

ANALISIS PERILAKU PERUNDUNGAN DILINGKUNGAN SEKOLAH DENGAN METODE KLASIFIKASI NAÏVE BAYES (STUDI KASUS : SDN KIARAPAYUNG I)

Nurmala Dewi, Mohamad Jajuli, Susilawati

Informatika, Universitas Singaperbangsa Karawang
HS.Ronggo Waluyo, Puseurjaya, Kec. Telukjambe Timur
Kab. Karawang, Jawa Barat, 41361, Karawang-Indonesia
2110631170121@student.unsika.ac.id

ABSTRAK

Perundungan atau *bullying* adalah tindakan kekerasan yang sering terjadi di lingkungan sekolah termasuk ditingkat sekolah dasar. Tindakan ini dapat berupa intimidasi fisik, verbal maupun tekanan psikologis yang dilakukan secara berulang dengan tujuan mendominasi korban. Peningkatan kasus perundungan di sekolah dalam beberapa tahun terakhir menunjukkan perlunya metode yang efektif dalam mendeteksi dan menganalisis perilaku perundungan. Penelitian ini bertujuan untuk mengklasifikasikan perilaku perundungan di SDN Kiarapayung 1 menggunakan metode klasifikasi *Naïve Bayes*. Prosedur penelitian dilakukan dengan pendekatan *Knowledge Discovery in Database (KDD)*, yang mencakup *data selection*, *data pre-processing*, *data transformation*, *data mining*, *evaluation*, serta *knowledge presentation*. Dataset yang digunakan diperoleh dari pengamatan langsung kepada siswa kelas I hingga VI, dengan variabel utama meliputi kekerasan fisik, verbal, psikologis, pengalaman perasaan siswa, dan kenyamanan siswa di sekolah. Data diklasifikasikan ke dalam dua kategori yaitu *Bully* dan *Non-Bully*. Hasil penelitian menunjukkan bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan perilaku perundungan dengan tingkat akurasi terbaik sebesar 85% pada rasio data latih dan uji 90:10. Evaluasi model menggunakan *confusion matrix* menunjukkan nilai *accuracy* 85%, *precision* 77%, *recall* 91% dan *F1-Score* 83%.

Kata kunci : *Naïve Bayes Classifier*, *Perundungan*, *Bully*, *Klasifikasi*, *KDD*

1. PENDAHULUAN

Perundungan ini dapat diartikan sebagai bentuk intimidasi fisik, serangan verbal, atau tekanan psikologis yang dilakukan dengan tujuan untuk menakut-nakuti, menekan, atau menyakiti korban tertentu dimana individu lebih kuat menindas mereka yang lemah. Aktivitas yang penuh tekanan ini biasanya terjadi pada anak-anak usia sekolah. Kekerasan yang dialami anak-anak disekolah sering disebut *school bullying* yaitu dimana keadaan seorang anak menerima perlakuan tidak menyenangkan yang umumnya terjadi secara intensif dan berulang-ulang dengan tujuan menunjukkan dominasi atau kekuasaan. Kondisi ini mengakibatkan perasaan tidak nyaman dan tertekan pada anak yang menjadi korbannya. Di sekolah, perundungan ini sering muncul dalam bentuk kekerasan fisik dan sosial, biasanya dilakukan oleh kakak kelas terhadap adik kelas atau antar teman kelas yang membedakan perlakuan terhadap teman-teman mereka [1]. Dalam beberapa tahun terakhir, kasus perundungan di lingkungan sekolah mengalami peningkatan signifikan. Dilansir dari databoks.co.id menurut laporan dari Global Federasi Guru Indonesia (FSGI) menunjukkan peningkatan kasus dari 21 kasus pada tahun 2022 menjadi 30 kasus pada tahun 2023 dengan mayoritas terjadi di sekolah-sekolah yang dinaungi oleh Kementrian, Pendidikan, Kebudayaan, Riset dan teknologi (KEMDIKBUDRISTEK). Perundungan ini terjadi di SMP sebanyak 50%, SD 30%, dan SMA/SMK masing-masing 10%. Oleh karena itu, diperlukan cara efektif untuk mendeteksi dan menganalisis perilaku perundungan guna mencegah terjadinya perundungan lebih lanjut di

lingkungan sekolah yang tentunya melibatkan komponen sekolah dengan tujuan dapat menyadarkan tentang bahaya dari perilaku perundungan [2]. Sekolah dasar merupakan jenjang pendidikan awal yang memiliki peran penting dalam membentuk karakter anak. Kebijakan sekolah ramah anak yang di keluarkan oleh Kementrian Negara Pemberdayaan Perempuan dan Perlindungan Anak Nomor 8 tahun 2014 menekankan pentingnya menciptakan lingkungan pendidikan yang aman dan mendukung perkembangan anak. Namun, adanya perilaku perundungan yang terjadi di sekolah justru bertolak belakang dengan konsep ini. Perundungan dapat menimbulkan dampak yang luas, seperti stress, depresi, penurunan prestasi akademik, hingga gangguan akademis [1]. SDN Kiarapayung I dipilih sebagai objek penelitian karena berdasarkan pengalaman salah satu siswa yang mengalami perundungan secara verbal yang dilakukan oleh teman sekelasnya yang menjadikan siswa tersebut merasa tidak nyaman. Kondisi ini menunjukkan perlunya tindakan preventif dan analisis mendalam untuk mengidentifikasi tingkat prevalensi perundungan di sekolah.

Berdasarkan data dan studi kasus tersebut, *data mining* menjadi solusi untuk permasalahan yang dihadapi sekolah untuk meminimalisir perilaku perundungan tersebut. Manfaat dari *data mining* ialah sebagai bahan atau metode untuk mendukung pengambilan keputusan serta menunjang kegiatan yang dibutuhkan [3]. Penelitian ini menggunakan algoritma *Naïve Bayes Classifier* yang bertujuan untuk mempermudah sekolah dalam mendeteksi dan menganalisis perilaku perundungan. *Knowledge*

Discovery in Database (KDD) merupakan proses untuk menemukan informasi dari data yang telah tersimpan di basis data yang saling terhubung. Hasil dari *KDD* dapat dimanfaatkan untuk mendukung pengambilan keputusan [4]. Terdapat penelitian terkait perilaku perundungan yang terjadi di dunia maya atau biasa disebut dengan *cyberbullying* menggunakan algoritma *Naïve Bayes Classifier* yang dilakukan dengan analisis sentimen dengan mengambil data di *platform* penyedia dataset seperti *kaggle* dan media sosial seperti *Instagram*, *Twitter*, dan *Youtube*. Penelitian tersebut memanfaatkan teknik pengolahan bahasa alami (*Neural Language Programming*) untuk mendeteksi dan menganalisis komentar pada sosial media yang mengindikasikan adanya perilaku *cyberbullying*. Seperti penelitian sebelumnya yang dilakukan oleh Denni Kurniawan & Muhammad Yasir yang melakukan optimalisasi analisis sentimen menggunakan metode *CRISP – DM* dan *Naïve Bayes Classifier* di media sosial *Twitter* mengenai *cyberbullying* yang mendapatkan *accuracy* 85,6%, *recall negative* 82,1% dan *recall positive* 90,23%, *precision negative* dan *positive* masing-masing sebesar 91,76% dan 79,18% [5]. Jasman Pardede, Yusuf Miftahudin & Wahidal Kahar juga telah melakukan penelitian mengklasifikasi komentar *cyberbullying* di media sosial dengan mengambil data dari *kaggle* dengan menggunakan model multinomial untuk pengambil keputusan apakah makna kata tersebut *bully* atau *non-bully* menggunakan *Naïve Bayes Classifier* yang menghasilkan nilai *accuracy* 80%, *precision* 81% dan *recall* 80% [6]. Penelitian terkait perundungan yang dilakukan dengan analisis sentimen pada media sosial dengan *Naïve Bayes Classifier* cukup baik, namun penelitian yang dilakukan dengan dataset yang diambil dari media sosial dan *platform kaggle* tersebut memiliki keterbatasan karena kecenderungan perilaku yang berbeda antara kehidupan maya dan kehidupan nyata [7]. Riset yang berfokus pada perilaku perundungan langsung di sekolah menggunakan algoritma *Naïve Bayes Classifier* masih terbatas. Oleh karena itu, keterbatasan tersebut bisa lebih di perbaiki dengan cara pendekatan secara langsung dengan metode algoritma *Naïve Bayes Classifier* pada data perundungan yang terjadi di sekolah menjadi penting untuk menutupi celah penelitian ilmiah. Berdasarkan pemaparan diatas, maka dilakukan penelitian dengan tujuan menerapkan algoritma *Naïve Bayes Classifier* untuk mengklasifikasikan perilaku perundungan di lingkungan sekolah dan mengevaluasi hasil penerapannya.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian sebelumnya dilakukan oleh Wibowo, Khairi, Humairoh, Fadhilah & Hartawan menggunakan algoritma *naïve bayes classifier*, *logistic regression* & untuk melihat hasil perbandingan algoritma dan menentukan komentar

pada *platform twitter* mengandung makna *bully* atau *non-bully*. Hasilnya menunjukkan bahwa model *naïve bayes* memiliki akurasi tertinggi dengan akurasi 78,26% dan disusul dengan model *KNN* dengan akurasi 59,90% [8].

Pada penelitian lain terkait *naïve bayes classifier* yang dilakukan oleh Alfarizi, Rizky, Ghufroni, Fathurahman, Saputra & Kurniawan melakukan analisis sentimen persepsi publik terhadap kasus *bullying* siswa di cilacap yang bertujuan untuk menerapkan algoritma klasifikasi dalam menganalisis sentimen komentar di media sosial. Penelitian ini menghasilkan tingkat *accuracy* sebesar 98,19% [9].

2.2. Data Mining

Data Mining adalah suatu alat yang memungkinkan pengguna mengakses data dalam jumlah besar dengan cepat. *Data mining* mencakup serangkaian teknik yang bertujuan mengidentifikasi pola-pola yang sebelumnya tidak diketahui dalam data yang terkumpul [10]. Secara garis besar ada 2 strategi untuk *data mining*, yaitu strategi prediktif dan strategi deskriptif yang masing-masing memiliki algoritma yang berbeda [11].

2.3. Klasifikasi

Klasifikasi adalah satu teknik dalam data mining yang bertujuan untuk mengelompokkan data kedalam kategori atau kelas tertentu berdasarkan atribut yang dimilikinya. Proses ini bertujuan untuk membangun model yang mampu memprediksi kategori atau label dari data dengan mengandalkan pola yang telah dipelajari dari data pelatihan. Dengan menganalisis atribut yang berkontribusi pada pengelompokkan data kedalam kategori tertentu, akan dapat memahami faktor apa yang mempengaruhi keputusan klasifikasi [12].

2.4. Naïve Bayes Classifier

Naïve Bayes classifier merupakan metode klasifikasi statistik yang digunakan untuk memprediksi probabilitas suatu tupel termasuk dalam kelas tertentu. Algoritma ini menunjukkan kecepatan dan akurasi, terutama ketika diterapkan pada dataset berukuran besar. Teori *Naïve Bayes* merupakan bagian dari teori statistik matematika yang memungkinkan pembuatan model ketidakpastian dari suatu kejadian yang terjadi dengan menggabungkan pengetahuan umum dan fakta dari hasil pengamatan. Statement teori *Naïve Bayes* sebagai berikut beserta penjelasannya [13].

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} \quad (1)$$

Keterangan :

X : Merupakan data dengan kelas yang belum diketahui.

H : Hipotesis bahwa data termasuk dalam suatu kelas tertentu.

P(H|X) : Probabilitas hipotesis H berdasarkan kondisi data X.

$P(H)$: Probabilitas awal dari hipotesis H.

$P(X|H)$: Probabilitas data X berdasarkan kondisi pada hipotesis H.

$P(X)$: Probabilitas X secara keseluruhan.

2.5. Confusion Matrix

Confusion matrix adalah salah satu metode yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil klasifikasi yang dilakukan sistem dengan klasifikasi yang seharusnya. Metode ini memberikan gambaran seberapa baik model dalam mengklasifikasi data kedalam berbagai kategori. Dalam penelitian ini, *confusion matrix* digunakan untuk menghitung parameter-parameter seperti *accuracy*, *precision*, *recall* dan *f-measure*. Terdapat 4 istilah yang menggambarkan hasil dari proses klasifikasi dalam pada *confusion matrix* [14] yaitu :

- True positive (TP).** Kondisi dimana prediksi benar sesuai dengan nilai yang sebenarnya yaitu sama-sama benar.
- True Negative (TN).** Kondisi dimana prediksi benar sesuai dengan nilai yang sebenarnya yaitu sama-sama salah.
- False Positive (FP).** Kondisi dimana prediksi benar, tetapi hasilnya yang sebenarnya salah.
- False Negative (FN).** Kondisi dimana prediksi salah, namun hasilnya yang sebenarnya benar.

Berikut adalah persamaan model *confusion matrix* [18] yaitu :

- Accuracy.** Mengukur sejauh mana prediksi yang dilakukan sistem mendekati kenyataan.

$$A = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

- Precision.** Mengukur seberapa tepat sistem memberikan jawaban yang diminta dibandingkan dengan hasil yang benar.

$$P = \frac{TP}{(TP+FP)} \quad (3)$$

- Recall.** Menunjukkan sejauh mana sistem dapat menemukan informasi yang relevan dari keseluruhan data.

$$R = \frac{TP}{(TP+FN)} \quad (4)$$

- F-Measure.** Evaluasi yang menggabungkan *precision* dan *recall* untuk memberikan gambaran umum dari kinerja sistem.

$$F - Measure = \frac{2 * P * R}{P + R} \quad (5)$$

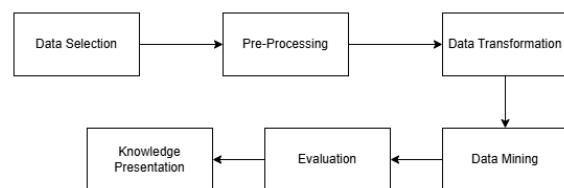
2.6. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah proses untuk menemukan informasi dari data yang telah tersimpan di basis data yang saling terhubung. Hasil dari *KDD* dapat dimanfaatkan untuk mendukung pengambilan keputusan. Istilah *KDD* dan *data mining* sering digunakan bergantian, namun sebenarnya berbeda. *Data mining* hanya merupakan salahsatu tahap dalam proses *KDD*. Langkah-langkah yang terlibat dalam proses *KDD* yaitu *data selection*, *data*

pre-processing, *data transformation*, *data mining*, *evaluation* dan *knowledge presentation* [4].

3. METODE PENELITIAN

Pada bagian ini memuat metode yang digunakan pada pembuatan skripsi. Metode penelitian yang digunakan pada penelitian ini yaitu menggunakan *KDD* yang terdiri dari beberapa tahapan untuk mengolah dan menganalisis data. Tahapan proses *KDD* yang digunakan dalam penelitian ini adalah *data selection*, *pre-processing*, *data transformation*, *data mining*, *evaluation* dan *knowledge presentation*. Berikut merupakan gambaran dari alur metode penelitian.



Gambar 1. Tahapan Metodologi

4. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data primer yang diperoleh melalui survei formulir yang dibagikan kepada siswa kelas I hingga VI SDN Kiarapayung I. survei tersebut mencakup beberapa pertanyaan yang berkaitan dengan perundungan. Total jumlah siswa yang terdaftar di SDN Kiarapayung I pada kelas I hingga VI adalah 270 siswa. Dari jumlah tersebut, berhasil terkumpul sebanyak 255 data siswa yang menjadi sampel data dalam penelitian ini. Proses pengambilan data dilakukan pada tanggal 6 hingga 7 Januari 2025. Data yang dikumpulkan mencakup periode satu bulan terakhir, yakni dari November hingga Desember 2024.

No	Nama	Kelas	Usia	Jenis Kelamin	Kekerasan Fisik	Kekerasan Verbal	Kekerasan Psikologis / Sosial	Pelaku Perundungan	Waktu Kejadian Perundungan	Pelaporan Guru	Tindakan Lanjut Pelaporan	Pengalaman Perasaan Siswa	Kenyamanan Siswa
1	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
2	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
3	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
4	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
5	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
6	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
7	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
8	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
9	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
10	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
11	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
12	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
13	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
14	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
15	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
16	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
17	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
18	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
19	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah
20	Adhika Pratomo	1	12 tahun	Laki-laki	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah	Tidak pernah

Gambar 2. Dataset Perundungan Siswa

Gambar 2 menyajikan gambaran data perundungan yang dialami oleh siswa selama periode tersebut. Dataset yang diperoleh terdiri dari 255 data dengan 13 atribut, yaitu Nama, Kelas, Usia, Jenis Kelamin, Kekerasan Fisik, Kekerasan Verbal, Kekerasan Psikologis / Sosial, Pelaku Perundungan, Waktu Kejadian Perundungan, Pelaporan Guru, Tindakan Lanjut Pelaporan, Pengalaman Perasaan Siswa dan Kenyamanan Siswa. Atribut yang dipilih mencakup variabel-variabel yang memiliki pengaruh langsung terhadap identifikasi dan klasifikasi perilaku perundungan siswa. Deskripsi data dapat dilihat pada tabel 1 berikut

Tabel 1. Deskripsi Atribut

Atribut	Tipe Data	Keterangan
Nama	Object	Nama siswa
Kelas	Int	Kelas siswa
Usia	Object	Usia siswa
Jenis Kelamin	Object	Jenis kelamin siswa laki-laki dan perempuan
Kekerasan Fisik	Object	Frekuensi perilaku perundungan secara fisik yang dialami siswa.
Kekerasan Verbal	Object	Frekuensi perilaku perundungan secara verbal yang dialami siswa.
Kekerasan Psikologis / Sosial	Object	Frekuensi perilaku perundungan secara psikologis / sosial yang dialami siswa.
Pelaku Perundungan	Object	Pelaku yang biasanya melakukan perundungan.
Waktu kejadian perundungan	Object	Waktu terjadinya mendapatkan perilaku perundungan yang sering terjadi.
Lapor guru	Object	Siswa melakukan pelaporan.
Tindaklanjut pelaporan	Object	Penindaklanjutan pelaporan siswa.
Pengalaman perasaan siswa	Object	Perasaan siswa setelah mengalami perundungan.
Kenyamanan siswa	Object	Kenyamanan siswa setelah mengalami perilaku perundungan.

4.1 Data Selection

Pada tahap *data selection* dilakukan pemilihan atribut data yang relevan karena tidak semua atribut data digunakan dan hanya berfokus pada perilaku perundungan. Atribut yang diseleksi dalam penelitian ini, yaitu Kekerasan Fisik, Kekerasan Verbal, Kekerasan Psikologis / Sosial, Pengalaman Perasaan Siswa dan Kenyamanan Siswa. Hasil data selection sebelum di seleksi dapat dilihat pada gambar berikut.

Hasil data selection sesudah diseleksi dapat dilihat pada gambar berikut :

	Kekerasan Fisik	Kekerasan Verbal	Kekerasan Psikologis / Sosial	Pengalaman Perasaan Siswa	Kenyamanan Siswa
0	Tidak pernah	Pernah 2 atau 3 kali	Pernah 1 kali	Tidak terganggu	Iya
1	Pernah lebih dari 3 kali	Pernah lebih dari 3 kali	Pernah 1 kali	Sangat terganggu	Tidak
2	Pernah lebih dari 3 kali	Pernah 1 kali	Pernah 1 kali	Tidak terganggu	Iya
3	Tidak pernah	Pernah 2 atau 3 kali	Pernah 1 kali	Terganggu sebentar	Iya
4	Pernah 1 kali	Tidak pernah	Pernah 1 kali	Tidak terganggu	Iya

Gambar 3. Atribut data sesudah di seleksi

4.2 Data Pre-Processing

Pada tahap *data pre-processing* dilakukan pengecekan untuk menghasilkan set data final yang akan digunakan dalam tahap selanjutnya. Melalui tahap *data pre-processing* ini, data dapat dipersiapkan dengan baik, diharapkan hasil analisis nantinya akan lebih akurat dan dapat memberikan wawasan yang berarti bagi pengambilan keputusan. Pada tahap ini, dilakukan pengecekan dan pembersihan data agar data yang diolah benar-benar bersih. Pembersihan data dilakukan dengan cara pengecekan *missing value* dan dilakukan penanganan jika ada. Selain itu, dilakukan juga pengecekan data pada atribut untuk memastikan tidak ada inkonsistensi data.

```
# Mengecek Missing Values
print("\nMissing Values per Kolom:")
(df_selected.isnull().sum())
```

Missing Values per Kolom:

Kekerasan Fisik	0
Kekerasan Verbal	0
Kekerasan Psikologis / Sosial	0
Pengalaman Perasaan Siswa	0
Kenyamanan Siswa	0

Gambar 4. Pengecekan *missing value*

```
Cek Data Tidak Valid atau Inkonsisten:
Kekerasan Fisik: ['Tidak pernah' 'Pernah lebih dari 3 kali' 'Pernah 1 kali'
'Pernah 2 atau 3 kali']
Kekerasan Verbal: ['Pernah 2 atau 3 kali' 'Pernah lebih dari 3 kali' 'Pernah 1 kali'
'Tidak pernah']
Kekerasan Psikologis / Sosial: ['Pernah 1 kali' 'Pernah lebih dari 3 kali' 'Tidak pernah'
'Pernah 2 atau 3 kali']
Pengalaman Perasaan Siswa: ['Tidak terganggu' 'Sangat terganggu' 'Terganggu sebentar']
Kenyamanan Siswa: ['Iya' 'Tidak']
```

Gambar 5. Pengecekan inkonsistensi data

Setelah dilakukan pengecekan *missing value* pada dataset dengan menggunakan fungsi *isnull()* seperti yang ditunjukkan dalam Gambar 4, dapat dilihat bahwa tidak terdapat *missing value* yang ditemukan. Hal ini menandakan bahwa setiap entri dalam dataset memiliki nilai yang lengkap dan tidak ada data yang hilang atau kosong. Selanjutnya dilakukan pengecekan inkonsistensi data seperti yang ditunjukkan dalam Gambar 5, hasil menunjukkan keseragaman data.

4.3 Data Transformation

Setelah dilakukan tahap data pre-processing untuk memastikan data bersih, langkah selanjutnya adalah tahap *data transformation*. Tahap ini bertujuan untuk mengubah data menjadi format yang lebih terstruktur dan dapat digunakan untuk proses modelling selanjutnya. Proses transformasi data diawali dengan melakukan perubahan nama atribut pada Kenyamanan Siswa menjadi label karena akan dijadikan kolom *class target* serta mengganti Iya dan tidak dengan 1 (*Non-Bully*) dan 0 (*Bully*). Hasil sesudah melakukan perubahan pada nama atribut Kenyamanan Siswa menjadi label serta mengubah data Iya menjadi 0 dan Tidak menjadi 1 dapat dilihat pada Gambar 6.

	Kekerasan Fisik	Kekerasan Verbal	Kekerasan Psikologis / Sosial	Pengalaman Perasaan Siswa	label
0	Tidak pernah	Pernah 2 atau 3 kali	Pernah 1 kali	Tidak terganggu	1
1	Pernah lebih dari 3 kali	Pernah lebih dari 3 kali	Pernah 1 kali	Sangat terganggu	0
2	Pernah lebih dari 3 kali	Pernah 1 kali	Pernah 1 kali	Tidak terganggu	1
3	Tidak pernah	Pernah 2 atau 3 kali	Pernah 1 kali	Terganggu sebentar	1
4	Pernah 1 kali	Tidak pernah	Pernah 1 kali	Tidak terganggu	1

Gambar 6. Atribut data sesudah proses konversi

Selanjutnya dilakukan encoding data menggunakan *Label Encoder*. Proses *encoding* ini diterapkan pada atribut yang sudah diseleksi pada tahap *data selection* yaitu diantaranya Kekerasan

Fisik, Kekerasan Verbal, Kekerasan Psikologis / Sosial, Pengalaman Perasaan Siswa dan label. Kriteria encoding yang digunakan dalam tahapan ini dapat dilihat pada Tabel 2.

Tabel 2. Kriteria *encoding*

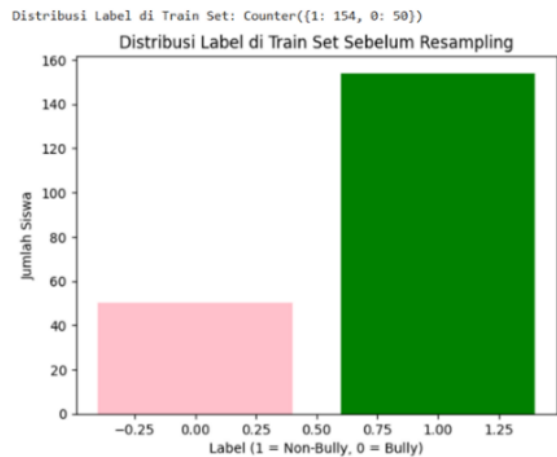
Atribut	kategori	Nilai <i>Encoding</i>
Kekerasan fisik	Tidak Pernah	3
	Pernah 1 kali	0
	Pernah 2 atau 3 kali	1
	Pernah lebih dari 3 kali	2
Kekerasan Verbal	Tidak Pernah	3
	Pernah 1 kali	0
	Pernah 2 atau 3 kali	1
	Pernah lebih dari 3 kali	2
Kekerasan Psikologis / Sosial	Tidak Pernah	3
	Pernah 1 kali	0
	Pernah 2 atau 3 kali	1
	Pernah lebih dari 3 kali	2
Pengalaman Perasaan Siswa	Tidak Terganggu	2
	Terganggu Sebesar	1
	Sangat Terganggu	0
label	Iya	0
	Tidak	1

Hasil *data transformation* sesudah di *encoding* dapat dilihat pada Gambar 7.

	Kekerasan Fisik	Kekerasan Verbal	Kekerasan Psikologis / Sosial	Pengalaman Perasaan Siswa	label
0	3	1	0	2	1
1	2	2	0	0	0
2	2	0	0	2	1
3	3	1	0	1	1
4	0	3	0	2	1

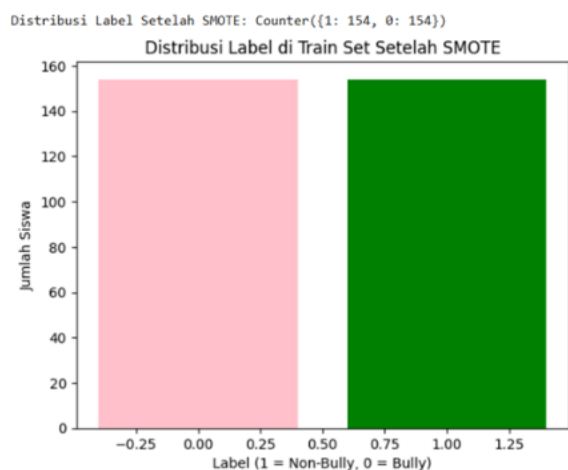
Gambar 7. Atribut data sesudah proses *encoding*

Setelah proses *encoding* data menggunakan *Label Encoder*, langkah selanjutnya memisahkan fitur dan label serta dataset *train-test*. Memisahkan fitur (X) dan label (y) dimana fitur terdiri dari atribut kekerasan fisik, kekerasan verbal, kekerasan psikologis / sosial, dan pengalaman perasaan siswa, sementara label adalah kenyamanan siswa yang diklasifikasikan sebagai *Bully* (0) atau *Non-Bully* (1). Setelah fitur dan label dipisahkan, data kemudian dibagi menjadi data latih (*train set*) dan data uji (*test set*) dengan rasio 80% untuk data latih dan 20% untuk data uji menggunakan fungsi *train_test_split*. Pada pembagian dataset dengan rasio 80% dan 20% didapat jumlah nya masing-masing sebanyak 204 data dan 51 data. Selanjutnya dilakukan pengecekan distribusi label pada data latih untuk mengetahui apakah terdapat ketidakseimbangan kelas (*imbalance class*) antara siswa yang termasuk dalam kategori *Bully* dan *Non-Bully*. Jika distribusi label tidak seimbang, maka model klasifikasi bisa menjadi bias terhadap kelas mayoritas. Oleh karena itu, dilakukan visualisasi label menggunakan diagram.



Gambar 8. Distribusi label di train set

Gambar 10 menunjukkan hasil visualisasi label yang menunjukkan ketidakseimbangan label. Jumlah siswa dengan label *Non-Bully* (label 1) jauh lebih banyak dibandingkan dengan siswa dengan label *Bully* (label 0). Distribusi label menunjukkan bahwa terdapat 154 siswa *Non-Bully* dan 50 siswa *Bully*. Perbedaan jumlah ini menunjukkan ketidakseimbangan data yang dapat menyebabkan model klasifikasi lebih cenderung memprediksi kelas mayoritas. Oleh karena itu, dilakukan teknik *oversampling SMOTE* (*Synthetic Minority Over-Sampling Technique*) untuk menyeimbangkan jumlah data antara kedua kelas sebelum *modelling Naïve Bayes Classifier*.



Gambar 9. Distribusi label di train set setelah menggunakan teknik SMOTE

Gambar 11 menunjukkan distribusi label setelah menerapkan teknik *SMOTE* jumlah data pada kedua kelas menjadi seimbang yaitu 154 data dan 154 data. Teknik *SMOTE* bekerja dengan mensintesis sampel baru untuk kelas minoritas, sehingga model akan lebih mengenali dan memprediksi kedua label dengan lebih akurat. Setelah data latih diseimbangkan menggunakan *SMOTE*, dilakukan pemisahan ulang fitur (*X_Train_final*) dan label (*y_train_final*) agar dapat digunakan dalam proses selanjutnya.

4.4 Data Mining

Pada tahap ini, dilakukan proses pemodelan untuk mengklasifikasikan perilaku perundungan berdasarkan data yang telah dikumpulkan dan melalui serangkaian tahapan pengolahan data, yaitu *data selection*, *data pre-processing*, dan *data transformation*. Proses klasifikasi menggunakan algoritma *Naïve Bayes Classifier* dengan bantuan *Google Colaboratory* sebagai *tools* pemodelan. Untuk menguji performa model, digunakan teknik *percentage split*, di mana dataset dibagi menjadi data *training* dan *testing* dengan rasio 90:10, 80:20, 70:30, 60:40, dan 50:50. Pembagian dataset ini bertujuan untuk mengevaluasi performa model dengan variasi jumlah data latih dan data uji, sehingga dapat dianalisis bagaimana perubahan proporsi data mempengaruhi akurasi klasifikasi. Pembagian dataset dapat dilihat pada Tabel 3.

Tabel 3. Jumlah Data Training dan Data Testing

RationData Training & Data Testing	Jumlah Data Training	Jumlah Data Testing
90 : 10	229	26
80 : 20	204	51
70 : 30	178	77
60 : 40	153	102
50 : 50	127	128

Setelah pembagian dataset dilakukan, tahap selanjutnya adalah membangun model *Naïve Bayes* dengan menggunakan fungsi model = *GaussianNB()*. Dalam implementasinya, *Gaussian Naïve Bayes* menghitung probabilitas dari setiap kelas berdasarkan data *training* dan menentukan kelas yang paling mungkin untuk setiap contoh dalam *data testing*.

4.5 Evaluation

Tahap *evaluation* merupakan tahap untuk melakukan evaluasi terhadap performa dari implementasi algoritma *Naïve Bayes Classifier* untuk melakukan klasifikasi dengan *confusion matrix*.

➡ Akurasi Model: 0.82

Laporan Klasifikasi:

	precision	recall	f1-score	support
0	0.76	0.72	0.74	18
1	0.85	0.88	0.87	33
accuracy			0.82	51
macro avg	0.81	0.80	0.80	51
weighted avg	0.82	0.82	0.82	51

Confusion Matrix:

```
[[13 5]
 [ 4 29]]
```

Gambar 10. Hasil evaluasi dengan rasio 80:20

Gambar 10 menunjukkan ketika data dibagi dengan rasio 80% untuk *data training* dan 20% untuk *data testing*, model berhasil mengklasifikasikan data dengan cukup baik. Berdasarkan *confusion matrix*, diperoleh hasil *accuracy* = 82%, *precision* 76%, *recall* 72%, *f1-score* 74%. Dari *confusion matrix*

diperoleh nilai *True Positive (TP)* = 13 yang berarti ada 13 siswa yang benar-benar tergolong sebagai *Bully* dan diklasifikasikan dengan benar sebagai *Bully* oleh model, *True Negative (TN)* = 29 yang berarti ada 29 siswa yang sebenarnya tergolong sebagai *Non-Bully* dan juga diprediksi dengan benar oleh model sebagai *Non-Bully*, *False Positive (FP)* = 4 yang berarti model salah memprediksi *Non-Bully* sebagai *Bully* dan *False Negative (FN)* = 5 yang berarti ada 4 siswa yang sebenarnya *Bully* tetapi salah diklasifikasikan sebagai *Non-Bully* oleh model.

➡ Akurasi Model: 0.85

Laporan Klasifikasi:

	precision	recall	f1-score	support
0	0.77	0.91	0.83	11
1	0.92	0.80	0.86	15
accuracy			0.85	26
macro avg	0.85	0.85	0.85	26
weighted avg	0.86	0.85	0.85	26

Confusion Matrix:

```
[[10 1]
 [ 3 12]]
```

Gambar 11. Hasil evaluasi dengan rasio 90:10

Gambar 11 menunjukkan bahwa ketika data dibagi dengan rasio 90% untuk *training* dan 10% untuk *testing*, model memperoleh hasil yang lebih baik dibandingkan dengan pembagian 80:20. Berdasarkan *confusion matrix*, diperoleh hasil *accuracy* = 85%, *precision*= 77, *recall*= 91%, dan *f1-score*= 83%. Dari *confusion matrix*, diperoleh nilai *True Positive (TP)*=10 yang berarti 10 siswa yang benar-benar tergolong sebagai *Bully* dan diklasifikasikan dengan benar sebagai *bully* oleh model, *True Negative (TN)* = 12 yang berarti ada 12 siswa yang sebenarnya tergolong sebagai *Non-Bully* dan juga diprediksi dengan benar sebagai *Non-Bully* oleh model, *False Positive (FP)* = 3 yang berarti model salah memprediksi *Non-Bully* sebagai *Bully* sebanyak 3 siswa, *False Negative (FN)* = 1 yang berarti ada 1 siswa yang sebenarnya *Bully* tetapi salah diklasifikasikan sebagai *Non-Bully* oleh model.

➡ Akurasi Model: 0.77

Laporan Klasifikasi:

	precision	recall	f1-score	support
0	0.59	0.70	0.64	23
1	0.86	0.80	0.83	54
accuracy			0.77	77
macro avg	0.73	0.75	0.73	77
weighted avg	0.78	0.77	0.77	77

Confusion Matrix:

```
[[16 7]
 [11 43]]
```

Gambar 12. Hasil evaluasi dengan rasio 70:30

Gambar 12 menunjukkan bahwa ketika data dibagi dengan rasio 70% untuk *training* dan 30% untuk *testing*, model memperoleh hasil cukup baik. Berdasarkan *confusion matrix*, diperoleh hasil *accuracy* = 77%, *precision*= 59%, *recall*=70%, dan

$f1\text{-score}$ = 64%. Dari *confusion matrix*, diperoleh nilai *True Positive (TP)* = 16 yang berarti 16 siswa yang benar-benar tergolong sebagai *Bully* dan diklasifikasikan dengan benar, *True Negative (TN)* = 43 yang berarti ada 43 siswa yang sebenarnya tergolong sebagai *Non-Bully* dan juga diprediksi dengan benar, *False Positive (FP)*=11 yang berarti model salah memprediksi *Non-Bully* sebagai *Bully* sebanyak 11 siswa, *False Negative (FN)*=7 yang berarti ada 7 siswa yang sebenarnya *Bully* tetapi salah diklasifikasikan sebagai *Non-Bully* oleh model.

Akurasi Model: 0.78

Laporan Klasifikasi:

	precision	recall	f1-score	support
0	0.59	0.68	0.63	28
1	0.87	0.82	0.85	74
accuracy			0.78	102
macro avg	0.73	0.75	0.74	102
weighted avg	0.80	0.78	0.79	102

Confusion Matrix:
[[19 9]
[13 61]]

Gambar 13. Hasil evaluasi dengan rasio 60:40

Gambar 13 menunjukkan bahwa ketika data dibagi dengan rasio 60% untuk training dan 40% untuk testing, model masih memperoleh hasil cukup baik. Berdasarkan *confusion matrix*, diperoleh hasil *accuracy* = 78%, *precision*= 59%, *recall*=68%, dan *f1-score*= 63%. Dari *confusion matrix*, diperoleh nilai *True Positive (TP)*=16 yang berarti 16 siswa yang benar-benar tergolong sebagai *Bully* dan diklasifikasikan dengan benar, *True Negative (TN)*=43 yang berarti ada 43 siswa yang sebenarnya tergolong sebagai *Non-Bully* dan juga diprediksi dengan benar, *False Positive (FP)*=7 yang berarti model salah memprediksi *Non-Bully* sebagai *Bully* sebanyak 7 siswa, *False Negative (FN)*=11 yang berarti ada 3 siswa yang sebenarnya *Bully* tetapi salah diklasifikasikan sebagai *Non-Bully* oleh model.

Akurasi Model: 0.77

Laporan Klasifikasi:

	precision	recall	f1-score	support
0	0.55	0.65	0.59	34
1	0.86	0.81	0.84	94
accuracy			0.77	128
macro avg	0.71	0.73	0.71	128
weighted avg	0.78	0.77	0.77	128

Confusion Matrix:
[[22 12]
[18 76]]

Gambar 14. Hasil evaluasi dengan rasio 50:50

Gambar 14 menunjukkan bahwa ketika data dibagi dengan rasio 50% untuk training dan 50% untuk testing, model memperoleh hasil cukup baik juga. Berdasarkan *confusion matrix*, diperoleh hasil *accuracy* = 77%, *precision*= 55%, *recall*=65%, dan *f1-score*= 59%. Dari *confusion matrix*, diperoleh nilai

True Positive (TP)=22 yang berarti 22 siswa yang benar-benar tergolong sebagai *Bully* dan diklasifikasikan dengan benar, *True Negative (TN)*=76 yang berarti ada 76 siswa yang sebenarnya tergolong sebagai *Non-Bully* dan juga diprediksi dengan benar, *False Positive (FP)*=18 yang berarti model salah memprediksi *Non-Bully* sebagai *Bully* sebanyak 18 siswa, *False Negative (FN)*=12 yang berarti ada 12 siswa yang sebenarnya *Bully* tetapi salah diklasifikasikan sebagai *Non-Bully* oleh model.

4.6 Knowledge Presentation

Berdasarkan hasil analisis data yang dikumpulkan dan di olah selama proses penelitian, data dapat divisualisasikan beberapa aspek dalam analisis perilaku perundungan di SDN Kiarapayung I, diantaranya :

- Pelaku perundungan. Dari data yang diperoleh, mayoritas kasus perundungan dilakukan oleh teman sekelas sebanyak 156 kejadian. Di posisi kedua, kakak kelas menjadi pelaku perundungan sebanyak 66 kejadian. Sementara itu sebanyak 24 siswa melaporkan tidak pernah mengalami perundungan dan adik kelas memiliki jumlah kejadian perundungan paling sedikit, yakni 9 kejadian.
- Waktu kejadian perundungan. Berdasarkan hasil visualisasi dari data yang diperoleh, perundungan paling sering terjadi yaitu pada saat siswa istirahat dengan jumlah 164 kejadian. Sementara itu, perundungan juga terjadi sebelum pulang sekolah sebanyak 35 kejadian dan saat jam pelajaran sebanyak 32 kejadian. Sebanyak 24 siswa menjawab tidak pernah pada poin ini dikarenakan siswa tersebut tidak pernah mengalami perundungan.
- Kekerasan fisik. Berdasarkan hasil visualisasi dari data yang diperoleh, kejadian kekerasan fisik menunjukkan variasi frekuensi yang berbeda dikalangan siswa. Jumlah siswa yang tidak pernah mengalami kekerasan fisik tercatat sebanyak 76 siswa yang merupakan kategori frekuensi dengan jumlah terbanyak. Sementara itu, 73 siswa melaporkan bahwa mereka pernah mengalami satu kali. 59 siswa melaporkan pernah mengalami kekerasan fisik 2 atau 3 kali dan 47 siswa mengalami kekerasan fisik lebih dari 3 kali.
- Kekerasan verbal. Berdasarkan hasil visualisasi dari data yang diperoleh, kejadian kekerasan verbal menunjukkan variasi frekuensi yang berbeda juga dikalangan siswa. Jumlah siswa yang tidak pernah mengalami kekerasan verbal tercatat sebanyak 81 siswa yang merupakan kategori frekuensi dengan jumlah terbanyak. Sementara itu, 72 siswa melaporkan bahwa mereka pernah mengalami satu kali. 59 siswa melaporkan pernah mengalami kekerasan fisik 2 atau 3 kali dan 43 siswa mengalami kekerasan fisik lebih dari 3 kali.
- Kekerasan psikologis / sosial. Berdasarkan hasil visualisasi dari data yang diperoleh, kejadian

kekerasan fisik menunjukkan variasi frekuensi yang berbeda dikalangan siswa. Jumlah siswa yang tidak pernah mengalami kekerasan fisik tercatat sebanyak 76 siswa yang merupakan kategori frekuensi dengan jumlah terbanyak. Sementara itu, 97 siswa melaporkan bahwa mereka pernah mengalami satu kali. 34 siswa melaporkan pernah mengalami kekerasan psikologis / sosial 2 atau 3 kali dan 48 siswa mengalami kekerasan fisik lebih dari 3 kali.

5. KESIMPULAN DAN SARAN

Hasil penelitian ini menunjukkan bahwa proses klasifikasi mencakup tahapan *data selection*, *data pre-processing*, *data transformation*, *data mining*, *evaluation* dan *knowledge presentation*. Model ini diuji dengan berbagai pembagian data training dan testing untuk mengukur kinerjanya dalam klasifikasi. Proses klasifikasi dalam penelitian ini dilakukan untuk mengklasifikasikan perilaku perundungan kedalam 2 kategori yaitu bully dan non-bully berdasarkan atribut kekerasan fisik, kekerasan verbal, kekerasan psikologis / sosial, pengalaman perasaan siswa dan kenyamanan siswa. Hasil evaluasi dari penerapan algoritma *Naïve Bayes Classifier* dalam mengklasifikasikan perilaku perundungan menunjukkan bahwa rasio 90:10 memberikan hasil terbaik dengan *accuracy* 85%, *precision* 77%, *recall* 91% dan *f1-score* 83%. Pengujian juga dilakukan dengan rasio lainnya seperti 80:20, 70:30, 60:40 dan 50:50 yang menunjukkan proporsi pembagian dataset sangat berpengaruh terhadap hasil evaluasi model. Selain itu, penggunaan *SMOTE* efektif dalam menyeimbangkan dataset yang meningkatkan kinerja model klasifikasi. Dari hasil penelitian yang telah dilakukan sebelumnya, saran yang diberikan untuk referensi penelitian selanjutnya dengan penggunaan algoritma klasifikasi lainnya selain *Naïve Bayes* dalam mengolah data kategorikal, data dan atribut pada dataset yang digunakan agar hasil klasifikasi yang diperoleh semakin akurat dan evaluasi dapat menggunakan metode lain selain *confusion matrix*, seperti *k-fold cross validation*.

DAFTAR PUSTAKA

- [1] Afrianda, C. R. P., Nayuko, A. M. P., & Lutfiah, P. (2024). Studi mengenai dampak bullying pada tingkat sekolah dasar di SDN Malabar Kota Bogor. *Indonesian Journal of Social Sciences and Humanities*, 4(1), 50-59.
- [2] Ahmad, N., Muslimin, A. A., & Sida, S. C. (2022). Analisis Perilaku Bullying Antar Siswa Terhadap Pembentukan Karakter Siswa di Sekolah Dasar Negeri Sangir Kecamatan Wajo Kota Makassar Sulawesi Selatan. *Naturalistic: Jurnal Kajian Dan Penelitian Pendidikan Dan Pembelajaran*, 7(1), 1318-1333.
- [3] Romli, I., & Zy, A. T. (2020). Penentuan jadwal overtime dengan klasifikasi data karyawan menggunakan algoritma C4. 5. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 4(2), 694-702.
- [4] Leza, M. A. A., Utami, N. W., & Dewi, P. A. C. (2024). Prediksi Prestasi Siswa Smas Katolik Santo Yoseph Denpasar Berdasarkan Kedisiplinan Dan Tingkat Ekonomi Orang Tua Menggunakan Metode Knowledge Discovery In Database Dan Algoritma Regresi Linier Berganda. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 373-379.
- [5] Kurniawan, D., & Yasir, M. (2022). Optimization Sentimen Analysis using CRISP-DM and Naive Bayes Methods Implemented on Social Media. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 6(2), 74-85.
- [6] Pardede, J., Miftahuddin, Y., & Kahar, W. (2020). Deteksi Komentar Cyberbullying Pada Media Sosial Berbahasa Inggris Menggunakan Naïve Bayes Classification. *Jurnal Informatika*, 7(1), 46-54.
- [7] Quroatun'Uyun, Z. (2020). Representasi Identitas Online-Offline dan Budaya Siber di lingkungan Akademik. *Jurnal Serambi Akademica*, 8(1), 61-69.
- [8] Wibowo, A. E., Khairi, A., Humairoh, H., Fadillah, M. I., & Hartawan, M. J. D. (2022). Text Mining: Sistem Prediksi Cyberbullying pada Platform Twitter menggunakan Logistic Regression, KNN, dan Naive Bayes. *Jurnal Rekayasa Elektro Sriwijaya*, 4(1), 17-23.
- [9] Alfari, M., Rizqy, M., Ghufro, R. I., Fathurahman, D., Saputra, R. D., & Kurniawan, F. (2023). Analisis Sentimen Persepsi Publik Terhadap Kasus Bullying Siswa Cilacap Menggunakan Pendekatan Machine Learning. *Journal of Information Technology Ampera*, 4(3), 265-276.
- [10] [14] Zai, C. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Jurnal Portal Data*, 2(3).
- [11] Arhami, M., Kom, M., & Muhammad Nasir, S. T. (2020). Data Mining-Algoritma dan Implementasi. Penerbit Andi.
- [12] Setiawan, Z., Fajar, M., Priyatno, A. M., Putri, A. Y. P., Aryuni, M., Yuliyanti, S., ... & Wijaya, A. (2023). Buku Ajar Data Mining. PT. Sonpedia Publishing Indonesia.
- [13] NURFA'IZAH, M. E. L. A. (2022). KLASIFIKASI PENENTUAN GIZI STUNTING PADA BALITA MENGGUNAKAN METODE NAÏVE BAYES (Doctoral dissertation, Universitas Islam Sultan Agung).
- [14] Erwin, I. M., Risnandar, E. P., & Sugiarto, B. (2020). Kayu7net: Identifikasi dan Evaluasi F-Measure Citra Kayu berbasis Deep Convolutional Neural Network (DCNN). *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(6).