

Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Menganalisis Pendapat Siswa Terhadap Edukasi Anti Perundungan

Petti Indrayati Sijabat¹, Eviyanti Br Barus², Nuraisana³, Yosia Laora S.Lingga⁴, Dwita Amelia Putri⁵

^{1,2,3,4,5}Teknik Informatika, Bisnis Digital, Teknik Informatika, Teknologi Informasi, Teknologi Informasi, STMIK Pelita Nusantara

Email: ¹Pettisijabat.mamaliem@gmail.com, ²Eviyanti94@gmail.com, ³Nuraisana94@gmail.com, ⁴Yosia@gmail.com,

⁵Dwita@gmail.com

Email Penulis Korespondensi: Pettisijabat.mamaliem@gmail.com

Abstrak

Perundungan di lingkungan pendidikan merupakan permasalahan sosial yang berdampak signifikan terhadap perkembangan psikologis dan akademik peserta didik. Edukasi anti perundungan telah banyak diterapkan di sekolah, namun efektivitasnya sering kali sulit diukur secara objektif. Penelitian ini bertujuan untuk menganalisis pendapat siswa terhadap program edukasi anti perundungan dengan menerapkan algoritma K-Nearest Neighbor (KNN) pada data berbasis teks. Data diperoleh melalui kuesioner terbuka yang diisi oleh siswa, kemudian dilakukan tahap preprocessing meliputi tokenizing, stopword removal, dan pembobotan Term Frequency–Inverse Document Frequency (TF-IDF). Kasus perundungan dalam dunia pendidikan telah menjadi isu sosial krusial yang mampu menghambat pertumbuhan akademik serta kesehatan mental siswa. Guna mengevaluasi respons peserta didik terhadap inisiatif pencegahan perundungan, studi ini memanfaatkan algoritma K-Nearest Neighbor (KNN) untuk mengolah data tekstual berisi opini siswa. Melalui teknik pembobotan TF-IDF dan nilai $k = 5$, sistem berhasil mencapai tingkat akurasi sebesar 86,4%, yang membuktikan efektivitas metode ini dalam memetakan kecenderungan opini secara objektif. Temuan ini diharapkan dapat menjadi dasar pengambilan keputusan dalam evaluasi serta pengembangan strategi edukasi anti perundungan di lingkungan sekolah.

Kata Kunci: K-Nearest Neighbor, Analisis Sentimen, Edukasi, Perundungan, Opini Siswa

Abstract

Bullying in educational settings is a social problem that significantly impacts students' psychological and academic development. Anti-bullying education has been widely implemented in schools, but its effectiveness is often difficult to measure objectively. This study aims to analyze students' opinions on anti-bullying education programs by applying the K-Nearest Neighbor (KNN) algorithm to text-based data. Data were obtained through open-ended questionnaires completed by students, followed by preprocessing stages including tokenizing, stopword removal, and Term Frequency–Inverse Document Frequency (TF-IDF) weighting. Bullying cases in education have become a crucial social issue that can hinder students' academic growth and mental health. To evaluate students' responses to bullying prevention initiatives, this study utilized the K-Nearest Neighbor (KNN) algorithm to process textual data containing students' opinions. Through the TF-IDF weighting technique and a k value of 5, the system successfully achieved an accuracy rate of 86.4%, proving the effectiveness of this method in objectively mapping opinion trends. These findings are expected to be the basis for decision-making in evaluating and developing anti-bullying education strategies in school environments.

Keywords: K-Nearest Neighbor, Sentiment Analysis, Education, Bullying, Student Opinion

1. PENDAHULUAN

Perundungan (*bullying*) merupakan salah satu permasalahan sosial yang masih sering terjadi di lingkungan pendidikan. Fenomena ini tidak hanya berdampak pada kondisi psikologis peserta didik, tetapi juga dapat menurunkan motivasi belajar dan prestasi akademik. Upaya pencegahan telah dilakukan melalui berbagai program edukasi anti perundungan, baik secara langsung maupun melalui media digital. Namun, efektivitas dari program tersebut sering kali sulit diukur secara objektif karena bergantung pada persepsi dan pendapat siswa sebagai peserta utama kegiatan edukasi. Permasalahan ini menimbulkan kebutuhan akan suatu pendekatan analitik berbasis data untuk memahami bagaimana siswa memandang edukasi anti perundungan tersebut.

Penelitian ini mengimplementasikan algoritma KNN untuk menganalisis respons siswa terhadap edukasi anti-perundungan melalui pendekatan pembelajaran mesin. KNN digunakan karena keunggulannya dalam melakukan klasifikasi teks berdasarkan kemiripan atribut antar data secara efektif. Proses ini melibatkan pembersihan data teks dan pemanfaatan metode TF-IDF untuk ekstraksi fitur, sehingga melalui optimasi nilai k , model dapat mengategorikan pendapat siswa ke dalam sentimen positif, negatif, dan netral dengan lebih presisi. Penelitian terkait telah banyak dilakukan untuk analisis sentimen berbasis pembelajaran mesin. Misalnya, penelitian oleh (Afriani et al., 2025) [1] menerapkan algoritma Naïve Bayes untuk analisis opini masyarakat terhadap kebijakan pendidikan dengan akurasi 81%. Studi oleh Santoso dan Prasetyo (Albab et al., 2023) [2] mengombinasikan Support Vector Machine dengan TF-IDF untuk analisis

sentimen siswa terhadap pembelajaran daring dengan akurasi 84%. Penelitian lain oleh (Dwi Wicaksono et al., 2024) [3] menggunakan metode Random Forest untuk klasifikasi komentar publik pada kampanye anti perundungan digital. Sementara itu, penelitian oleh (Efendi & Latifah, 2021)[4] menerapkan KNN pada analisis ulasan produk edukatif dengan akurasi 85%. Terakhir, studi terbaru oleh (Fudhail Ferio Supeli & Setiaji, 2023)[5] menunjukkan bahwa kombinasi KNN-TFIDF dapat meningkatkan akurasi analisis opini hingga 87%.

Berdasarkan penelitian-penelitian tersebut, dapat diidentifikasi bahwa sebagian besar fokus pada konteks opini umum atau pembelajaran daring, sedangkan analisis terhadap pendapat siswa mengenai edukasi anti perundungan masih terbatas. Hal ini menjadi research gap yang melatarbelakangi penelitian ini. Oleh karena itu, penelitian ini bertujuan untuk menerapkan algoritma KNN dalam menganalisis pendapat siswa terhadap program edukasi anti perundungan, dengan harapan dapat memberikan hasil klasifikasi yang akurat serta menjadi dasar dalam evaluasi dan pengembangan strategi pencegahan perundungan di sekolah.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan untuk menganalisis pendapat siswa terhadap edukasi anti perundungan dengan menggunakan algoritma K-Nearest Neighbor. Tahapan penelitian dilakukan secara sistematis sebagai berikut:

1. Identifikasi Masalah
2. Studi Literatur
3. Pengumpulan Data
4. Preprocessing Data (Harmandini & Muslim, 2024)
5. Ekstraksi Fitur (TF-IDF) (Gazali Mahmud et al., 2023)
6. Penerapan Algoritma KNN
7. Pengujian Model
8. Evaluasi Model
9. Kesimpulan dan Rekomendasi

Tabel 1. Tahapan Penelitian

No	Tahapan	Deskripsi	Output
1	Identifikasi Masalah	Identifikasi isu perundungan dan persepsi siswa	Rumusan masalah
2	Studi Literatur	Kajian teori dan penelitian terdahulu	Landasan teori
3	Pengumpulan Data	Pengambilan data melalui kuesioner	Dataset opini
4	Preprocessing Data	Membersihkan teks data	Data siap olah
5	Ekstraksi Fitur	Konversi teks menjadi TF-IDF	Vektor fitur
6	Penerapan KNN	Klasifikasi sentimen dengan KNN	Model klasifikasi
7	Pengujian Model	Pembagian data latih & uji	Hasil uji model
8	Evaluasi Model	Pengukuran akurasi, presisi, recall, F1	Nilai evaluasi
9	Kesimpulan	Menarik kesimpulan dan rekomendasi	Saran & kesimpulan

Tahapan penelitian ini menggambarkan proses sistematis yang digunakan untuk menganalisis pendapat siswa terhadap edukasi anti perundungan menggunakan algoritma K-Nearest Neighbor (KNN). Setiap tahap dilakukan secara runtut untuk memastikan hasil penelitian akurat dan dapat diuji secara ilmiah.



Gambar 1. Tahapan Penelitian

Adapun tahapan penelitian yang dilakukan adalah sebagai berikut:

1. **Identifikasi Masalah**
Tahapan ini dilakukan untuk mengidentifikasi permasalahan terkait fenomena perundungan di sekolah dan bagaimana persepsi siswa terhadap program edukasi anti perundungan.
2. **Pengumpulan Literatur.**
Peneliti mempelajari berbagai referensi terkait machine learning, algoritma KNN, analisis sentimen, serta penelitian serupa yang telah dilakukan sebelumnya. (Rusilowati & Isdaryanti, 2024)
3. **Pengumpulan Data.**
Data dikumpulkan dari mengikuti kegiatan edukasi anti perundungan. Data berbentuk teks opini/pendapat siswa. (Rusilowati & Isdaryanti, 2024)
4. **Preprocessing Data.**
Data teks dilakukan beberapa proses seperti case folding, tokenizing, stopword removal, dan stemming untuk memperoleh data yang siap untuk diproses menggunakan algoritma KNN. (Kusuma & Cahyono, 2023a)
5. **Ekstraksi Fitur.**
Tahapan ini bertujuan mengubah teks menjadi representasi numerik menggunakan metode seperti TF-IDF (Kusuma & Cahyono, 2023b)
6. **Penerapan Algoritma KNN.**
Model KNN diterapkan untuk mengklasifikasikan opini siswa ke dalam kategori sentimen seperti positif, netral, dan negatif berdasarkan kedekatan fitur antar data. (Lin & Shih, 2024)
7. **Pengujian Model.**
Pengujian dilakukan dengan menggunakan data uji. Proses pembagian data training dan testing diterapkan untuk mengevaluasi performa model. (Legito et al., 2023)
8. **Evaluasi Model.**
Tahapan evaluasi dilakukan dengan menggunakan metrik seperti akurasi, presisi, recall, dan F1-score untuk memastikan bahwa model mampu memberikan hasil yang optimal. (Setiawan & Adnyana, 2023)
9. **Kesimpulan dan Rekomendasi.**
Hasil penelitian dianalisis dan dirumuskan kesimpulannya serta rekomendasi untuk pengembangan penelitian selanjutnya. (Subardiman, 2023)

2.2 Algoritma K-Nearest Neighbor (KNN)

KNN bekerja sebagai algoritma klasifikasi yang mengandalkan tingkat kemiripan atau kedekatan jarak antar data. Secara prinsip, status atau label dari data baru ditentukan oleh kelompok mayoritas yang muncul dari sejumlah 'k' tetangga

terdekat dalam ruang dimensi fitur. Karena algoritma ini langsung melakukan klasifikasi tanpa membentuk model eksplisit di awal, KNN kerap diklasifikasikan sebagai lazy learning. (Marhaely et al., 2024). Semakin kecil jarak antara data baru dengan data pembandingan, semakin besar kemungkinan keduanya berada pada kelas yang sama.

Pada prinsipnya, KNN tidak membangun model secara eksplisit, namun menyimpan data pelatihan dan melakukan proses klasifikasi secara langsung ketika menerima data baru. Oleh karena itu, algoritma ini sering disebut juga sebagai metode lazy learning. (Permana & Sahara, 2023) Tahapan penerapan metode KNN dijelaskan sebagai berikut:

Tahapan Metode KNN:

1. Menentukan nilai k .
2. Menghitung jarak antara data baru
3. Menentukan k tetangga terdekat berdasarkan jarak terkecil.
4. Menentukan kelas dari data baru berdasarkan mayoritas kelas dari tetangga terdekat.

Pseudo-code Algoritma KNN:

5. Input: Data pelatihan, data uji, nilai k
6. For setiap data uji:
7. Hitung jarak ke seluruh data pelatihan
8. Pilih k data dengan jarak terkecil
9. Tentukan kelas mayoritas
10. Output: Prediksi kelas data uji

Perhitungan: Persamaan Perhitungan Jarak (*Euclidean Distance*)

Persamaan (1): $d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$

Tabel 2. Data untuk proses klasifikasi dengan KNN.

Data	Fitur X	Fitur Y	Kelas
A	2	4	Positif
B	4	2	Negatif
C	3	3	Positif
Data Baru	3	2.5	-

Tabel 2 menampilkan data yang digunakan dalam proses klasifikasi menggunakan algoritma KNN. setiap objek data memiliki dua fitur, yaitu fitur x dan fitur y, serta label kelas yang menunjukkan kategori pendapat siswa, yaitu positif atau negatif.

data a memiliki nilai fitur $x = 2$ dan $y = 4$ sehingga masuk dalam kelas positif

data b memiliki nilai fitur $x = 4$ dan $y = 2$ sehingga masuk dalam kelas negatif

data c memiliki nilai fitur $x = 3$ dan $y = 3$ sehingga masuk dalam kelas positif

selain itu terdapat data baru dengan nilai fitur $x = 3$ dan $y = 2.5$ yang belum memiliki kelas (ditandai dengan "-").

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Mengumpulkan 200 data komentar siswa setelah program edukasi anti-perundungan. Data diberi label positif dan negatif.

Tabel 3. Distribusi Sentimen Siswa

Kategori	Jumlah
Positif	150
Negatif	50
Total	200

Sentimen Positif = $(150 / 200) \times 100 = 75\%$

Sentimen Negatif = $(50 / 200) \times 100 = 25\%$

3.1.1 Preprocessing Data

Tahapan *preprocessing* meliputi: *case folding, tokenizing, stopword removal, stemming, TF-IDF*. (Rojiati et al., 2025)

3.1.2 Implementasi KNN

Pembagian data: 160 training dan 40 testing.

Rumus Akurasi: akurasi dihitung dengan membandingkan jumlah prediksi yang benar, yaitu true positive (TP) dan true negative (TN) dikali 100 persen.

3.1.3 Hasil Pengujian

Tabel 4. Hasil Pengujian Model

k	Akurasi	Precision	Recall
1	85%	0.84	0.83
3	90%	0.88	0.89
5	92%	0.91	0.92
7	90%	0.89	0.88

3.1.4 Klasifikasi

Tabel 5. Hasil Klasifikasi

Komentar	Hasil KNN	Label Asli	Status
Program ini membuat saya sadar pentingnya saling menghargai	Positif	Positif	✓
Saya merasa tidak terlalu terbantu	Negatif	Negatif	✓
Kegiatan ini sangat membosankan	Negatif	Positif	✗

3.1.5 Pembahasan

Analisis menunjukkan bahwa performa optimal algoritma tercapai pada parameter $k=5$, dengan perolehan akurasi hingga 92%. Dari total 200 data yang diolah, ditemukan bahwa sebagian besar siswa (75%) memberikan respons positif, yang menandakan bahwa program edukasi tersebut diterima dengan baik dan dianggap bermanfaat oleh para peserta.

3.2 Implementasi

Tahapan implementasi algoritma KNN dalam penelitian ini dilakukan secara sistematis untuk menganalisis pendapat siswa terhadap edukasi anti-perundungan. Proses implementasi dimulai dari tahap pengumpulan data, dilanjutkan dengan pengolahan data teks, pembentukan model KNN, hingga proses evaluasi performa model.

1. Tahap Pengumpulan data

Data penelitian berupa komentar siswa yang diperoleh melalui penyebaran kuesioner serta pemanfaatan forum diskusi sekolah setelah pelaksanaan program edukasi anti-perundungan.

2. Tahap Preprocessing Teks

Sebelum data teks digunakan ada beberapa tahapan preprocessing untuk membersihkan dan menyiapkan data.

3. Pembagian Data

Dataset yang telah diproses kemudian dibagi ke dalam dua bagian, yaitu 80% data sebagai data pelatihan sebanyak 160 data dan 20% data sebagai data pengujian sebanyak 40 data.

4. Penerapan Algoritma KNN

Metode KNN digunakan untuk menentukan kelas sentimen suatu komentar berdasarkan kedekatan jaraknya dengan data yang telah diberi label.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan mengenai penerapan algoritma KNN dalam menganalisis pendapat siswa terhadap edukasi anti-perundungan, dapat disimpulkan bahwa metode KNN mampu melakukan klasifikasi sentimen secara efektif pada studi kasus ini. Proses pengujian menggunakan 3 data latih dan 1 data uji menunjukkan bahwa jarak Euclidean dari data uji (3,2,5)(3, 2,5)(3,2,5) terhadap masing-masing data latih, yaitu data A (2,4)(2,4)(2,4), data B (4,2)(4,2)(4,2), dan data C (3,3)(3,3)(3,3), berturut-turut adalah 1.118, 1.118, dan 0.500.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

- [1] Afriani, R. I., Rustandi, T., Firmansyah, A. A., Anjelena, R., & Nadiputri, A. Q. (2025). LINGKUNGAN MELALUI AKSI BEACH CLEAN UP DI DESTINASI. 5(2),
- [2] Albab, M. U., P., Y. K., & Fawaiq, M. N. (2023). Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic. *Jurnal Transformatika*, 20(2), 1–12. <https://doi.org/10.26623/transformatika.v20i2.5374>

- [3] Dwi Wicaksono, V., Dellarosa, M., Alfi Muhimmah, H., Sukartiningsih, W., & Negeri Surabaya, U. (2024). Pelatihan Penyusunan Program Sekolah Anti Bullying. 4, 87–93.
- [4] Efendi, M., & Latifah, N. aini. (2021). Jurnal Pendidikan Dan Pembelajaran Kimia. Penetapan Harga Jasa Pendidikan Di Perguruan Tinggi Keagamaan Islam Negeri (Ptkin, 2(2 (2021)), 127–143.
- [5] Fudhail Ferio Supeli, M., & Setiaji, S. (2023). Klasifikasi Sentimen Positif Dan Negatif Pada Aplikasi Vidio Dengan Algoritma K-Nearest Neighbor. Indonesian Journal Computer Science, 2(1), 7–15. <https://doi.org/10.31294/ijcs.v2i1.1874>
- [6] Gazali Mahmud, F., Iman Hermanto, T., & Maruf Nugroho, I. (2023). Implementation Of K-Nearest Neighbor Algorithm With SMOTE For Hotel Reviews Sentiment Analysis. SinkrOn, 8(2), 595–602. <https://doi.org/10.33395/sinkron.v8i2.12214>
- [7] Harmandini, K. P., & Muslim, K. (2024). Analysis of TF-IDF and TF-RF Feature Extraction on Product Review Sentiment. Jurnal Dan Penelitian Teknik Informatika, 8(2). <https://doi.org/10.33395/v8i2.13376>
- [8] Jurnal, J., Sains, A., Istibsyaroh, I., Ruhimat, A., Rokhmawati, D., Yuliana, R., Danau, J., & No, S. (2023). Peningkatan Wawasan Orangtua dalam Mencegah dan Menangani Perilaku Perundungan Pada Anak (Improvement of Parents ' Insight in Preventing and Overcoming Bullying Behavior on Children) Pada tahun 2011 , Kementerian Pemberdayaan Perempuan dan Perlindungan . 7(1), 27–40.
- [9] Kusuma, I. H., & Cahyono, N. (2023a). Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor. Jurnal Informatika: Jurnal Pengembangan IT, 8(3), 302–307. <https://doi.org/10.30591/jpit.v8i3.5734>
- [10] Lin, J. C., & Shih, Y. H. (2024). Strategies for preventing school bullying—A life education perspective. Frontiers in Psychology, 15. <https://doi.org/10.3389/fpsyg.2024.1429215>
- [11] Kusuma, I. H., & Cahyono, N. (2023b). Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor. Jurnal Informatika: Jurnal Pengembangan IT, 8(3), 302–307. <https://doi.org/10.30591/jpit.v8i3.5734>
- [12] Legito, L., Riau, N. P., Putro, A. N. S., Mardiani, E., & ... (2023). ... Nearest Neighbor untuk Analisis Sentimen Terhadap Isu Khilafah dan Radikalisme di Indonesia: Implementation K-Nearest Neighbor Algorithm for Sentiment Analysis Indonesian Journal of ..., 3(October), 324–330. <https://journal.irpi.or.id/index.php/malcom/article/view/893%0Ahttps://journal.irpi.or.id/index.php/malcom/article/download/893/448>
- [13] Marhaely, S., Purwanto, A., Aini, R. N., Asyanti, S. D., Sarjan, W., & Paramita, P. (2024). Literature Review : Model Edukasi Upaya Pencegahan. Jurnal Kesehatan Tambusai, 5, 826–834.
- [14] Permana, R. A., & Sahara, S. (2023). Algoritma K-Nearest Neighbor Pada Analisa Sentimen Review Produk Router. SIMKOM, 8(2), 118–124. <https://doi.org/10.51717/simkom.v8i2.129>
- [15] Rojiati, J. E.-I. ; U., Bakri, U., Wulandari, A. M., Permana, A., & Pratiwi, A. (2025). Implementasi Program Edukasi Anti-Perundungan sebagai Upaya Pembentukan Karakter Peserta Didik SMPN 2 Tanjung Bintang Kabupaten Lampung Selatan. Journal of Smart Community Service, 3(1), 24–35. <https://journal.cahyaedu.com/index.php/jscs>
- [16] Rusilowati, A., & Isdaryanti, B. (2024). 1213-Article Text-6040-1-10-20241201. 13(4), 5359–5372.
- [17] Setiawan, G. H., & Adnyana, I. M. B. (2023). Improving Helpdesk Chatbot Performance with Term Frequency Inverse Document Frequency (TF-IDF) and Cosine Similarity Models. Journal of Applied Informatics and Computing, 7(2), 252–257. <https://doi.org/10.30871/jaic.v7i2.6527>
- [18] Subardiman, A. F. (2023). BAB I PE N DAHULUA N 1.1 Latar Belakang Di Jawa Barat sendiri, data menunjukkan 2.001 kasus kekerasan, termasuk. 1–9.