / 52 = 0,8269 = 82,69%
Recall	TP / (TP + FN)	43 / (43 + 0) = 43 / 43 = 1,0000 = 100,00%
F1-Score	2 × (Presisi × Recall) / (Presisi + Recall)	2 × (0,8269 × 1,0000) / (0,8269 + 1,0000) = 1,6538 / 1,8269 = 0,9053 = 90,53%

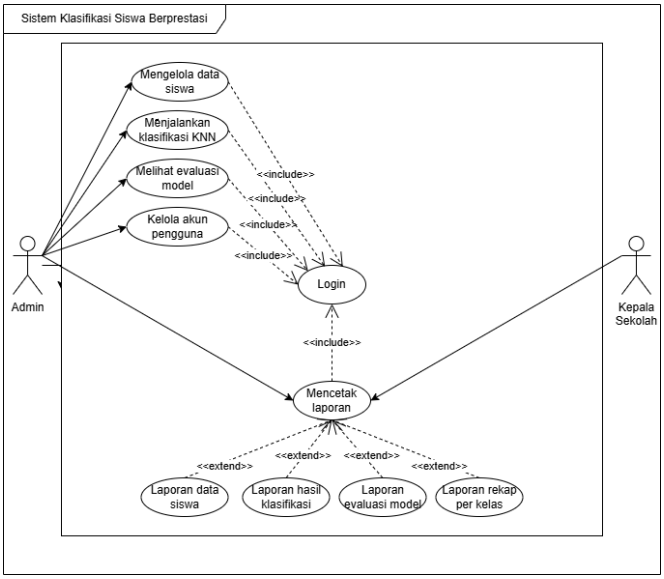
Berdasarkan hasil evaluasi pada Tabel 4…., model KNN dengan nilai K = 5 yang diterapkan pada penelitian ini menghasilkan:

1. *Accuracy* sebesar 85,48%. Nilai ini telah melampaui batas minimum akurasi yang ditetapkan dalam penelitian, yaitu ≥80%.

2. *Precision* sebesar 82,69% mengindikasikan bahwa terdapat sejumlah kecil siswa yang seharusnya masuk kategori Tidak Berprestasi namun diklasifikasikan sebagai Berprestasi oleh model, yang tercermin dari delapan kasus *False Positive*.
3. *Recall* yang mencapai 100% menunjukkan bahwa model berhasil mendeteksi seluruh siswa yang benar-benar Berprestasi tanpa ada yang terlewatkan.
4. *F1-Score* sebesar 90,53% menunjukkan keseimbangan yang baik antara Presisi dan *Recall*.

Dapat disimpulkan bahwa model KNN yang dibangun memiliki kinerja yang baik dalam melakukan klasifikasi siswa berprestasi di SDN Baru 06 Pagi Pasar Rebo.

Pemodelan perangkat lunak pada penelitian ini digunakan untuk menggambarkan rancangan sistem secara visual sebelum diimplementasikan ke dalam bentuk kode program. Pemodelan dilakukan menggunakan notasi *Unified Modeling Language* (UML) yaitu :

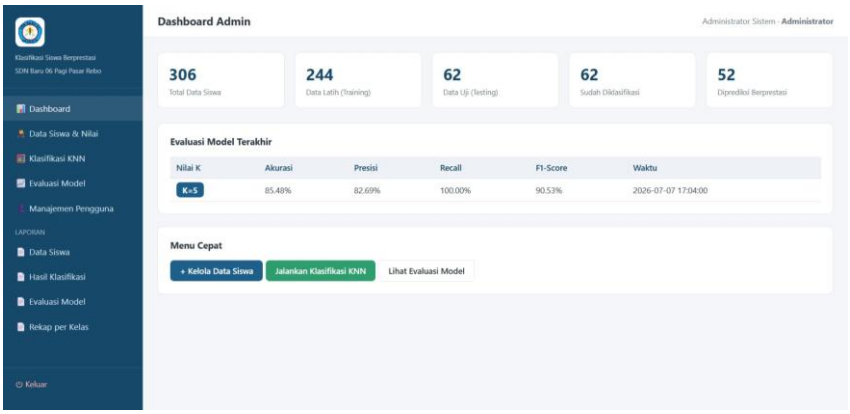


Gambar 1. *Use Case Diagram*

Use case diagram pada Gambar 1 menggambarkan interaksi antara aktor dengan Sistem Klasifikasi Siswa Berprestasi Menggunakan Metode *K-Nearest Neighbor* (KNN). Sistem ini melibatkan dua aktor, yaitu Admin dan Kepala Sekolah, yang memiliki hak akses berbeda sesuai dengan tugas dan tanggung jawabnya.

Activity diagram klasifikasi KNN menggambarkan alur proses Admin dalam menjalankan algoritma *K-Nearest Neighbor*.

Dashboard merupakan halaman utama setelah pengguna berhasil melakukan *login*. Halaman ini menyajikan informasi ringkas mengenai sistem, seperti total data siswa, jumlah data latih (training), jumlah data uji (*testing*), jumlah data yang telah diklasifikasikan, serta jumlah siswa yang diprediksi berprestasi. Selain itu, *dashboard* juga menampilkan hasil evaluasi model klasifikasi terbaru yang meliputi nilai K, akurasi, presisi, *recall*, dan *F1-score*. Pada bagian bawah tersedia menu cepat yang memudahkan Admin untuk mengakses fitur Kelola Data Siswa, Jalankan Klasifikasi KNN, dan Lihat Evaluasi Model secara langsung. Tampilan layar *Dashboard* dapat dilihat pada Gambar 2.



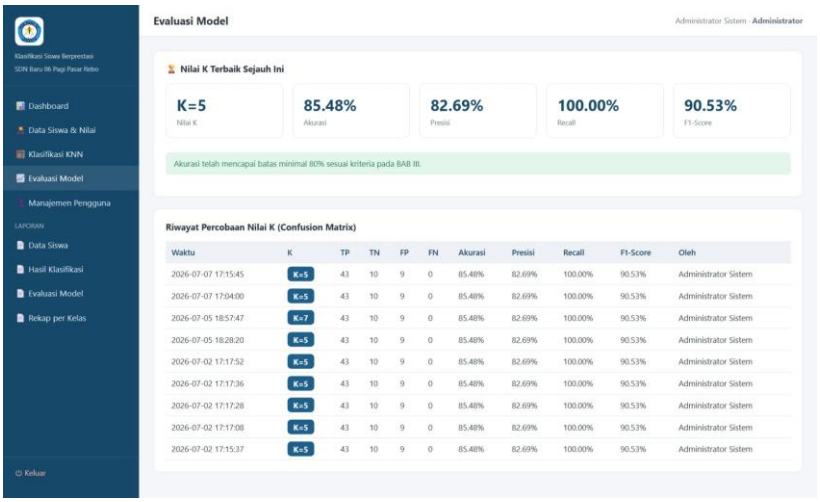
Gambar 2. Tampilan Layar *Dashboard*

Tampilan Laporan Hasil digunakan untuk menyajikan hasil prediksi siswa berdasarkan metode *K-Nearest Neighbor* (KNN) dalam format yang siap dicetak atau disimpan sebagai PDF. Tampilan tersebut dapat dilihat pada Gambar 3.



Gambar 3. Tampilan Layar Laporan Hasil Klasifikasi

Halaman Evaluasi Model digunakan untuk menampilkan hasil evaluasi kinerja metode *K-Nearest Neighbor* (KNN). Pada bagian atas halaman ditampilkan nilai K terbaik beserta nilai akurasi, presisi, *recall*, dan *F1-score* yang diperoleh dari proses klasifikasi. Informasi tersebut membantu pengguna dalam menentukan nilai K yang memberikan performa terbaik. Selain itu, halaman ini juga menampilkan riwayat percobaan nilai K dalam bentuk tabel yang berisi waktu pengujian, nilai K, *Confusion Matrix* (TP, TN, FP, dan FN), nilai akurasi, presisi, *recall*, *F1-score*, serta pengguna yang menjalankan proses evaluasi. Dengan demikian, Admin dapat membandingkan hasil setiap percobaan dan mengevaluasi performa model klasifikasi yang digunakan. Tampilan layar Evaluasi Model dapat dilihat pada Gambar 4.



Gambar 4. Tampilan Layar Evaluasi Model

SIMPULAN

Aplikasi ini memudahkan pihak sekolah dalam mendokumentasikan hasil klasifikasi melalui empat jenis laporan cetak resmi, sehingga dapat langsung digunakan sebagai arsip maupun bahan pengambilan keputusan sekolah secara objektif.

Metode KNN dengan nilai K=5 menghasilkan *accuracy* 85,48%, *precision* 82,69%, *recall* 100%, dan *F1-Score* 90,53%, seluruhnya melampaui batas minimum akurasi yang ditetapkan (≥80%), sehingga terbukti tepat digunakan sebagai dasar pendukung keputusan klasifikasi siswa berprestasi.

DAFTAR PUSTAKA

Ardiaz Fahrezi Maulana Rusli. (2025). *Penerapan Data Mining untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor pada Toko Fatih*. https://library.unindra.ac.id/skripsi/index.php?author=&colltype=0&gmd=Teknik+Informatika&isbn=&location=0&page=37&search=search&searchtype=advance&subject=&title=

Barus, R., Fardila, A., Zulaikha, S., & Takdir, M. (2025). Peran Teknologi Informasi dalam Pengambilan Keputusan Strategis di Lembaga Pendidikan: Kajian Sistematis. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, *3*(4).

Husein, A. M., & Hutauruk, R. E. H. (2022). Penerapan Algoritma C4.5 Dalam Pemilihan Siswa Berprestasi di SMPN 10 Medan. *Digital Transformation Technology*, *2*(1), 8–11.

Kurniawan, R., Hidayat, T., & Pratama, D. (2022). Implementasi Metode K-Nearest Neighbor dalam Klasifikasi Data Akademik Siswa. *Jurnal Sistem Informasi*, *14*(1), 88–97.

Nofriansyah, D., & Nurcahyo, G. W. (2021). *Algoritma Data Mining dan Pengujian*. Deepublish.

Noviyanto, A., Auliya, R., & Rahmat, B. (2021). Klasifikasi Data Nasabah Menggunakan Metode K-Nearest Neighbor pada Perusahaan Asuransi. *JURIKOM: Jurnal Riset Komputer*, *8*(3), 113–119.

Nur Mauliza. (2025). *Sistem Klasifikasi Data Penjualan Barang di Zahra Market Menggunakan Metode Naive Bayes dan K-Nearest Neighbor (KNN)*. https://rama.unimal.ac.id/id/eprint/10054/

Rahayu, N. D., Anshor, A. H., & Afriantoro, I. (2024). Penerapan Data Mining untuk Pemetaan Siswa Berprestasi Menggunakan Metode Clustering K-Means. *JUKI: Jurnal Komputer Dan Informatika*, *6*(1).

Ramadhani, S., & Sensuse, D. I. (2023). Perbandingan Algoritma Klasifikasi Data Mining untuk Prediksi Prestasi Belajar Mahasiswa. *Jurnal Ilmu Komputer Dan Informasi*, *16*(1), 21–30.

Wahyuni, E. S., & Wahyudi, U. (2019). Pengujian Akurasi Algoritma K-Nearest Neighbor (KNN) pada Data Nilai Siswa. *Jurnal Nasional Informatika Dan Teknologi Jaringan*, *4*(1), 10–16.