

Machine Learning

## Penerapan Algoritma *K-Nearest Neighbor* dalam Sistem Absensi untuk Klasifikasi Kehadiran Siswa pada Pondok Pesantren Mawaridussalam

Alwi Maulana Siregar \*, Khairuddin Nasution, Antoni

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Islam Sumatera Utara, Medan, Indonesia

### INFORMASI ARTIKEL

Diterima Redaksi: 23 Juli 2024  
Revisi Akhir: 03 September 2024  
Diterbitkan Online: 24 Maret 2025

### KATA KUNCI

Algoritma *K-Nearest Neighbor* (KNN)  
Sistem Absensi  
Klasifikasi Kehadiran  
Pondok Pesantren Mawaridussalam

### KORESPONDENSI (\*)

Phone: +62 813-2870-2932  
E-mail: [aagyms10042000@gmail.com](mailto:aagyms10042000@gmail.com)

### A B S T R A K

Penelitian ini dilatarbelakangi oleh permasalahan absensi siswa di Pondok Pesantren Mawaridussalam yang masih dilakukan secara manual, sehingga sering terjadi kesalahan dalam pencatatan kehadiran siswa. Penelitian ini bertujuan untuk mengembangkan sistem absensi berbasis algoritma *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan kehadiran siswa. Proses penelitian dimulai dengan observasi dan wawancara di Pondok Pesantren Mawaridussalam untuk mengumpulkan data absensi siswa. Data yang dikumpulkan meliputi nama, NISN, dan catatan kehadiran siswa. Sistem yang dirancang menggunakan algoritma KNN untuk mengklasifikasikan kehadiran siswa ke dalam tiga kategori: Rajin, Cukup Rajin, dan Tidak Rajin. Hasil penelitian menunjukkan bahwa algoritma KNN efektif dalam mengklasifikasikan data absensi siswa. Pada pengujian sistem dengan 10 sampel siswa, salah satu siswa bernama Aziz Guntara dengan jumlah kehadiran 17, sakit 1, izin 0, dan absen 0, dikategorikan sebagai Rajin. Hasil klasifikasi ini membantu guru dalam memonitor kehadiran siswa secara lebih akurat dan efisien, serta mengurangi kesalahan pencatatan yang sering terjadi pada sistem manual.

### PENDAHULUAN

Permasalahan utama yang dihadapi Pondok Pesantren Mawaridussalam saat ini adalah pendataan kehadiran santri yang masih dilakukan secara manual. Proses manual ini tidak hanya memakan waktu dan tenaga, tetapi juga rentan terhadap kesalahan pencatatan dan pengelolaan data yang tidak efisien. Kesulitan dalam mengelola data kehadiran secara tepat waktu dan akurat dapat berdampak negatif terhadap efektivitas pembelajaran dan manajemen pendidikan di pesantren. Maka, penelitian ini bertujuan untuk menerapkan sistem absensi *online* di Pondok Pesantren Mawaridussalam untuk mempermudah proses absensi para santri, sistem absensi ini akan menggunakan algoritma *K-Nearest Neighbor* (KNN) untuk mempermudah mengklasifikasikan absensi para santri. Algoritma *K-Nearest Neighbor* (KNN) telah terbukti efektif dalam berbagai aplikasi klasifikasi, termasuk dalam pengelompokan data kehadiran santri. Dengan menerapkan algoritma KNN dalam sistem absensi, Pondok Pesantren Mawaridussalam dapat mengoptimalkan proses pemantauan kehadiran santri secara otomatis melalui *website* tanpa harus menggunakan kertas. Penggunaan sistem absensi ini akan dilakukan melalui *website*, guru mata pelajaran dapat menggunakan sistem absensi ini untuk memantau kehadiran para santri yg hadir dan tidak.

Kelebihan dari KNN yaitu dapat menghasilkan data yang kuat atau jelas dan efektif jika digunakan pada data yang besar. Sedangkan kekurangan dari KNN yaitu membutuhkan nilai  $k$  sebagai parameter, jarak dari data percobaan tidak dapat jelas dengan tipe jarak yang digunakan dan dengan atribut yang digunakan untuk memperoleh hasil terbaik [1].

Namun, meskipun konsep dan potensi aplikasi algoritma KNN dalam sistem absensi sudah terbukti, implementasi yang tepat dalam konteks Pondok Pesantren Mawaridussalam memerlukan penyesuaian yang spesifik. Faktor seperti pola kehadiran santri di pesantren perlu dipertimbangkan dengan cermat.

Dengan menggabungkan kebutuhan praktis dari Pondok Pesantren Mawaridussalam dengan keunggulan teknologi algoritma KNN, penelitian ini bertujuan untuk merancang dan mengimplementasikan sebuah sistem absensi yang canggih dan handal. Diharapkan sistem ini tidak hanya dapat mengoptimalkan pengelolaan kehadiran santri secara efisien, tetapi juga dapat memberikan kontribusi positif terhadap pengalaman pendidikan santri di Pondok Pesantren Mawaridussalam.

## TINJAUAN PUSTAKA

### *Algoritma*

Algoritma berasal dari kata algoris dan ritmis yang pertama kali diperkenalkan oleh Abu Ja'far Muhammad Ibn Musa Al Khwarizmi pada 825 M di dalam buku *Al-Jabr Wa-al Muqabla*. Dalam bidang pemrograman, algoritma didefinisikan sebagai metode yang terdiri dari serangkaian langkah yang terstruktur dan sistematis untuk menyelesaikan masalah dengan bantuan komputer [3]. Algoritma menurut [4] adalah suatu upaya dengan urutan operasi yang disusun secara logis dan sistematis untuk menyelesaikan suatu masalah untuk menghasilkan suatu output tertentu.

### *K-Nearest Neighbor (KNN)*

*K-Nearest Neighbor* merupakan salah satu metode untuk mengambil keputusan menggunakan pembelajaran terawasi dimana hasil dari data masukan yang baru diklasifikasi berdasarkan terdekat dalam data nilai. Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek yang berdasarkan dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN merupakan algoritma *supervised learning* dimana hasil dari *query instance* yang baru diklasifikasi berdasarkan mayoritas dari kategori pada algoritma KNN. Dimana kelas yang paling banyak muncul yang nantinya akan menjadi kelas hasil dari klasifikasi [5].

### *Euclidean Distance*

Kedekatan didefinisikan dalam jarak metrik, seperti jarak *Euclidean*. Proses perhitungan dalam Algoritma k-NN ini merupakan perhitungan *Euclidean Distance* yaitu metode pencarian antara dua titik variabel, semakin dekat dan mirip maka semakin kecil jarak antara dua titik tersebut. *Euclidean Distance* dikatakan baik jika data baru memiliki jarak minimum dan memiliki kemiripan yang tinggi [6]. Jarak *Euclidean* dapat dicari dengan menggunakan persamaan berikut ini.

$$D_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

### *Absensi*

Menurut [7] "Absensi adalah sebuah pembuatan data untuk daftar kehadiran yang bisa digunakan bagi sebuah lembaga atau instansi yang sangat perlu membutuhkan sistem seperti ini".

Menurut [8] "Absensi merupakan daftar hadir sekumpulan orang dari suatu kelompok orang banyak yang tergabung dalam sebuah instansi secara resmi yang mempunyai peraturan-peraturan, ketentuan-ketentuan, serta batasan-batasan, dan orang-orang yang terlibat di dalamnya terikat oleh peraturan tersebut. Jika sekelompok orang tersebut melanggarnya maka akan dikenakan sanksi sebagai hukuman dari pelanggaran yang dilakukan oleh orang tersebut sesuai dengan peraturan dan ketentuan yang diterapkan oleh instansi tersebut".

### *PHP*

"PHP (*HyperText PreProcessor*) merupakan bahasa pemrograman yang di proses di server, Fungsi utama PHP dalam membangun website adalah untuk melakukan pengelolaan data dalam database" [9].

"PHP (*HyperText PreProcessor*) merupakan suatu bahasa pemrograman yang digunakan untuk menerjemahkan baris kode program menjadi kode mesin yang dapat di mengerti oleh komputer yang bersifat serverside yang dapat di tambahkan ke dalam HTML" [10].

## METODOLOGI

### *Teknik Pengumpulan Data*

Teknik pengumpulan data yang digunakan pada penelitian ini, yaitu:

1. Observasi  
Penelitian ini dilakukan dimana peneliti meninjau langsung bagaimana para guru melakukan kegiatan absensi untuk memulai pembelajaran.
2. Studi Literatur  
Studi literatur atau studi kepustakaan dilakukan dengan mengambil data dari buku, jurnal ataupun artikel mengenai teori dan langkah-langkah dalam pembuatan aplikasi. Sehingga dapat menggunakannya sebagai dasar landasan teori dan perancangan pembuatan aplikasi sistem absensi untuk klasifikasi kehadiran santri.
3. Wawancara  
Melakukan tanya jawab dan interaksi langsung dengan guru di Pondok Pesantren Mawaridussalam mengenai sistem absensi santri yang ada di tempat.

### *Implementasi Algoritma K-Nearest Neighbor*

1. Menentukan Kategori.  
Sebelum masuk ke tahapan perhitungan manual algoritma, tahap pertama yang harus dilakukan adalah menentukan kategori yang akan digunakan sebagai acuan klasifikasi. Pada penelitian ini, penulis menggunakan tiga kategori klasifikasi yaitu:

Tabel 1. Tabel Kategori

No	Kategori
1	Rajin
2	Cukup Rajin
3	Tidak Rajin

2. Input Data  
Setelah mengetahui kategori yang akan digunakan, tahap selanjutnya adalah melakukan penginputan data sampel (data *training*) dan data *testing*. Tahap ini dilakukan sebagai tahap awal untuk melakukan perhitungan manual algoritma.

#### a. Penginputan Data *Training*

Data *training* adalah data sampel atau data percobaan yang menjadi acuan klasifikasi algoritma *K-Nearest Neighbor* dalam melakukan perhitungan data *testing* nantinya. Data *training* didapat berdasarkan data absensi tahun sebelumnya. Data ini didapat dari data kehadiran santri pada tahun 2023 di Pondok Pesantren Mawaridussalam.

Tabel 2. Tabel Data *Training*

No	Nama	H	S	I	A	Kategori
1	Santri 1	15	2	1	0	Cukup Rajin
2	Santri 2	16	1	1	0	Rajin
3	Santri 3	15	1	2	0	Cukup Rajin
4	Santri 4	14	2	2	0	Cukup Rajin
5	Santri 5	15	1	1	1	Cukup Rajin
6	Santri 6	12	2	1	3	Tidak Rajin
7	Santri 7	11	2	1	4	Tidak Rajin
8	Santri 8	13	1	1	3	Tidak Rajin
9	Santri 9	18	0	0	0	Rajin
10	Santri 10	16	2	0	0	Rajin
11	Santri 11	15	2	1	0	Cukup Rajin

12	Santri 12	16	1	1	0	Rajin
13	Santri 13	15	1	2	0	Cukup Rajin
14	Santri 14	14	2	2	0	Cukup Rajin
15	Santri 15	15	1	1	1	Cukup Rajin
16	Santri 16	12	2	1	3	Tidak Rajin
17	Santri 17	11	2	1	4	Tidak Rajin
18	Santri 18	13	1	1	3	Tidak Rajin
19	Santri 19	18	0	0	0	Rajin
20	Santri 20	16	2	0	0	Rajin
21	Santri 21	15	2	1	0	Cukup Rajin
22	Santri 22	16	1	1	0	Rajin
23	Santri 23	15	1	2	0	Cukup Rajin
24	Santri 24	14	2	2	0	Cukup Rajin
25	Santri 25	15	1	1	1	Cukup Rajin

Sumber: Pondok Pesantres Mawaridussalam

b. Penginputan Data *Testing*

Data *testing* adalah data sampel yang nilainya akan dicari oleh algoritma. Pada pembahasan ini, data *testing* adalah data santri di Pondok Pesantren Mawaridussalam pada mata pelajaran Bahasa Arab berjumlah 10 orang yang akan digunakan untuk menjadi data percobaan perhitungan manual algoritma *K-Nearest Neighbor*. Data ini didapat dari Pondok Pesantren Mawaridussalam pada mata pelajaran Bahasa Arab.

Tabel 3. Tabel Data *Testing*

No	Nama	Mata Pelajaran	H	S	I	A	Kategori
1	Aziz Guntara	Bahasa Arab	17	1	0	0	?
2	Farhan Fadillah	Bahasa Arab	11	4	2	1	?
3	Arif Alfi Nazri	Bahasa Arab	14	1	2	1	?
4	Aditya Wahyudi	Bahasa Arab	15	2	0	1	?
5	Rayyan Imtiaz Masga	Bahasa Arab	12	4	1	1	?
6	Muhammad Husnul Afif	Bahasa Arab	10	3	3	2	?
7	Muhammad Khairul Izmi	Bahasa Arab	15	2	0	1	?
8	Zahrani Khairullah	Bahasa Arab	16	1	0	1	?
9	Affandi Aulia Sambo	Bahasa Arab	18	0	0	0	?
10	Rizky Septiawan	Bahasa Arab	13	3	1	1	?

3. Menentukan Nilai K

Tahap selanjutnya adalah dengan menentukan nilai K. Nilai K adalah nilai yang berfungsi sebagai titik atau batas nilai yang akan diambil untuk hasil akhir. Penentuan nilai K tidak ada rumus pastinya, tetapi ada tips dalam menentukan nilai K, yaitu:

- Jika jumlah kategori adalah genap, maka sebaiknya  $k = \text{ganjil}$ .
- Jika jumlah kategori adalah ganjil, maka sebaiknya  $k = \text{genap}$ .

Pada penelitian ini jumlah kategori adalah 3 (ganjil), maka nilai  $k$  sebaiknya adalah genap. Dengan demikian, penulis menentukan nilai  $k = 2$ .

4. Menghitung Jarak *Euclidean*

Tahap selanjutnya adalah menghitung jarak *Euclidean* dengan menggunakan rumus pada persamaan 1. Pada perhitungan jarak *Euclidean* ini digunakan data *training* pada tabel 2 dan data *testing* pada tabel 3 sebagai bahan perhitungan yang akan digunakan. Tahap perhitungan yang akan dilakukan adalah mengurangi data *training*

tiap keterangan absensi dengan data *testing* dengan keterangan absensi yang sama, lalu setelah itu dipangkatkan dua.

Rumus perhitungan pada penelitian ini:

$$d = \sqrt{(H_2 - H_1)^2 + (S_2 - S_1)^2 + (I_2 - I_1)^2 + (A_2 - A_1)^2} \quad (2)$$

Keterangan:

$H_2$  = Keterangan hadir data training pada tabel 3.2.

$H_1$  = Keterangan hadir data testing pada tabel 3.3.

$S_2$  = Keterangan sakit data training pada tabel 3.2.

$S_1$  = Keterangan sakit data testing pada tabel 3.3.

$I_2$  = Keterangan izin data training pada tabel 3.2.

$I_1$  = Keterangan izin data testing pada tabel 3.3.

$A_2$  = Keterangan absen data training pada tabel 3.2.

$A_1$  = Keterangan absen data testing pada tabel 3.3.

Untuk perhitungan jarak pada bab ini akan dilakukan dengan menggunakan satu sampel data testing, yaitu data milik Aziz Guntara.

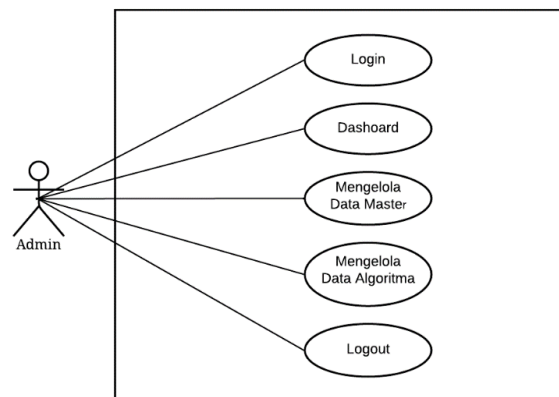
$$d = \sqrt{(15 - 17)^2 + (2 - 1)^2 + (1 - 0)^2 + (0 - 0)^2}$$

$$d = \sqrt{(-2)^2 + 1^2 + 1^2 + 0^2}$$

$$d = \sqrt{6} = 2,45 = 2$$

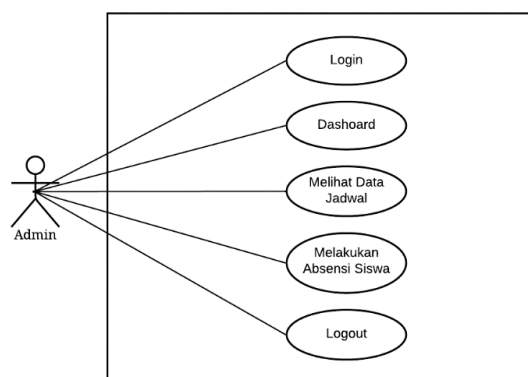
## Perancangan Sistem

### Use Case Diagram



Gambar 1. Use Case Diagram

Pada gambar 1 merupakan use case diagram, admin dapat melakukan serangkaian kegiatan didalam menu admin. Untuk masuk ke dalam menu admin, admin harus melakukan *login* terlebih dahulu dengan memasukkan *username* dan *password* yang sudah ada dan sudah terinput didalam *database*. Setelah berada di halaman admin, admin dapat melihat menu *dashboard* yang ada, lalu admin dapat mengelola data master yang meliputi data admin, data guru dan data santri. Admin juga dapat mengelola data algoritma absensi yang ada.



Gambar 2. Use Case Diagram Guru

Gambar 2 di atas merupakan *use case* diagram guru, dimana guru dapat melakukan beberapa aktivitas didalam halaman guru. Sebelum guru dapat masuk ke dalam halaman guru, guru harus terlebih dahulu melakukan *login*, guru harus memasukkan *username* dan *password* yang sudah dimasukkan admin ke dalam *database* sebelumnya. Setelah masuk ke dalam halaman guru, guru dapat melihat tampilan *dashboard*, dapat melihat jadwal guru mengajar dan dapat melakukan proses absensi santri sesuai dengan mata pelajaran tiap guru.

## HASIL DAN PEMBAHASAN

Pada penelitian ini, penulis sudah menganalisa pengklasifikasian absensi santri Pondok Pesantren Mawaridussalam. Penelitian ini dilakukan selama satu bulan pada bulan Februari 2024 di Pondok Pesantren Mawaridussalam dan menggunakan 10 sampel data *testing* santri. Informasi yang penulis dapatkan berupa nama, nisn dan data absensi santri. Untuk mendapatkan hasil analisa pengklasifikasian absensi santri pada penelitian ini, guru akan melakukan pengisian absensi terhadap santri atau anak didik sesuai dengan mata pelajaran yang di ajar. Untuk memnuhi proses pengklasifikasian, penulis menggunakan tiga kategori pengklasifikasian yaitu Rajin, Cukup Rajin, dan Tidak Rajin. Dimana nanti hasil akhir dari setiap santri yang sudah di absensi akan otomatis memiliki kategori masing-masing sesuai dengan kriteria kehadiran santri tersebut.

Tabel 4. Hasil Data *Testing*

No	Nama	H	S	I	A	Kategori
1	Aziz Guntara	17	1	0	0	Rajin
2	Farhan Fadillah	11	4	2	1	Tidak Rajin
3	Arif Alfi Nazri	14	1	2	1	Cukup Rajin
4	Aditya Wahyudi	15	2	0	1	Cukup Rajin
5	Rayyan Imtiaz Masga	12	4	1	1	Tidak Rajin
6	Muhammad Husnul Afif	10	3	3	2	Tidak Rajin
7	Muhammad Khairul Izmi	15	2	0	1	Cukup Rajin
8	Zahran Khairullah	16	1	0	1	Cukup Rajin
9	Affandi Aulia Sambo	18	0	0	0	Rajin
10	Rizky Septiawan	13	3	1	1	Cukup Rajin

## KESIMPULAN DAN SARAN

Berdasarkan hasil dari 25 sampel data training dan 10 sampel data testing pada penelitian ini, algoritma K-NN memprediksi santri dengan jumlah hadir 17, sakit 1, izin 0, dan absen 0 yang kemudian memiliki nilai  $k = 2$  menghasilkan 2 kategori Cukup Rajin dan 3 kategori Rajin. Hasil akhir algoritma tersebut adalah Cukup Rajin < Rajin, maka santri tersebut masuk ke dalam kategori Rajin. Dan Disarankan untuk melakukan evaluasi secara menyeluruh terhadap implementasi algoritma k-NN pada sistem absensi guna mengidentifikasi area-area yang membutuhkan optimalisasi untuk meningkatkan kinerja dan akurasi sistem.

## DAFTAR PUSTAKA

- [1] Aji, B.A. (2019). Implementasi Metode K - Nearest Neighbor Dalam Pencarian Lokasi Rumah Makan.. Dipetik 02 Maret 2024 dari [https://lib.unnes.ac.id/36881/1/5302412114\\_Optimized.pdf](https://lib.unnes.ac.id/36881/1/5302412114_Optimized.pdf)
- [2] Muhammad Fathurrohman dan Sulistyorini, Implementasi Manajemen Peningkatan Mutu Pendidikan Islam Peningkatan Lembaga Pendidikan Islam Secara Holistik, Yogyakarta: Teras, 2012, pp. 189.
- [3] Jando. E. & Nani. P. Algoritma dan Pemograman Berbasis Java. Yogyakarta: CV Andi, 2018, pp.
- [4] Kani. M. Algoritma dan Pemograman. Tangerang Selatan: Universitas Terbuka, 2020, pp 1-19.
- [5] S. S. Keerthi dan S. K. Shevade. "K-Nearest Neighbor: Metode Klasifikasi untuk Data Berukuran Besar dan Berdimensi Tinggi," Jurnal Ilmiah Komputer dan Informatika, Vol. 21, No. 2, Juni 2023.

- [6] I. Riadi, A. Fadlil, dan P. Annisa, "Identifikasi Tulisan Tangan Huruf Katakana Jepang dengan Metode Euclidean," J-SAKTI (Jurnal Sains Komput. dan Inform., vol. 4, no. 1, p. 29, 2020.
- [7] Santoso dan Yulianto. Analisa Dan Perancangan Sistem Absensi Siswa Berbasis Web Dan Sms Gateway. Jurnal Matrik Vol. 16, No. 2, 2017, ISSN: 1858 – 4144 Hal. 65-75.
- [8] Juhartini. SISTEM INFORMASI ABSENSI SISWA MENGGUNAKAN PHP DAN MYSQL BERBASIS WEB PADA MAN 2 UNGGULAN MATARAM, Vol. 10, No.1, 2020. Pages. 60–64.
- [9] EMS, Tim. "Pengantar PHP dan MySQL". Jakarta : PT.Elex Media Komplitudo, 2012.
- [10] Abdulloh, Rohi. Easy & Simple Web Programing. Jakarta: PT Elex Media Komputindo,2016.