

IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBORS DALAM MENENTUKAN KATEGORI PRESTASI SISWA BERDASARKAN NILAI AKADEMIK

Jordan Ryano Fadillah^{1*}, Rizky Pradana²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ^{1*}2011500028@student.budiluhur.ac.id, ²rizky.pradana@budiluhur.ac.id
(* : corresponding author)

Abstrak- Madrasah Tsanawiyah Annajah, sebuah sekolah Islam swasta tingkat menengah pertama, saat ini belum banyak menerapkan digitalisasi dalam operasionalnya. Sistem penilaian dan ranking yang digunakan saat ini mengandalkan perhitungan manual di spreadsheet, yang dikelompokkan berdasarkan kelas dan disimpan di direktori guru. Dokumen tersebut kemudian dilaporkan kepada orang tua siswa sebagai feedback. Penelitian ini mengaplikasikan algoritma K-Nearest Neighbor (K-NN) untuk memprediksi peringkat siswa dan membandingkannya dengan peringkat semester sebelumnya. K-NN adalah metode algoritma terbimbing yang mengklasifikasikan sampel pengujian baru ke dalam kelas berdasarkan kedekatannya dengan data yang sudah ada. Sistem penilaian manual yang ada saat ini tidak hanya memakan waktu, tetapi juga rentan terhadap kesalahan manusia, yang membatasi kemampuan sekolah dalam menganalisis data akademik secara mendalam dan cepat. Penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menyediakan metode yang lebih efisien dan akurat untuk memprediksi peringkat siswa. Hasil penelitian menunjukkan bahwa untuk semester ganjil, akurasi cross-validation adalah 76.6%, precision, recall, dan f1-score masing-masing 77.6%. Sedangkan untuk semester genap, akurasi cross-validation adalah 75.2%, dengan precision, recall, dan f1-score masing-masing 75.2%.

Kata Kunci: Algoritma K-Nearest Neighbor, Peringkat, Cross-Validation accuracy, Precision, Recall, F1_score

IMPLEMENTATION OF THE K-NEAREST NEIGHBORS ALGORITHM IN DETERMINING STUDENT ACHIEVEMENT CATEGORIES BASED ON ACADEMIC GRADES

Abstract- Madrasah Tsanawiyah Annajah, a private Islamic junior high school, currently has not widely implemented digitalization in its operations. The existing grading and ranking system relies on manual calculations in spreadsheets, which are grouped by class and stored in teachers' directories. These documents are then reported to students' parents as feedback. This research applies the K-Nearest Neighbor (K-NN) algorithm to predict student rankings and compare them with the previous semester's rankings. K-NN is a supervised learning algorithm that classifies new test samples into classes based on their proximity to existing data. The current manual grading system is not only time-consuming but also prone to human error, which limits the school's ability to analyze academic data deeply and quickly. This study aims to address these issues by providing a more efficient and accurate method for predicting student rankings. The research results show that for the odd semester, cross-validation accuracy is 76.6%, with precision, recall, and f1-score each at 77.6%. For the even semester, cross-validation accuracy is 75.2%, with precision, recall, and f1-score each at 75.2%.

Keywords: K-Nearest Neighbor Algorithm, Ranking, Cross-Validation accuracy, Precision, Recall, F1_score

1. PENDAHULUAN

Kinerja siswa dalam pendidikan merupakan indikator pengajaran yang efektif. Salah satu tantangannya adalah penilaian yang akurat terhadap kinerja siswa di kurikulum siswa. Analisis data telah menjadi alat penting untuk menganalisis dan memprediksi kinerja siswa. Madrasah Tsanawiyah Annajah adalah salah satu sekolah tingkat menengah pertama Islam swasta yang masih belum banyak menerapkan digitalisasi dalam operasional mereka.

Klasifikasi dapat dijelaskan sebagai pembelajaran terawasi memprediksi label kelas, digunakan untuk menemukan model yang menjelaskan atau membedakan kelas data dan memprediksi kelas dari data yang belum diketahui[1]. Tujuan dari penelitian ini adalah menggunakan KNN dalam konteks pendidikan untuk lebih memahami pola perilaku siswa dan faktor-faktor yang mempengaruhinya.

Penelitian ini berfokus pada Implementasi algoritma K-Nearest Neighbor dalam menentukan nilai kategori prestasi siswa berdasarkan nilai akademik. Algoritma ini dipilih karena kemampuannya untuk mengklasifikasikan kemiripan atribut yang dimiliki data siswa, data nilai dan data peringkat siswa. Namun pendekatannya memiliki

pembatasan, yaitu cenderung tidak mempertimbangkan variabel lain yang mempengaruhi siswa. seperti faktor sosial, faktor ekonomi dan lain – lain. Hasil dari algoritma ini seringkali bergantung pada pemilihan parameter, seperti nilai k yang dapat memengaruhi nilai akurasi. Oleh karena itu penelitian ini juga mengeksplorasi berbagai parameter dan teknik untuk meningkatkan akurasi prediksi peringkat kategori prestasi siswa.

Berdasarkan penelitian sebelumnya, penulis memakai metode KNN untuk mencari, membandingkan, menganalisis perbedaan metode atau perubahan algoritma yang digunakan pada penelitian sebelumnya terkait dengan penelitian penulis pada kategori siswa menurut aktivitas sekolah. peneliti [10] menggunakan metode KNN untuk memprediksi kelulusan siswa.

Proses pertama dimulai dengan definisi sistem dan evaluasi serta pemodelan menggunakan data eksperimen yang telah dipilih sebelumnya dari aplikasi tools orange. KNN dikaitkan dengan alat prediksi dan matriks ketidakpastian. Nanti penulis bandingkan dengan [10]. Berdasarkan perbandingan pada hasil penelitian sebelumnya, memprediksi kelulusan siswa yang dikembangkan dalam riset ini memiliki perbedaan pada langkah penggunaan tools pada proses algoritma K-Nearest Neighbor saja, untuk proses selanjutnya memiliki langkah-langkah yang mirip. Manfaat dari penelitian ini adalah sekolah dan guru dapat lebih cepat dan akurat mengidentifikasi siswa yang membutuhkan dukungan tambahan.

2. METODE PENELITIAN

2.1 Data Mining

menurut Gartner Group, merupakan tahapan menganalisis banyak data dalam database untuk menemukan pola, hubungan, dan tren dengan menggunakan teknik matematika dan statistik daripada metode markup. Penambangan data memungkinkan Anda menghasilkan nilai tambahan yang berasal dari pengetahuan yang belum Anda ketahui sebelumnya [4].

2.2 K-Nearest Neighbor

K-NN adalah metode yang sangat mudah digunakan dan efektif untuk data yang lebih besar. Ini juga dapat mengklasifikasikan data dengan cepat [3]. Algoritma K-NN juga merupakan teknik yang digunakan untuk mengklasifikasikan data sampel pengujian ke dalam kelas K-NN berdasarkan data atribut dan data latih. Algoritma ini menggunakan ingatan pola jawaban dan titik uji untuk menemukan nilai K suatu benda di dekat titik latihan. Pengklasifikasi menggunakan suara mayoritas di antara objek K dan menggunakan klasifikasi natural sebagai nilai prediksi sampel uji baru. Jarak Euclidean digunakan untuk menghitung tetangga terdekat dan terjauh, sehingga kita dapat mengelompokkan siswa ke dalam kelas yang berbeda dan membantu guru merancang program yang lebih spesifik [5]. Data mining telah digunakan dalam beberapa bidang, seperti pendidikan, pemasaran, teknik, keuangan [2].

2.3 Confusion Matrix

Adalah matriks yang menyediakan campuran contoh kelas yang diprediksi vs. aktual. Hal ini memungkinkan definisi berbagai metrik kinerja (misalnya, akurasi, presisi) atau teknik seperti karakteristik operasi penerima. [8].

2.3.1 Presisi (*Precision*)

dengan membagi prediksi positif sejati dengan jumlah perkiraan positif sejati dan positif palsu. Metrik ini menghitung keakuratan prediksi kelas positif dengan mencari kasus positif yang antisipasinya benar [6].

$$Recall = \frac{true\ positives}{True\ positives + False\ Positive} \quad (1)$$

2.3.2 Recall (*True Positive*)

adalah ukuran kemampuan mengklasifikasi untuk menemukan semua yang relevan dalam data. Metrik ini sangat penting karena tujuannya adalah meminimalkan negatif palsu, memastikan Sebagian besarditangkap oleh modelnya [7].

$$Recall = \frac{true\ positives}{True\ positives + False\ Negative} \quad (2)$$

2.3.3 *True Negative*

Adalah ketika model mengidentifikasi sampel negatif dengan benar, artinya contoh yang tidak termasuk dalam kelas positif diklasifikasikan dengan benar [9].

2.4 Praproses data

Langkah – Langkah praproses adalah :

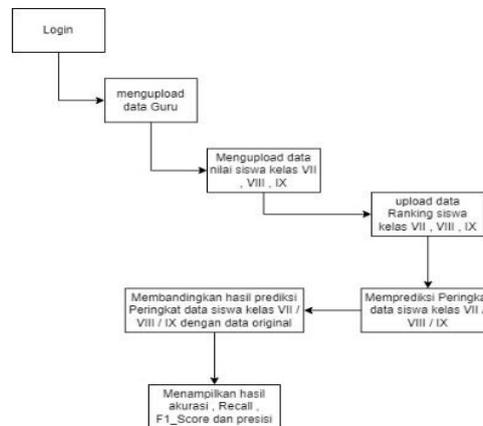
- Pengumpulan data : Dalam penelitian ini data yang digunakan menganalisis kategori prestasi siswa dikumpulkan dari 1 sumber data. Data dari nilai mata Pelajaran dalam satu semester dan nilai peringkat siswa dalam satu semester. Data nilai mata Pelajaran berisi NIS, Nisn, Nama, JK, dan mata pelajarannya. Untuk data peringkat berisi NIS, Nama, JK, jumlahUts, jumlahUas, SUM, AVARAGE dan peringkat. Pengumpulan data dilakukan dengan mengakses database dan mempunyai data excel yang digunakan nantinya untuk mengupload ke dalam aplikasi
- Normalisasi data : Untuk menghindari dominasi atribut tertentu yang mungkin mempengaruhi hasil analisis, nilai numerik dinormalisasi untuk memastikan bahwa semua atribut memiliki skala yang sama.
- Pembagian data : data dibagi menjadi dua. Seperti data testing dan data training. Data training digunakan untuk melatih model. Model akan belajar dari data ini untuk memahami pola dan hubungan. sedangkan data testing untuk menguji model setelah pelatihan. Data ini tidak digunakan selama pelatihan, tetapi digunakan untuk mengukur kinerja model pada data yang baru dipelajari.

2.5 Implementasi program

Tahapan ini, dilakukan pembuatan Penerapan Metode, rancangan menu, rancangan , serta pembuatan program pada website.

2.5.1 Penerapan Metode

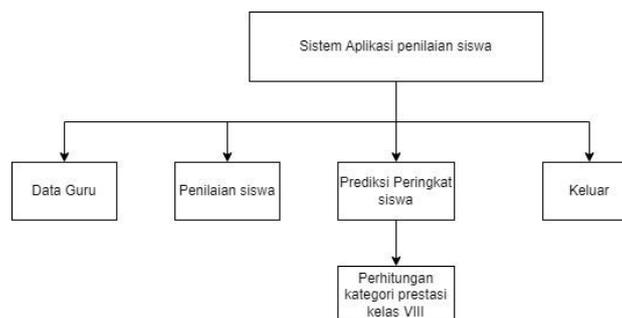
Pada penelitian ini, dalam implementasi algoritma K-Nearest Neighbor untuk memprediksi peringkat siswa terdapat beberapa tahap yang akan dijalankan. Pada gambar 1 merupakan tahapan yang digunakan.



Gambar 1. Penerapan Metode

2.5.2 Rancangan Menu

Rancangan menu ini merupakan tahapan bagaimana menyusun menu yang ingin dibangun, seperti yang terlihat pada Gambar 2:



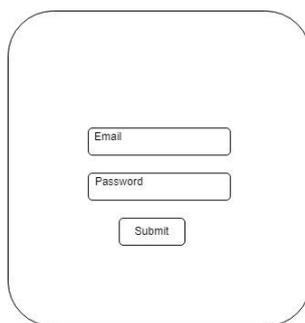
Gambar 2. Rancangan Menu

2.5.3 Rancangan Layar

merupakan sangat penting untuk pembuatan aplikasi dengan berbagai antarmuka. Demikian rancangan layar harus bisa dipahami oleh pengguna untuk alasan efektivitas dan efisiensi, pengguna tidak akan tertarik dengan tampilan layar aplikasi yang berantakan. Berikut adalah rancangan layar sistem yang akan dibuat:

a. Rancangan Layar Halaman Login

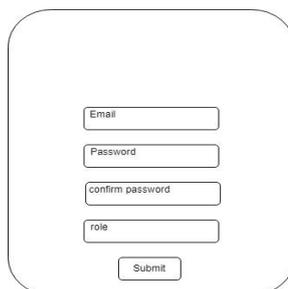
Pada Gambar 3 merupakan rancangan layar yang akan ditampilkan pertama kali dimana *User* dapat melihat serta Mengisi Email dan password.


 A wireframe of a login page. It features a rounded rectangular container with three input fields: 'Email', 'Password', and a 'Submit' button below them.

Gambar 3. Rancangan Layar Halaman Login

b. Rancangan Layar Halaman Register

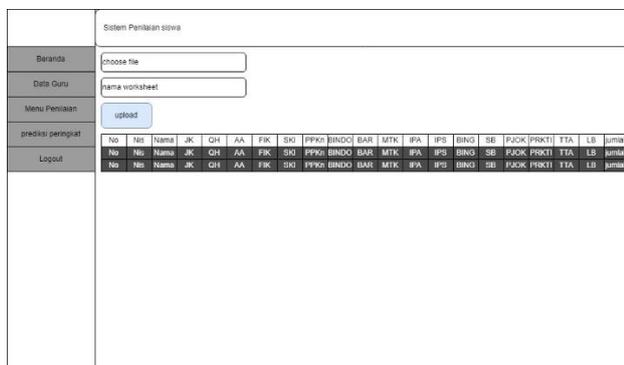
Pada Gambar 4, merupakan rancangan layar yang tampilan dipergunakan untuk mengisi data Email, Password, confirm password dan rolenya:


 A wireframe of a registration page. It features a rounded rectangular container with four input fields: 'Email', 'Password', 'confirm password', and 'role', followed by a 'Submit' button.

Gambar 4. Rancangan Layar Halaman Register

c. Rancangan Layar Halaman Penilaian Siswa

Pada Gambar 5, merupakan tampilan dimana *Guru* dapat melihat dan mengupload data excel nilai siswa. berikut ini Menu Penilaian siswa:


 A screenshot of a web application interface for student evaluation. It includes a sidebar menu with options like 'Beranda', 'Data Guru', 'Menu Penilaian', 'prediksi pernapas', and 'Logout'. The main content area is titled 'Sistem Penilaian siswa' and contains two input fields: 'choose file' and 'nama worksheet', followed by an 'upload' button. Below the input fields is a table with a header row and a data row, both containing a long list of subject abbreviations.

Gambar 5. Rancangan Layar Halaman Data Penilaian Siswa

d. Rancangan Layar Halaman Ranking

Tampilan pada Gambar 6, merupakan tampilan dimana Guru dapat upload, mengisi worksheet yang akan ditampilkan dan sistem akan melakukan perhitungan menggunakan algoritma K-NN.

Sistem Penilaian siswa																									
Beranda	<input type="text" value="choose file"/>																								
Data Guru	<input type="text" value="nama worksheet"/>																								
Menu Penilaian	<input type="button" value="upload"/>																								
prediksi peringkat	<table border="1"> <thead> <tr> <th>No</th> <th>Nis</th> <th>Nama</th> <th>JK</th> <th>kelas</th> <th>semester</th> <th>jumlahUts</th> <th>jumlahJas</th> <th>SUM</th> <th>AVARAGE</th> <th>Peringkat</th> <th>Predictedperingkat</th> </tr> </thead> <tbody> <tr> <td>No</td> <td>Nis</td> <td>Nama</td> <td>JK</td> <td>kelas</td> <td>semester</td> <td>jumlah Uts</td> <td>jumlahJas</td> <td>SUM</td> <td>AVARAGE</td> <td>Peringkat</td> <td>Predictedperingkat</td> </tr> </tbody> </table>	No	Nis	Nama	JK	kelas	semester	jumlahUts	jumlahJas	SUM	AVARAGE	Peringkat	Predictedperingkat	No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	SUM	AVARAGE	Peringkat	Predictedperingkat
No	Nis	Nama	JK	kelas	semester	jumlahUts	jumlahJas	SUM	AVARAGE	Peringkat	Predictedperingkat														
No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	SUM	AVARAGE	Peringkat	Predictedperingkat														
Logout	<table border="1"> <thead> <tr> <th>No</th> <th>Nis</th> <th>Nama</th> <th>JK</th> <th>kelas</th> <th>semester</th> <th>jumlah Uts</th> <th>jumlahJas</th> <th>BAR</th> <th>AVARAGE</th> <th>Peringkat</th> <th>Predictedperingkat</th> </tr> </thead> <tbody> <tr> <td>No</td> <td>Nis</td> <td>Nama</td> <td>JK</td> <td>kelas</td> <td>semester</td> <td>jumlah Uts</td> <td>jumlahJas</td> <td>BAR</td> <td>AVARAGE</td> <td>Peringkat</td> <td>Predictedperingkat</td> </tr> </tbody> </table>	No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	BAR	AVARAGE	Peringkat	Predictedperingkat	No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	BAR	AVARAGE	Peringkat	Predictedperingkat
No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	BAR	AVARAGE	Peringkat	Predictedperingkat														
No	Nis	Nama	JK	kelas	semester	jumlah Uts	jumlahJas	BAR	AVARAGE	Peringkat	Predictedperingkat														

Gambar 6. Rancangan Layar Halaman Data Penilaian Siswa

3. HASIL DAN PEMBAHASAN

3.1 Algoritma Data prediksi

Merupakan serangkaian langkah atau teknik penyelesaian masalah yang tersusun dengan baik. Berikut prediksi garis model menggunakan metode K-NN.

Algoritma 3.1. Algoritma Data Prediksi

```

1 BEGIN
2 IMPORT PhpSpreadsheet library
3 DEFINE FUNCTION euclideanDistance(data1, data2)
4 SET distance TO 0.0
5 FOR i FROM 0 TO LENGTH(data1) - 1
6 ADD (data1[i] - data2[i])^2 TO distance
7 END FOR
8 RETURN SQUARE ROOT OF distance
9 END FUNCTION
10 DEFINE FUNCTION knnPredict(trainingData, testData, k)
11 INITIALIZE empty list distances
12 FOR EACH trainRow IN trainingData
13 CALCULATE distance BETWEEN trainRow features AND testData using
euclideanDistance
14 ADD [distance, label] TO distances
15 END FOR
16 SORT distances BY distance
17 SELECT top k elements FROM distances AS neighbors
18 COUNT labels IN neighbors AND STORE IN labelsCount
19 RETURN the label WITH MAXIMUM count in labelsCount
20 END FUNCTION
21 DEFINE FUNCTION normalizeData(data)
22 INITIALIZE empty list normalizedData
23 SET columns TO NUMBER OF columns IN data
24 FOR i FROM 0 TO columns - 1
25 GET columnData AS values OF column i FROM data
26 SET min TO MINIMUM OF columnData
27 SET max TO MAXIMUM OF columnData
28 FOR EACH value IN data
29 CALCULATE normalized value AS (value[i] - min) / (max - min)
30 ADD normalized value TO normalizedData[i]
31 END FOR
32 END FOR
33 FOR EACH row IN data
34 COPY non-normalized columns (NIS, Nama, JK, Kelas, Semester, Peringkat) TO
normalizedData
35 END FOR
36 RETURN normalizedData
37 END FUNCTION
38 DEFINE FUNCTION crossValidation(data, k, folds)
39 SET foldSize TO SIZE OF data DIVIDED BY folds
40 SET accuracySum TO 0
41 INITIALIZE truePositives, falsePositives, falseNegatives, trueNegatives TO 0
42 FOR i FROM 0 TO folds - 1
43 SET testData TO SLICE of data FROM i * foldSize TO (i + 1) * foldSize
44 SET trainData TO REMAINDER OF data NOT IN testData
45 SET correctPredictions TO 0
46 FOR EACH testRow IN testData
47 GET testFeatures FROM testRow
48 SET predictedLabel TO knnPredict(trainData, testFeatures, k)
49 IF predictedLabel == ACTUAL label IN testRow
50 INCREMENT correctPredictions AND truePositives
51 ELSE
52 INCREMENT falsePositives AND falseNegatives
53 END IF
54 END FOR
55 ADD correctPredictions / SIZE OF testData TO accuracySum
56 END FOR
57 CALCULATE precision AS truePositives / (truePositives + falsePositives)
58 CALCULATE recall AS truePositives / (truePositives + falseNegatives)
59 CALCULATE f1Score AS 2 * (precision * recall) / (precision + recall)
60 RETURN accuracySum / folds * 100, precision, recall, f1Score

```

3.2 Langkah Pengujian

3.2.1 Implementasi Program

a. Tampilan Login

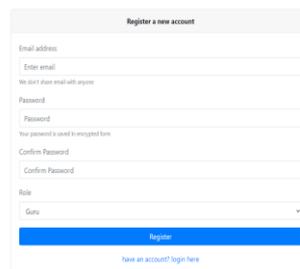
Pada Gambar 7 merupakan halaman login berisi *Email* dan *Password*, pengguna memasukkan *email* dan *password* agar memasuki sebagai guru.



Gambar 7. Halaman Login

b. Tampilan Register

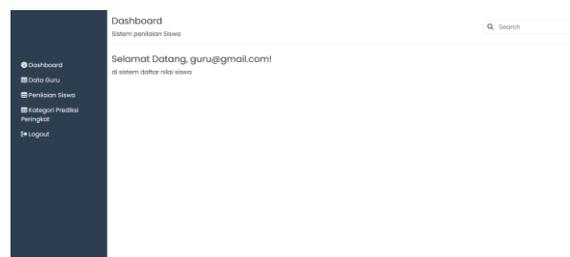
Pada Gambar 8 merupakan tampilan halaman register berisi *Email*, *Password*, *Confirmed password* untuk memastikan *password*nya dan memasuki *role* yaitu sebagai guru. *User* memasukkan *email*, *password*, *confirmed password* dan *role* yang benar agar bisa memasuki sebagai guru.



Gambar 8. Halaman Register

c. Tampilan beranda Guru

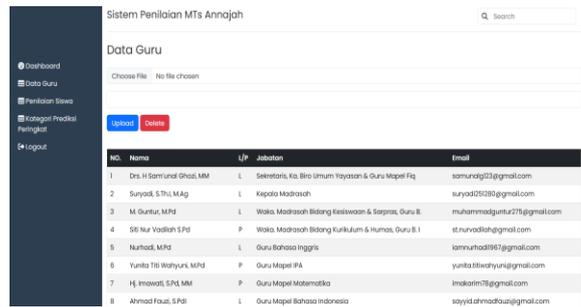
Pada Gambar 9 merupakan halaman beranda guru. Di halaman ini guru dapat memilih menu data guru, data nilai dan kategori prediksi siswa.



Gambar 9. Halaman Beranda Guru

d. Tampilan Data Guru

Pada Gambar 10 merupakan halaman data guru yang berisi No., Nama, L/P, Jabatan dan Email.

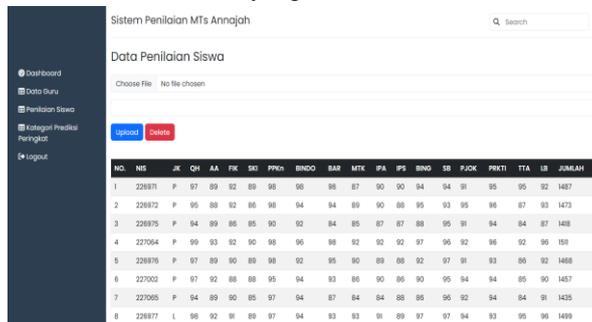


NO.	Nama	L/P	Jabatan	Email
1	Drs. H Samudir Ghaz, MM	L	Selektaris, Ka. Bina Utama Yayasan & Guru Mapel PA	samudir2@gmail.com
2	Suryadi, S.Thi, M.Ag	L	Kepala Madrasah	suryadi2529@gmail.com
3	M. Guntur, MPH	L	Waka. Madrasah Bidang Kesiswaan & Sarpras, Guru B.I	muhammadguntur275@gmail.com
4	Siti Nur Yadhik S Pd	P	Waka. Madrasah Bidang Kurikulum & Humas, Guru B.I	sitnurvadiah@gmail.com
5	Nurhadi, MPH	L	Guru Bahasa Inggris	iamnurhadi91@gmail.com
6	Yunika Tri Wahyuni, MPH	P	Guru Mapel IPA	yunika.triwahyuni@gmail.com
7	H. Imawati, S.Pd, MM	P	Guru Mapel Matematika	imawati7@gmail.com
8	Ahmad Fauzi, S.Pd	L	Guru Mapel Bahasa Indonesia	sayyid.ahmadfauzi@gmail.com

Gambar 10. Halaman Data Guru

e. Tampilan Data Nilai

Pada Gambar 11 merupakan halaman data nilai yang berisi data nilai siswa setiap semester.

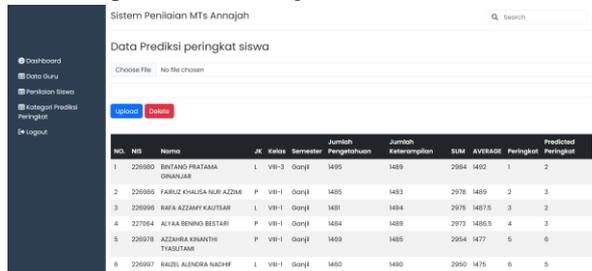


NO.	NIK	JK	QH	AA	FK	SD	PPKn	BNDO	BAR	MTK	IPA	IPS	BING	SB	FJOK	PKRT	ITA	LE	JUMLAH
1	226871	P	97	89	92	89	96	98	96	87	90	90	94	94	91	95	95	92	1487
2	226872	P	95	88	92	86	98	94	84	89	90	88	95	93	95	96	87	93	1473
3	226875	P	94	89	86	85	90	92	84	85	87	87	88	95	91	94	84	87	1488
4	227064	P	99	93	92	90	96	96	98	92	92	92	97	96	92	96	92	96	158
5	226876	P	97	89	90	88	98	92	95	90	89	88	92	97	91	93	88	92	1458
6	227002	P	97	92	88	88	95	94	93	86	90	88	93	94	94	85	90	1457	
7	227005	P	94	89	90	85	97	94	87	84	84	88	86	90	92	94	84	91	1425
8	226877	L	98	92	91	89	97	94	93	93	91	89	97	97	94	93	95	96	1499

Gambar 10. Halaman Data Nilai

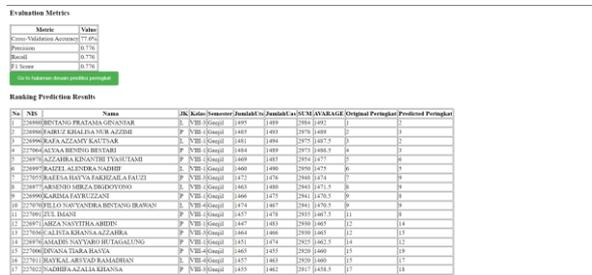
f. Tampilan Data Prediksi

Pada Gambar 11 merupakan halaman data ranking yang berisi data ranking. Untuk gambar 12 dan 13 ada tampilan hasil evaluasi metrik dan tampilan data ranking.



NO.	NIK	Nama	JK	Kelas	Semester	Jumlah Peringkat	Jumlah Keterampilan	SUM	AVERAGE	Peringkat	Predicted Peringkat
1	226980	BHITANG PRATAMA GUNANTAR GUNANAR	L	VIII-3	Ganjil	1495	1489	2984	1492	1	2
2	226989	FABUL KHALISA NURI AZZAH AZZAHAR	P	VIII-1	Ganjil	1485	1483	2970	1485	2	3
3	226998	RAFA AZZAHY KAUSAR	L	VIII-1	Ganjil	1481	1484	2976	1487.5	3	3
4	227064	ALYIA BENING BESTARI	P	VIII-1	Ganjil	1484	1489	2973	1486.5	4	3
5	226978	AZZAHBA KURNATHI TIASUTAMA	P	VIII-1	Ganjil	1469	1485	2954	1477	5	6
6	226987	RAULI ALENORA NADHF	L	VIII-1	Ganjil	1460	1480	2920	1475	6	5

Gambar 11. Halaman Data Ranking



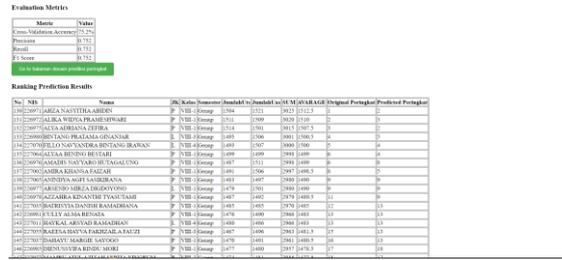
Evaluation Metrics

Metric	Value
Cronbach's Alpha	0.7746
Reliability	0.776
P2	0.776

Ranking Prediction Results

No	NIK	Nama	JK	Kelas	Semester	Jumlah Peringkat	Jumlah Keterampilan	SUM	AVERAGE	Original Peringkat	Predicted Peringkat
1	226980	BHITANG PRATAMA GUNANTAR	L	VIII-3	Ganjil	1495	1489	2984	1492	1	2
2	226989	FABUL KHALISA NURI AZZAH	P	VIII-1	Ganjil	1485	1483	2970	1485	2	3
3	226998	RAFA AZZAHY KAUSAR	L	VIII-1	Ganjil	1481	1484	2976	1487.5	3	3
4	227064	ALYIA BENING BESTARI	P	VIII-1	Ganjil	1484	1489	2973	1486.5	4	3
5	226978	AZZAHBA KURNATHI	P	VIII-1	Ganjil	1469	1485	2954	1477	5	6
6	226987	RAULI ALENORA NADHF	L	VIII-1	Ganjil	1460	1480	2920	1475	6	5
7	227002	YUNIKA TRI WAHYUNI	P	VIII-1	Ganjil	1473	1478	2946	1474	7	6
8	227005	YUNIKA TRI WAHYUNI	P	VIII-1	Ganjil	1463	1488	2941	1471.5	8	9
9	226986	ARABIA FATMATEZZAN	P	VIII-1	Ganjil	1466	1473	2939	1475.5	9	8
10	227001	LO SAVANORA BHITANG BRANAN	L	VIII-1	Ganjil	1474	1487	2941	1476.5	9	9
11	227003	DELLA BAHRI	P	VIII-1	Ganjil	1477	1478	2940	1474.5	11	8
12	226979	AMBA NUSYTRA ARDIN	P	VIII-1	Ganjil	1467	1483	2930	1465	12	14
13	227004	ALYIA KRANSIA AZZAHRA	P	VIII-1	Ganjil	1484	1486	2939	1463	12	13
14	226985	SAYURAH SUTAYORU TRIANGSUNG	P	VIII-1	Ganjil	1451	1474	2925	1462.5	14	12
15	227063	WIDYANITA HARAS RAISA	P	VIII-1	Ganjil	1465	1453	2928	1460	15	19
16	227007	IRUKA KRETIYOP KAMARISMA	L	VIII-1	Ganjil	1457	1489	2928	1460	15	17
17	227008	ADRIANA AZALIA KRANSIA	P	VIII-1	Ganjil	1453	1482	2917	1458.5	17	18

Gambar 12. Halaman Data Ranking Semester Ganjil



Evaluation Metrics

Metric	Value
Cross-Validation Accuracy	75.2%
Precision	0.76
Recall	0.752
F1 Score	0.757

Ranking Prediction Results

No.	NIS	Nama	JK	Kelas	Semester	Jumlah Uts	Jumlah Uas	SUM	AVARAGE	Original Peringkat	Predicted Peringkat
1	20220001	JARZA NAYITHA ARDIEN	P	VIII	Genap	1204	1202	1212.3	1	1	2
2	20220002	TALHA WITSA PRANIS PERDIBI	P	VIII	Genap	1211	1209	1209.5	2	2	1
3	20220003	SALIA ABRARZA ZIFRA	P	VIII	Genap	1214	1201	1207.5	3	3	2
4	20220004	RIYAN PRALAMA GUNASARI	L	VIII	Genap	1203	1206	1204.5	4	4	3
5	20220005	WIBLO NAYANDRA BINTANG BAWAN	L	VIII	Genap	1203	1207	1205	5	5	4
6	20220006	MALINDA RIZKA BELAL	P	VIII	Genap	1203	1203	1203	6	6	5
7	20220007	MALINDA NAYANDRA TALALUNG	P	VIII	Genap	1203	1211	1204	7	7	6
8	20220008	JARZA KHUSNA FALAM	P	VIII	Genap	1203	1208	1205.5	8	8	7
9	20220009	ANINDYA AGNI LASIRABANA	P	VIII	Genap	1203	1207	1204	9	9	8
10	20220010	RIYAN HEBAL ARIYANTO	L	VIII	Genap	1203	1203	1203	10	10	9
11	20220011	ALZABRA KRANTIR YUANYANI	P	VIII	Genap	1203	1202	1202.5	11	11	10
12	20220012	BACHRYA DANIS RAMADHANA	P	VIII	Genap	1203	1203	1203	12	12	11
13	20220013	EVELY ALBA RIZKA	P	VIII	Genap	1203	1209	1206	13	13	12
14	20220014	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1206	1204.5	14	14	13
15	20220015	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1207	1209	1208	15	15	14
16	20220016	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1203	1203	16	16	15
17	20220017	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1203	1203	17	17	16
18	20220018	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1203	1203	18	18	17
19	20220019	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1203	1203	19	19	18
20	20220020	RIYAN ARYAN RAMADHAN	L	VIII	Genap	1203	1203	1203	20	20	19

Gambar 13. Halaman Data Ranking Semester Genap

3.2.2 Hasil riset

Pambar 14 dan 15 pengujian dan hasil sistem digunakan untuk mengetahui, menganalisa dan evaluasi apakah kumpulan nilai cross-validation, precision, recall dan f1_score. Algoritma KNN digunakan untuk menguji dan mengetahui prediksi peringkat siswa.

1	Yang pertama membuat file excel dan isi dengan No, NIS, Nama, JK, kelas, semester, jumlahUts, jumlahUas, SUM, AVARAGE, Peringkat dan Predictedperingkat. Untuk SUM menggunakan rumus SUM yang di excel misalkan : =SUM(G8:H8), jumlahUts, jumlahUas itu ada di kolom G ke 8 dan jumlah Uas ada di kolom H ke 8. AVARAGE mirip bedanya AVERAGE(G8:H8). Untuk Peringkat rumusnya =RANK(J8,\$J\$8:\$J\$136,0).
2	Yang kedua login sebagai guru, memilih menu prediksi peringkat, pilih file excel dan ketik nama worksheet yang akan di tampilkan, setelah itu klik upload button
3	Setelah di klik sistem akan langsung melakukan fungsi upload file, melakukan perhitungan K-Nearest Neighbor.

Gambar 14. Langkah Hasil Prediksi Peringkat

No.	Cuplikan kode	Proses perhitungan
1	<pre>// Variabel untuk menghitung metrik evaluasi \$truePositives = 0; >falsePositives = 0; >falseNegatives = 0; // Melakukan cross-validation dengan membagi data ke dalam beberapa fold for (\$i = 0; \$i < \$folds; \$i++) { // Pemisahan data latihan dan data uji \$testData = array_slice(\$data, \$i * \$foldSize, \$foldSize); \$trainData = array_diff_key(\$data, array_flip(array_keys(\$testData))); foreach (\$testData as \$testRow) { \$testFeatures = array_slice(\$testRow, 0, 4); \$predictedLabel = knnPredict(\$trainData, \$testFeatures, \$k); // Menghitung True Positives, False Positives, dan False Negatives if (\$predictedLabel == end(\$testRow)) { \$truePositives++; } else { \$falsePositives++; \$falseNegatives++; } } }</pre>	<p>Mencari dan mengumpulkan data untuk menghitung True Positives (TP), False Positives (FP), dan False Negatives (FN) selama proses cross-validation</p>
2	<pre>Rumus untuk perhitungan Precision dan f1_score : // Menghitung metrik evaluasi \$precision = \$truePositives / (\$truePositives + \$falsePositives); \$recall = \$truePositives / (\$truePositives + \$falseNegatives); \$f1Score = 2 * ((\$precision * \$recall) / (\$precision + \$recall));</pre>	<p>Precision = $\frac{TP}{TP+FP}$</p> <p>Precision = $\frac{76}{76+24} = 0.76 = 76\%$</p> <p>Recall = $\frac{TP}{TP+FN}$</p> <p>Recall = $\frac{76}{76+25} = 0.752 = 75.2\%$</p> <p>F1_score = $2 \times \frac{precision \times recall}{precision + recall}$</p> <p>F1_score = $2 \times \frac{0.76 \times 0.752}{0.76+0.752} = 2 \times \frac{0.57152}{1.512}$</p> <p>= 0.756 = 75.6%</p> <p>berdasarkan jumlah TP, FP, dan FN yang dihitung selama proses cross-validation. Precision dan recall menunjukkan performa model dalam mendeteksi kelas positif dengan tepat dan benar. F1-score memberikan keseimbangan antara precision dan recall.</p>

Gambar 15. Langkah Hasil Prediksi Proses Perhitungan Peringkat

4. KESIMPULAN

Berdasarkan hasil uji dari prediksi peringkat siswa dengan dengan metode algoritma K-NN dapat disimpulkan bahwa Berdasarkan data yang ada dari data nilai: ini data kelas VIII.1 sampai VIII.4 yang nilai pengetahuan dan keterampilan , data Peringkat(NIS, Nama, Nama, JK, jumlah, jumlah, SUM, AVARAGE, Peringkat). Dari hasil prediksi peringkat yang didapatkan kan adalah untuk semester ganjil Cross-validation Accuracy adalah dengan nilai 76%, untuk Precision dengan nilai 76% , untuk nilai recall dengan nilai 76% dan untuk f1-score nilainya adalah 76 %. Untuk semester genap Cross-validation Accuracy nilainya adalah 75.2%, Precision dengan nilai 75.2%, recall untuk nilainya adalah 75.2% dan untuk f1-score nilainya adalah 75.2%. Untuk hasil riset ini menunjukkan bahwa algoritma knn dapat menghasilkan nilai prediksi dan menghasilkan kinerja yang memenuhi persyaratan yang diberikan serta lebih efisien dibandingkan dengan cara manual.

Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi algoritma lain seperti algoritma *Support Vector Machine* dan lain – lain. Guna membandingkan kinerja atau akurasi. Penelitian lebih lanjut juga perlu mempertimbangkan parameter dan penerapan modal di lingkungan pendidikan untuk memastikan keefektifan dan relevansinya.

DAFTAR PUSTAKA

- [1] D. Prasetyawan, R. Gatra, K. Kunci, P. Akademik, and K. Pearson, "Algoritma K-Nearest Neighbor untuk Memprediksi Prestasi Mahasiswa Berdasarkan Latar Belakang Pendidikan dan Ekonomi," *Jurnal Informatika Sunan Kalijaga*, vol. 7, no. 1, 2022.
- [2] W. Lastari, "Penerapan Data Mining Untuk Memprediksi Prestasi Siswa SMA Pada Dinas Pendidikan Provinsi Jambi," vol. 8, no. 2, 2023.
- [3] S. Nuraeni, S. P. A. Syam, M. F. Wajdi, B. Firmansyah, and M. Malkan, "Implementasi Metode K-NN Untuk Menentukan Jurusan Siswa di SMAN 02 Manokwari," *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 1, pp. 89–95, 2023, doi: 10.33379/gtech.v7i1.1905.
- [4] S. Widaningsih and S. Yusuf, "Penerapan Data Mining Untuk Memprediksi Siswa Berprestasi Dengan Menggunakan Algoritma K Nearest Neighbor," *Jurnal Teknik Informatika Dan Sistem Informasi*, vol. 9, no. 3, 2022.
- [5] D. Supriadi, E. Suryana, J. Meranti Raya Nomor, and S. Lebar Bengkulu, "Penerapan K-Nearest Neighbor Untuk Klasifikasi Tingkat Kelulusan Pada Siswa SMA Negeri 11 Kota Bengkulu," *Jurnal Media Infotama*, vol. 19, no. 2, 2023.
- [6] K. Taunk, S. De, S. Verma, and A. Swetapadma, "Optimization Strategies for the k-Nearest Neighbor Classifier," *SN Computer Science*, vol. 2, no. 4, 2021.
- [7] J. Chung and K. Lee, "Credit Card Fraud Detection: An Improved Strategy for High Recall Using KNN, LDA, and Linear Regression," *Sensors*, vol. 23, no. 18, pp. 7788, 2023.
- [8] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, "Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem," *Technologies*, vol. 9, no. 4, 2021, doi: 10.3390/technologies9040081.
- [9] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of k-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Sci. Rep.*, vol. 12, no. 1, pp. 1-11, 2022, doi: 10.1038/s41598-021-03233-3.
- [10] S. Kasus, K. Barat, Desti, P. Sari, S. Shofia Hilabi, and A. Hananto, "Penerapan Data Mining Metode K-Nearest Neighbor Untuk Memprediksi Kelulusan Siswa Sekolah Menengah Pertama," *SMARTICS J.*, vol. 9, no. 1, pp. 14-19, 2023, doi: 10.21067/smartics.v9i1.8088.