SKRIPSI

Analisis Sentimen Terhadap Kekerasan Anak Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes

Jenis Skripsi:
Penelitian Eksperimental

Disusun Sebagai Salah Satu Syarat Memperoleh Gelar Sarjana Teknik (S.Kom.)

Disusun oleh:
Zella Brimanda Bela Corsanti
NIM. 18.0504.0019

Pembimbing:
Nuryanto, S.T., M.Kom.
NIDN. 0605037002

Pembimbing:
Maimunah, S.Si., M.Kom.
NIDN. 0612117702

Program Studi Teknik Informatika
Fakultas Teknik
Universitas Muhammadiyah Magelang
Tahun 2025

Bab 1 Pendahuluan

1.1 Latar Belakang Masalah

Era media sosial saat ini, platform seperti Twitter berperan sebagai sumber informasi yang signifikan untuk memahami tren, persepsi, dan reaksi masyarakat terhadap berbagai isu sosial, termasuk kekerasan terhadap anak. Teknologi internet telah mendorong masyarakat untuk menyuarakan opini mereka melalui media sosial, sehingga menciptakan peluang baru dalam menganalisis opini publik secara real-time (Dedi Darwis et al., 2020).

Salah satu pendekatan yang dapat digunakan untuk memahami persepsi publik secara sistematis adalah analisis sentimen. Analisis ini merupakan metode untuk mengidentifikasi dan mengkategorikan opini masyarakat terhadap suatu topik, seperti layanan publik, kebijakan pemerintah, atau isu sosial lainnya. Dengan demikian, analisis sentimen dapat digunakan sebagai dasar dalam pengambilan keputusan yang lebih berbasis data, khususnya dalam mengevaluasi respons publik terhadap suatu kebijakan atau fenomena sosial.

Dalam konteks kekerasan terhadap anak, yang merupakan isu sosial yang kompleks dan memiliki dampak jangka panjang terhadap kesejahteraan masyarakat, analisis sentimen dapat memberikan gambaran menyeluruh tentang respons emosional dan opini masyarakat. Twitter sebagai media sosial yang aktif digunakan dalam diskusi publik menjadi sumber data yang tepat untuk dianalisis lebih lanjut.

Penelitian ini menggunakan metode *Naive Bayes* karena kemampuannya dalam melakukan klasifikasi teks dengan efisien dan akurat. Pemilihan metode ini didasari oleh keinginan peneliti untuk memperdalam pengambilan keputusan berbasis data, dengan mengelompokkan tweet ke dalam kategori sentimen positif, negatif, dan netral. *Naive Bayes* merupakan algoritma probabilistik yang telah terbukti efektif dalam berbagai studi terdahulu dalam domain Text Mining dan Sentiment Analysis.(Dedi Darwis et al., 2020) (Hasri & Alita, 2022)

Beberapa penelitian sebelumnya juga menunjukkan efektivitas metode ini. Misalnya (Suryono et al., 2018), menunjukkan bahwa penggunaan *Naive Bayes* menghasilkan akurasi sebesar 66,79% dalam klasifikasi sentimen Twitter. Penelitian dari Dea Oktavia (Oktavia et al., 2023) berhasil mengklasifikasikan sentimen masyarakat pada sosial media Twitter mengenai penerapan e-Tilang. (Agustian et al., 2022) juga melaporkan akurasi 80% dalam menganalisis opini masyarakat tentang kendaraan listrik. Hal ini memperkuat alasan pemilihan metode ini dalam penelitian ini.

Melalui pendekatan ini, peneliti berharap dapat menggali pola, tren, dan respons publik secara lebih terstruktur, sehingga dapat memberikan masukan yang relevan bagi pembuat kebijakan, lembaga perlindungan anak, atau pihak terkait lainnya. Penelitian ini juga diharapkan mampu menghasilkan wawasan yang dapat mendukung pengambilan keputusan yang lebih tepat dalam penanganan isu kekerasan terhadap anak di Indonesia.

1.2 Perumusan Masalah

1. Bagaimana penerapan algoritma *Naive Bayes* dalam melakukan analisis sentimen terhadap tweet yang berkaitan dengan kekerasan pada anak?
2. Bagaimana menganalisis tingkat akurasi dari algoritma *Naive Bayes* dalam mengklasifikasikan data tweet berdasarkan sentiment?

1.3 Tujuan

1. Menerapkan algoritma *Naive Bayes* untuk melakukan klasifikasi sentimen terhadap tweet yang membahas kekerasan pada anak.
2. Mengukur dan menganalisis tingkat akurasi dari algoritma *Naive Bayes* dalam mengklasifikasikan data tweet berdasarkan sentimen.

1.4 Manfaat

Adapun manfaat yang dapat diperoleh dari penelitian ini adalah:

1. Manfaat Akademis: Menambah literatur dan referensi ilmiah dalam bidang text mining dan analisis sentimen, khususnya dalam penerapan algoritma *Naive Bayes* terhadap isu sosial di media sosial.
2. Manfaat Praktis: Memberikan wawasan bagi pemerintah, organisasi perlindungan anak, maupun masyarakat umum mengenai persepsi publik terhadap isu kekerasan pada anak yang dapat menjadi dasar pengambilan kebijakan atau tindakan lebih lanjut.

Bab 2 Studi Literatur

2.1 Tinjauan Penelitian Terdahulu

Penelitian dari (Nurjoko & Rahardi, 2024) Menunjukkan kekerasan verbal di media sosial, seperti Twitter, merupakan fenomena yang semakin sering terjadi dan berpotensi menimbulkan dampak psikologis serius terhadap individu atau kelompok. model Indo-BERT dikembangkan untuk mengidentifikasi perilaku kekerasan verbal di Twitter melalui pendekatan analisis sentimen bertema *Verbal Violence Behavior* (VVB). Penelitian ini menunjukkan bahwa kekerasan verbal sering berupa ujaran yang merendahkan, menghina, atau mengancam, yang sayangnya masih lazim ditemukan di platform daring.

Penelitian yang dilakukan oleh (Hasri & Alita, 2022) dengan judul "Penerapan Metode Naïve Bayes Classifier Dan *Support Vector Machine* Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter" Membahas mengenai penerapan metode *Naive Bayes* Classifier dan *Support Vector Machine* untuk mengklasifikasikan sentimen masyarakat terhadap dampak dari covid-19. Saat melakukan analisis cross validation diketahui prediksi salah dalam hal pengujian, dimana kategori seharusnya netral tetapi berlabel positif maupun sebaliknya. Kesalahan ini mencerminkan keterbatasan dari algoritma *Naive Bayes* yang mengasumsikan independensi antar fitur dan kurang mampu memahami konteks kalimat secara utuh.

Selanjutnya penelitian dari (Agustian et al., 2022) dengan judul "Penerapan Analisis Sentimen Dan *Naive Bayes* Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter" Membahas tentang analisis sentimen terhadap kendaraan listrik. Data dikumpulkan menggunakan Twitter API, kemudian dilakukan proses pra-pemrosesan seperti pembersihan data dan penerjemahan ke dalam bahasa Inggris. Selanjutnya, klasifikasi sentimen dilakukan menggunakan algoritma *Naïve Bayes*, dengan bantuan pustaka Python seperti TextBlob dan WordCloud untuk visualisasi. Hasil penelitian menunjukkan bahwa sebagian besar tanggapan masyarakat terhadap kendaraan listrik bersifat positif, dengan nilai akurasi sebesar 80%, presisi 82%, dan recall 44%. Penelitian ini membuktikan bahwa analisis sentimen berbasis media sosial efektif untuk menangkap persepsi publik terhadap isu-isu teknologi dan kebijakan pemerintah.

Penelitian lain dari (Soemedhy et al., 2022) yang berjudul "Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube "Kekerasan Seksual")" Penelitian ini bertujuan untuk menganalisis respons masyarakat terhadap isu pelecehan seksual dengan menggunakan pendekatan analisis sentimen. Sumber data dalam penelitian ini diperoleh dari komentar pengguna pada platform YouTube. Metode yang digunakan melibatkan algoritme pembelajaran mesin, yaitu *Support Vector Machine* (SVM), Naive Bayes, dan Random Forest. Ketiga algoritma tersebut kemudian dibandingkan untuk mengevaluasi kinerja masing-masing dalam mengklasifikasikan sentimen publik. Berdasarkan hasil pengujian, algoritme Random Forest menunjukkan performa terbaik dengan akurasi mencapai 78%, melampaui algoritme SVM dan *Naive Bayes* dalam mengolah data komentar terkait topik pelecehan seksual..

Lalu penelitian dari (Romadhoni & Holle, 2022) berjudul "Analisis Sentimen Terhadap PERMENDIKBUD No.30 pada Media Sosial Twitter Menggunakan Metode *Naive Bayes* dan LSTM" Penelitian ini membandingkan performa metode *Naïve Bayes* dan *Long Short-Term Memory (LSTM)* dalam analisis sentimen. Hasil menunjukkan bahwa meskipun jumlah data dan tahapan *preprocessing* yang digunakan sama, performa LSTM lebih unggul dibandingkan Naïve Bayes. Namun, metode LSTM membutuhkan waktu komputasi yang lebih lama dibandingkan Naïve Bayes. Kekurangan dalam penelitian ini adalah jumlah data yang terbatas, padahal metode deep learning seperti LSTM idealnya memerlukan dataset yang besar. Selain itu, perbandingan antar metode dinilai kurang seimbang karena membandingkan algoritma dari kategori yang berbeda (deep learning dan machine learning), sehingga disarankan untuk membandingkan antar metode yang sejenis pada penelitian selanjutnya.

Berdasarkan hasil analisis yang telah dilakukan, dapat disimpulkan bahwa metode Naïve Bayes merupakan pendekatan yang efisien dan efektif untuk melakukan klasifikasi sentimen pada data Twitter terkait isu kekerasan terhadap anak. Meskipun metode ini tergolong sederhana dibandingkan algoritma lain seperti *Support Vector Machine* (SVM) atau *Long Short-Term Memory* (LSTM), namun *Naïve Bayes* tetap mampu menghasilkan tingkat akurasi klasifikasi yang baik dengan waktu komputasi yang relatif cepat.

Dibandingkan dengan SVM yang memerlukan penyesuaian parameter, dan LSTM yang membutuhkan jumlah data besar serta waktu pelatihan yang lama, Naïve Bayes lebih unggul dalam hal kemudahan implementasi dan efisiensi proses, khususnya pada dataset berukuran kecil hingga menengah seperti tweet. Oleh karena itu, metode ini tepat digunakan dalam penelitian **ini** karena dapat memberikan hasil yang relevan dengan sumber daya yang tersedia.

2.2 Kajian Teoretis

2.2.1 Analisis sentimen

Analisis sentimen merupakan proses menganalisis teks digital untuk menentukan apakah emosi atau sikap yang diekspresikan dalam pesan tersebut bersifat positif, negatif, atau netral. Istilah lain yang sering digunakan adalah *opinion mining* atau *emotion AI*, yaitu teknik yang memakai pemrosesan bahasa alami (NLP) untuk mendeteksi dan mengukur informasi subjektif dan keadaan emosional dari suatu teks. Menurut (Medhat et al., 2014) analisis sentimen adalah proses otomatis untuk mengetahui apakah suatu teks bernada positif, negatif, atau netral. Mereka menyebutkan bahwa analisis sentimen bisa dilakukan dengan beberapa pendekatan, seperti menggunakan kamus kata (lexicon-based), algoritma pembelajaran mesin (machine learning), atau gabungan keduanya (hybrid). Sementara itu, Liu dalam artikelnya *Sentiment Analysis and Opinion Mining* menjelaskan bahwa analisis sentimen bukan hanya tentang klasifikasi teks, tapi juga memahami siapa yang memberikan opini, kepada siapa opini itu ditujukan, serta dalam konteks waktu dan tempat tertentu. Dalam praktiknya, analisis sentimen sering digunakan oleh perusahaan untuk mengetahui pendapat konsumen terhadap produk atau layanan mereka, membantu evaluasi, dan mendukung pengambilan keputusan dalam strategi bisnis.

2.2.2 Text Mining

Text mining adalah cabang dari *data mining* yang fokus pada penggalian informasi dari data berbentuk teks atau dokumen. Proses ini menggunakan alat dan teknik dari data mining untuk menelusuri kata-kata penting serta mengekstrak informasi yang bermanfaat. Text mining juga termasuk dalam bidang *Artificial Intelligence* karena melibatkan pemrosesan otomatis terhadap data teks untuk menghasilkan pengetahuan (Idris et al., 2023). Text Mining menggabungkan berbagai teknik dari bidang lain seperti pengambilan informasi, data mining, statistik, matematika, machine learning, pemrosesan bahasa alami (NLP), linguistik, dan visualisasi data. Dalam penerapannya, text mining memiliki beberapa tahapan penting, yaitu: ekstraksi dan dokumentasi teks, pra-pemrosesan data, pengumpulan data statistik dan proses indexing, serta analisis isi untuk memperoleh informasi yang dibutuhkan(Idris & Mustofa, 2022)

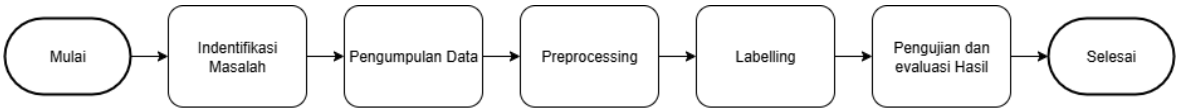
2.2.3 Naïve Bayes

Salah satu algoritma klasifikasi yang paling banyak digunakan dalam penggalian data dan teks adalah *Naive Bayes* Classifier. Algoritma ini didasarkan pada teorema Bayes bahwa setiap kegiatan memberikan kontribusi yang sama penting atau saling bebas untuk memilih kelas tertentu. Dalam analisis sentimen, *Naive Bayes* Classifier sering digunakan untuk membagi teks atau dokumen ke dalam kategori positif, negatif, atau netral(Fakhri et al., 2023). *Naive Bayes* dikenal sebagai metode yang sederhana dan efisien, namun tetap efektif untuk menangani dataset berukuran besar. Metode ini banyak digunakan dalam berbagai aplikasi, seperti klasifikasi teks, deteksi spam, analisis sentimen, hingga bidang kesehatan seperti diagnosis penyakit.

Bab 3 Metode Penelitian

3.1 Prosedur Penelitian

Prosedur penelitian merupakan langkah-langkah yang ditempuh secara sistematis untuk mencapai tujuan penelitian. Dalam penelitian ini, prosedur dimulai dari identifikasi masalah, pengumpulan data, pra-pemrosesan, pembagian data, transformasi data menjadi bentuk numerik, hingga klasifikasi menggunakan algoritma.Alasan penggunaan algoritma Naïve Bayes adalah karena algoritma ini sederhana, efisien, dan terbukti efektif dalam klasifikasi teks, khususnya analisis sentimen. Naïve Bayes mampu bekerja dengan baik pada data berdimensi tinggi seperti teks, serta memiliki kemampuan generalisasi yang baik meskipun dengan jumlah data terbatas



Gambar 3. 1 Flowchart Penelitian

Penelitian dimulai dengan identifikasi masalah, dilanjutkan dengan pengumpulan data melalui proses *crawling* dari Twitter.Setelah data terkumpul, dilakukan pra-pemrosesan yang mencakup *cleansing*, *case folding*, *tokenizing*, dan *stemming* untuk menyiapkan data agar siap dianalisis kemudian dilakukan pelabelan sentimen (positif, negatif, atau netral). Data kemudian dibagi menjadi data training dan testing. Selanjutnya, data diubah menjadi bentuk numerik menggunakan metode TF-IDF, lalu diklasifikasikan menggunakan algoritma Naive Bayes. Terakhir, dilakukan pengujian model untuk mengetahui akurasi, dan hasil akhir dianalisis untuk menarik kesimpulan.

1. Identifikasi Masalah

Kekerasan pada anak masih menjadi permasalahan sosial yang serius dan kompleks, namun pemanfaatan media sosial seperti Twitter sebagai sumber informasi publik terkait isu ini belum optimal. Di sisi lain, opini masyarakat yang tersebar di media sosial dapat menjadi sumber data penting untuk memahami persepsi dan reaksi terhadap kekerasan anak. Masalahnya, masih minim metode yang mampu menganalisis data tersebut secara efisien dan akurat. Oleh karena itu, dibutuhkan pendekatan analisis sentimen yang tepat, seperti metode Naive Bayes, untuk mengelompokkan tweet berdasarkan kemiripan sentimen sehingga dapat membantu dalam mengungkap pola opini publik yang dominan terhadap isu kekerasan pada anak.

2. Pengumpulan data

Penelitian ini menggunakan model alur dengan langkah awal yang melibatkan pengambilan data dari Twitter API melalui proses *crawling*. Untuk mengakses Twitter, *API KEY* diperlukan dalam pengambilan data Twitter. Untuk proses penggalian data dari twitter menggunakan bantuan library *Twit-harvest*. Data tersebut dikumpulkan dengan menggunakan kata kunci "kekerasan pada anak" dalam rentang tanggal mulai dari januari 2020 hingga maret 2025. Hasil pengambilan data Twitter ini menghasilkan sebanyak 1589 tweet . Data Twitter ini mencakup informasi tentang nama pengguna (*username*), waktu pembuatan tweet (*tweetcreated*),

dan teks dari tweet itu sendiri. Selanjutnya, data Twitter yang berhasil diambil akan diolah menggunakan bahasa pemrograman Python. Hasil dari proses scraping data Twitter kemudian disimpan ke dalam drive untuk pengolahan lebih lanjut.

3. Preprocessing

Tahapan selanjutnya yaitu melakukan *text preprocessing* yang bertujuan untuk mengurangi gangguan (*noise*) dalam data teks agar dapat diolah secara optimal. Proses ini mencakup berbagai tahapan rutin yang diperlukan untuk menyiapkan data sebelum digunakan dalam proses penemuan pengetahuan (knowledge discovery) dalam sistem text mining. Data teks akan melalui beberapa tahap seperti case folding, tokenisasi, filtering, dan stemming. Melalui tahap ini, data yang dihasilkan menjadi lebih bersih dan terstruktur, sehingga mempermudah proses analisis dan pemrosesan lebih lanjut dalam sistem(Yuniar et al., 2022).

Tahap *text preprocessing* mencakup beberapa proses penting yang berfungsi untuk membersihkan data teks sebelum analisis lebih lanjut. Proses pertama adalah *case folding*, yaitu mengubah seluruh huruf dalam teks menjadi huruf kecil untuk menghindari ketidakkonsistenan dalam penulisan huruf kapital(Sasongko et al., 2023). Selanjutnya adalah proses *filtering*, yang dilakukan setelah tokenisasi, bertujuan untuk menyaring kata-kata penting dan membuang kata-kata yang tidak memiliki makna signifikan atau dikenal sebagai *stopword*. *Stopword* merupakan kata-kata umum seperti "dan", "atau", "yang", dan sebagainya yang tidak mencerminkan isi dokumen secara spesifik. Kata-kata tersebut dihapus berdasarkan daftar *stoplist* yang telah ditentukan sebelumnya. Proses terakhir adalah *stemming*, yaitu mengubah kata menjadi bentuk dasarnya dengan menghilangkan imbuhan seperti prefiks, sufiks, maupun konfiks. Pada penelitian ini, *stemming* dilakukan berdasarkan kaidah morfologi bahasa Indonesia agar kata-kata yang digunakan menjadi lebih seragam dan relevan untuk analisis lebih lanjut(Arista et al., 2023).

4. Data Labelling

Setelah melalui tahap *preprocessing*, data kemudian diklasifikasikan ke dalam atribut kelas sentimen. Berbeda dengan metode tradisional yang menggunakan kamus leksikon dalam format .txt untuk mendeteksi kata-kata positif, negatif, dan netral, penelitian ini menggunakan model BERT (*Bidirectional Encoder Representations from Transformers*) yang mampu memahami konteks kalimat secara lebih mendalam. Dengan memanfaatkan kemampuan representasi kontekstual dari BERT, proses klasifikasi sentimen dilakukan secara otomatis berdasarkan pembelajaran dari data latih. Adapun kelas sentimen dalam penelitian ini terdiri atas tiga kategori, yaitu positif, negatif, dan netral(Armaeni et al., 2024)

5. Term Frequency – Inverse Document Frequency TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah metode pembobotan kata atau istilah yang memberikan bobot berbeda pada setiap kata dalam sebuah dokumen (Cahyani & Patasik, 2021), berdasarkan seberapa sering kata tersebut muncul dalam satu dokumen dan seberapa jarang kata itu muncul di seluruh kumpulan dokumen. Metode ini digunakan dalam penelitian karena mampu memberikan kinerja yang lebih baik, khususnya dalam meningkatkan nilai *recall* dan *precision*. Model TF-IDF terdiri dari empat tahapan utama.

Model TF-IDF menunjukkan kinerja yang lebih baik dibandingkan dengan model Word2Vec, terutama dalam kondisi jumlah data yang tidak seimbang pada setiap kelas emosi. Beberapa kelas, seperti emosi terkejut, termasuk dalam kategori minoritas karena memiliki

jumlah data yang jauh lebih sedikit dibandingkan kelas lainnya. Dalam kondisi data yang terbatas, Word2Vec mengalami kesulitan dalam menangkap informasi semantik dan sintaktik dari kata-kata secara optimal, karena model tersebut memerlukan data pelatihan dalam jumlah besar untuk membangun representasi kata yang akurat. Sebaliknya, TF-IDF mampu menghasilkan akurasi yang baik meskipun digunakan pada dataset dengan jumlah data yang relatif kecil(Cahyani & Patasik, 2021).

6. Splitting Data

Pada penelitian ini, pembagian data dilakukan dengan dua teknik pembagian yang berbeda, yaitu sebagai berikut.

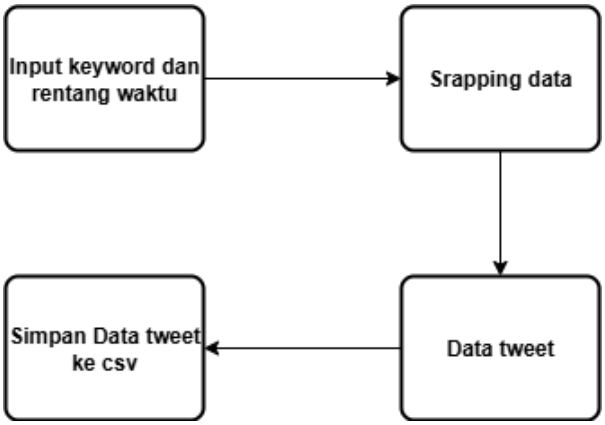
a. Hold-Out

Metode Hold-Out merupakan salah satu teknik pembagian data yang digunakan dalam proses pengujian model, di mana dataset dibagi menjadi dua bagian, yaitu data latih dan data uji. Dalam metode ini, satu bagian data digunakan untuk melatih model, sementara bagian lainnya digunakan untuk menguji kinerja model yang telah dilatih.

b. K-Fold Cross Validation

K-Fold Cross Validation adalah salah satu teknik validasi yang digunakan untuk mengukur tingkat akurasi sebuah model prediksi atau klasifikasi berdasarkan dataset tertentu. Dalam metode ini, data dibagi menjadi *K* bagian atau *fold* yang kurang lebih sama besar. Proses pelatihan dan pengujian dilakukan sebanyak *K* kali, di mana pada setiap iterasi, satu bagian digunakan sebagai data uji dan sisanya sebagai data latih. Hasil dari setiap iterasi kemudian dirata-rata untuk memperoleh performa akhir model yang lebih stabil dan dapat diandalkan.

3.2 Analisa sistem

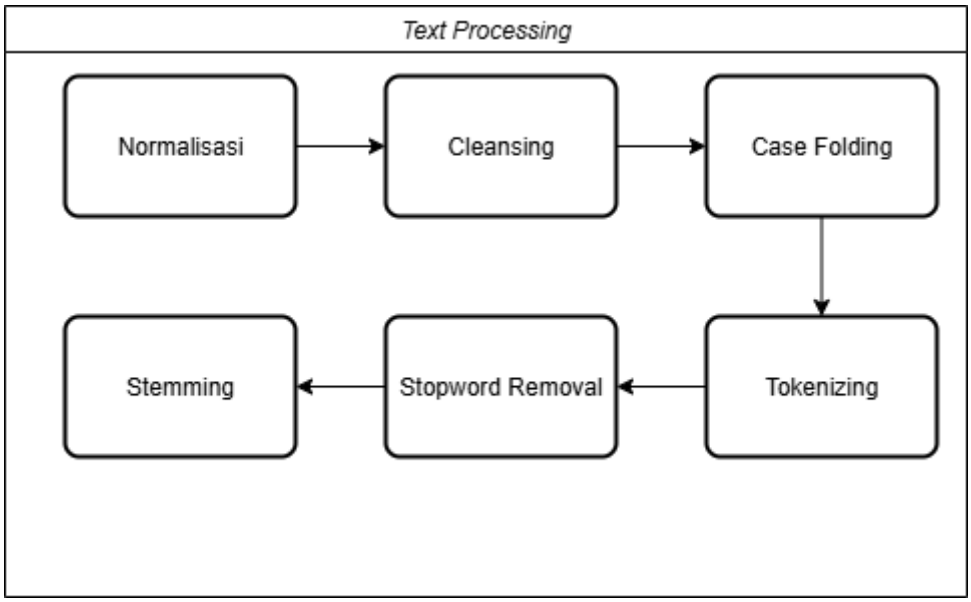


Gambar 3. 2 Alur Scrapping data twitter

Sampel pada penelitian ini adalah tweet netizen Indonesia yang terkait kekerasan pada anak yaitu yang mengandung kata-kata kunci "kekerasan anak" OR "pelecehan pada anak". Periode pengambilan data tweet dari Januari 2020 sampai Maret 2025. Hasil webscraping (gambar 3.1) pada Twitter didapatkan jumlah tweet 1306 tweet berbahasa Indonesia kemudian dilakukan tahap preprocessing.

Data yang digunakan dalam penelitian ini berupa kumpulan tweet dari platform Twitter yang mengandung kata kunci terkait kekerasan pada anak. Data ini dikumpulkan menggunakan metode crawling, salah satunya dengan bantuan Twit Harvest, yang memungkinkan pengambilan data publik dari Twitter menggunakan autentikasi API. Setelah data terkumpul dan disimpan dalam format CSV, langkah selanjutnya adalah melakukan proses pra-pemrosesan teks (text preprocessing) untuk mempersiapkan data sebelum dianalisis lebih lanjut menggunakan algoritma analisis sentimen.

Proses preprocessing ini meliputi beberapa tahapan, dimulai dari cleaning (pembersihan data), normalisasi, tokenisasi, hingga stemming. Pada tahap cleaning, dilakukan pembersihan teks dari elemen-elemen yang tidak relevan seperti URL, mention (kata yang diawali dengan simbol @), hashtag, angka, tanda baca, karakter khusus, hingga emoji, dengan menggunakan bantuan library re (regular expression). Selanjutnya, seluruh teks diubah menjadi huruf kecil (lowercase) untuk menyeragamkan format kata. Setelah teks dibersihkan, dilakukan proses normalisasi, yaitu mengganti kata-kata tidak baku atau singkatan (seperti "gk", "tdk", "sm", "yg") menjadi bentuk bakunya dengan bantuan kamus normalisasi. Kemudian dilakukan tokenisasi menggunakan library seperti nltk atau Sastrawi, yaitu memecah kalimat menjadi potongan-potongan kata (token) agar bisa dianalisis per kata. Tahap terakhir adalah stemming, yaitu mengubah kata ke bentuk dasarnya menggunakan library Sastrawi yang secara khusus dikembangkan untuk Bahasa Indonesia. Stemming penting dilakukan agar kata-kata yang memiliki makna serupa (misalnya "memakan", "dimakan", "makanannya") dapat dikenali sebagai satu bentuk kata dasar, yaitu "makan". Semua tahapan ini bertujuan untuk menyederhanakan dan menstandarkan data teks agar dapat dianalisis secara akurat dalam proses klasifikasi sentimen. Tahapan-tahapan bisa dilhat pada gambar 3.2



Gambar 3. 3 Text Preprocessing

Setelah proses pra-pemrosesan teks selesai dilakukan, tahap berikutnya adalah pelabelan sentimen terhadap setiap tweet yang telah dibersihkan. Dalam penelitian ini, digunakan model BERT (*Bidirectional Encoder Representations from Transformers*) untuk melakukan klasifikasi sentimen secara otomatis. BERT merupakan model deep learning berbasis transformer yang dikembangkan oleh Google dan mampu memahami konteks kata dalam sebuah kalimat secara dua arah, sehingga sangat efektif dalam memahami nuansa bahasa alami, termasuk Bahasa Indonesia. Untuk keperluan penelitian ini, digunakan model BERT yang telah dilatih untuk Bahasa Indonesia, seperti IndoBERT atau IndoBERTweet, yang tersedia melalui pustaka transformers dari Hugging Face. Model ini digunakan untuk mengklasifikasikan setiap tweet ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Proses pelabelan dilakukan dengan cara memuat model BERT yang telah dilatih (*pretrained*), kemudian memberikan input berupa teks tweet yang telah melalui tahap preprocessing. Model ini akan mengeluarkan probabilitas untuk masing-masing kelas sentimen, dan kelas dengan probabilitas tertinggi akan dijadikan label akhir untuk tweet tersebut. Keunggulan penggunaan BERT adalah kemampuannya dalam memahami konteks kata secara lebih mendalam dibandingkan metode klasik seperti *Naive Bayes* atau SVM, sehingga diharapkan hasil pelabelan sentimen yang diperoleh lebih akurat dan sesuai dengan konteks yang sebenarnya. Proses ini dilakukan menggunakan bantuan library transformers, torch, dan pandas dalam lingkungan Python.

Pengujian model dilakukan dengan menghitung metrik evaluasi seperti akurasi, precision, recall, dan F1-score. Metrik ini memberikan gambaran sejauh mana model mampu mengklasifikasikan sentimen dengan benar. Hasil pengujian dari ketiga model kemudian dibandingkan untuk menentukan model mana yang memberikan performa terbaik dalam menganalisis sentimen masyarakat terhadap program makan gratis pemerintah. Secara keseluruhan, tahapan ini bertujuan untuk memperoleh model yang paling akurat dan andal dalam menginterpretasi opini publik berbasis data media sosial.

3.3 Simulasi

3.3.1 Preprocessing

Langkah ini bertujuan membersihkan teks sebelum masuk ke algoritma. berikut adalah contoh simulasi tahapan preprocessing.

Contoh teks: "Kekerasan seksual pada anak-anak adalah kejahatan kemanusiaan. Siapa pun pelakunya apalagi pejabat kepolisian harus dihukum seberat-beratnya"

Cleansing	Tokenizing	Stopword	Stemming
kekerasan seksual anak anak kejahatan kemanusiaan pelaku pejabat kepolisian hukum berat	kekerasan, seksual, anak, anak, kejahatan, kemanusiaan, pelaku, pejabat, kepolisian, hukum, berat	kekerasan, seksual, anak, kejahatan, kemanusiaan, pelaku, pejabat, kepolisian, hukum, berat	keras, seksual, anak, anak, jahat, manusia, laku, jabat, polisi, hukum, berat

a. Dataset latih

Kelas negatif

- D1 (neg): kekerasan seksual anak pelaku dihukum → stems: [keras, seksual, anak, pelaku, hukum]
- D2 (neg): pejabat polisi terlibat kejahatan dihukum → [jabat, polisi, terlibat, jahat, hukum]

Kelas netral

- D3 (neu): laporan kasus dilaporkan polisi investigasi berjalan → [lapor, kasus, polisi, investigasi, jalan]
- D4 (neu): penanganan korban mendapatkan perlindungan → [tangan, korban, dapat, lindung]

Kelas positif

- D5 (pos): penegakan hukum proses pelaku hukuman tegas → [penegak, hukum, proses, pelaku, tegas]
- D6 (pos): hak anak dilindungi hukum layanan bantuan → [hak, anak, lindung, hukum, layanan, bantu]

Total = 6 (2 tiap kelas) → **Prior P(class) = 2/6 = 1/3 ≈ 0.3333** untuk tiap kelas (sederhana/imbang).

3.3.2 Splitting data dan Implementasi Algoritma Naïve Bayes

Proses pembagian data pada penelitian ini dibagi menjadi dua bagian untuk melakukan percobaan klasifikasi, yaitu menggunakan hold-out dan cross validation.

a. Hold-Out

Proses pembagian data training dan data testing yang dilakukan dengan menggunakan metode HoldOut membagi data training sebesar 80% dan data testing sebesar 20%. Proses pembagian data ini dilakukan dengan menggunakan bahasa pemrograman Python.

Pengujian menggunakan confusion matrix dilakukan untuk menguji model yang diimplementasikan pada data training dan data testing sehingga menghasilkan matriks berukuran 3x3. Tabel confusion matrix dapat dilihat pada tabel 1.

Tabel 1 Confusion Matrix

Kelas Sebelumnya	Kelas Prediksi		
		1	0
	1	TP	FN
	0	FP	TN

Keterangan:

TP (True Positive) = jumlah dokumen dari kelas 1 yang benar diklasifikasikan sebagai kelas 1

TN (True Negative) = jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0

FP (False Positive) = jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1

FN (False Negative) = jumlah dokumen dari kelas 1yang salah diklasifikasikan sebagai kelas 0

Rumus confusion matrix untuk menghitung accuracy, precision, dan recall seperti berikut.

$$Accuracy = \frac{TP + TN}{Total} \quad Precision = \frac{TP}{FP + FP} \quad recall = \frac{TP}{TP + FN} \qquad (3.1)$$

b. K-Cross Validation

Pembagian data dengan menggunakan metode K-fold cross validation dimana data akan diproses sebanyak K kali. Contoh dataset yang digunakan adalah 100 data, kemudian dibagi menjadi 5 bagian secara merata. Sehingga setiap fold terdiri dari 20 data. Hasil precision, recall, f1-score, dan akurasi dari algoritma Naïve Bayes dengan metode pembagian data menggunakan 5-fold cross valdation

Langkah-langkah:

1. Proses dilakukan sebanyak 5 iterasi, dengan skema sebagai berikut:

Tabel 2 Iteras ke-1

Iterasi	Data Latih (80%)	Data Uji (20%)
1	$D_2 + D_3 + D_4 + D_5$	D_1
2	$D_1 + D_3 + D_4 + D_5$	D_2
3	$D_1 + D_2 + D_4 + D_5$	D_3
4	$D_1 + D_2 + D_3 + D_5$	D_4
5	$D_1 + D_2 + D_3 + D_4$	D_5

2. Untuk setiap iterasi:
   a. Model Naïve Bayes dilatih pada data latih
   b. Model diuji pada data uji
   c. Dihitung nilai TP, TN, FP, dan FN.
   d. Dihitung metrik evaluasi: Accuracy, Precision, Recall, F1-Score

3. Contoh perhitungan Evaluasi (Iterasi ke-1).

Tabel 3 Perhitungan Evaluasi

	Prediksi Positif	Prediksi Negatif
Aktual Positif	TP = 16	FN = 4
Aktual Negatif	FP = 3	TN = 17

4. Rumus evaluasi:

a. $Accuracy = \frac{TP+TN}{Total} = \frac{16+17}{16+17+3+4} = \frac{33}{40} = 0.825$ (3.2)

b. $Precision = \frac{TP}{TP+FP} = \frac{16}{16+3} = \frac{16}{19} = 0.842$ (3.3)

c. $Recall = \frac{TP}{TP+FN} = \frac{16}{16+4} = \frac{16}{20} = 0.80$ (3.4)

d. $F1 - Score = 2\ x\ \frac{Precision\ x\ recall}{Precision+recall} = 2\ x\ \frac{0.842\ x\ 0.80}{0.842+0.80} = 0.820$ (3.5)

5. Ulangi untuk semua fold.
Proses perhitungan yang sama dilakukan pada setiap fold (fold ke-2 sampai ke-5)

Tabel 4 Perhitungan iterasi fold

Fold	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
1	16	17	3	4	0.825	0.842	0.80	0.820
2	15	18	2	5	0.825	0.882	0.75	0.810
3	17	16	4	3	0.825	0.809	0.85	0.829
4	14	19	1	6	0.825	0.933	0.70	0.800
5	15	18	2	5	0.825	0.882	0.75	0.810

6. Hitung rata-rata K-fold

$Accuracy\ avg = \frac{0.825+\ 0.825+\ 0.825+\ 0.825+\ 0.825}{5} = 0.825$ (3.6)

$Precision\ avg = \frac{0.842+\ 0.882+\ 0.809+\ 0.933+\ 0.882}{5} = 0.8696$ (3.7)

$Recall\ avg = \frac{0.80+\ 0.75+\ 0.85+\ 0.70+\ 0.75}{5} = 0.77$ (3.8)

$F1\ avg = \frac{0.820+\ 0.810+\ 0.829+\ 0.800+\ 0.810}{5} = 0.8139$ (3.9)

Contoh perhitungan K-Fold Cross Validation di atas bertujuan untuk memberikan gambaran bagaimana proses evaluasi model klasifikasi dilakukan menggunakan algoritma Naïve Bayes. Dengan membagi data menjadi beberapa fold dan melakukan pelatihan serta pengujian secara bergantian, peneliti dapat memperoleh nilai rata-rata metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *f1-score* sebagai acuan awal performa model.

Bab 5 Penutup

5.1 Kesimpulan

Berdasarkan hasil analisis sentimen menggunakan algoritma Naive Bayes terhadap tweet yang berkaitan dengan isu kekerasan pada anak, diperoleh akurasi klasifikasi sebesar 0.87 atau 86,6%. Model mampu melakukan klasifikasi dengan baik pada kelas negative dan neutral, masing-masing ditunjukkan oleh nilai F1-score sebesar 0,88 dan 0,89. Namun, pada kelas positive performa model rendah akibat ketidakseimbangan jumlah data. Secara umum, hasil penelitian ini menunjukkan bahwa sentimen masyarakat di media sosial Twitter terhadap isu kekerasan pada anak didominasi oleh sentimen negatif, yang mencerminkan adanya kecaman, penolakan, dan keprihatinan yang kuat dari pengguna terhadap tindakan kekerasan tersebut. Sebagian tweet bersifat netral, sedangkan sentimen positif muncul dalam jumlah yang jauh lebih sedikit dibandingkan dua kelas lainnya.

5.2 Saran

Berdasarkan hasil evaluasi, kinerja model pada kelas positive sangat rendah akibat ketidakseimbangan data. Oleh karena itu, untuk penelitian selanjutnya disarankan melakukan perbaikan distribusi data agar lebih seimbang, misalnya dengan metode oversampling seperti SMOTE atau menambah data positive secara manual. Selain itu, dapat dipertimbangkan penggunaan algoritma lain yang lebih mampu menangani data imbalanced, seperti Logistic Regression, Random Forest, atau model berbasis deep learning seperti LSTM

Referensi

Agustian, A., Tukiro, T., & Nurapriani, F. (2022). Penerapan Analisis Sentimen Dan *Naive Bayes* Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter. *Jurnal TIKA*, *7*(3), 243–249. https://doi.org/10.51179/tika.v7i3.1550

Arista, T. D., Yusra, Y., Fikry, M., & Oktavia, L. (2023). Klasifikasi Sentimen Masyarakat di Twitter terhadap Kenaikan Harga BBM dengan Metode K-NN. *JUKI: Jurnal Komputer Dan …*, *5*, 140–150. https://doi.org/https://doi.org/10.53842/juki.v5i1.189

Armaeni, P. P., Arya, I. K., Wiguna, G., Gede, W., & Parwita, S. (2024). *Sentiment Analysis of YouTube Comments on the Closure of TikTok Shop Using Naïve Bayes and Decision Tree Method Comparison*. *1*(2), 70–80.

Cahyani, D. E., & Patasik, I. (2021). Performance comparison of tf-idf and word2vec models for emotion text classification. *Bulletin of Electrical Engineering and Informatics*, *10*(5), 2780–2788. https://doi.org/10.11591/eei.v10i5.3157

Dedi Darwis, Nery Siskawati, & Zaenal Abidin. (2020). Penerapan Algoritma *Naive Bayes* untuk Analisis Sentimen Review Data Twitter BMKG Nasional. *Jurnal TEKNO KOMPAK*, *15*(1), 131–145.

Fakhri, J., Sunge, A. S., & Turmudi zy, A. (2023). Perancangan Klasifikasi Algoritma *Naive Bayes* Pada Data Pemilihan Jurusan Siswa. *JTT (Jurnal Teknologi Terpadu)*, *11*(2), 260–269. https://doi.org/10.32487/jtt.v11i2.1823

Hasri, C. F., & Alita, D. (2022). Penerapan Metode Naã□ Ve Bayes Classifier Dan *Support Vector Machine* Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, *3*(2), 145–160. https://doi.org/10.33365/jatika.v3i2.2026

Idris, I. S. K., & Mustofa, Y. A. (2022). Typo Checking Menggunakan Algoritma Rabin-Karp. *Jambura Journal of Electrical and Electronics Engineering*, *4*(1), 87–91. https://doi.org/10.37905/jjeee.v4i1.12150

Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Mengunakan Algoritma *Support Vector Machine* (SVM). *Jambura Journal of Electrical and Electronics Engineering*, *5*(1), 32–35. https://doi.org/10.37905/jjeee.v5i1.16830

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Nurjoko, & Rahardi, A. (2024). Model Indo-BERT Untuk Identifikasi Sentimen Kekerasan Verbal Di Twitter. *Jurnal Teknika*, *18*(2), 583–593. https://jurnal.polsri.ac.id/index.php/teknika/article/view/8897

Oktavia, D., Ramadahan, Y. R., & Minarto, M. (2023). Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma *Support Vector Machine* (SVM). *KLIK: Kajian Ilmiah Informatika Dan Komputer*, *4*(1), 407–417. https://doi.org/10.30865/klik.v4i1.1040

Romadhoni, Y., & Holle, K. F. H. (2022). Analisis Sentimen Terhadap PERMENDIKBUD No.30 pada Media Sosial Twitter Menggunakan Metode *Naive Bayes* dan LSTM. *Jurnal Informatika: Jurnal Pengembangan IT*, *7*(2), 118–124. https://doi.org/10.30591/jpit.v7i2.3191

Sasongko, M. F., Fauziah, F., & Hayati, N. (2023). Analisis Sentimen Opini Masyarakat Terhadap Kuliner DKI Jakarta dengan Metode Naïve Baiyes dan Support Vector Machine. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, *7*(3), 241. https://doi.org/10.30998/string.v7i3.13931

Soemedhy, C. A. A., Trivetisia, N., Winanti, N. A., Martiyaningsih, D. P., Utami, T. W., & Sudianto, S. (2022). Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube "Kekerasan Seksual"). *Jurnal Informatika: Jurnal Pengembangan IT*, *7*(2), 80–84. https://doi.org/10.30591/jpit.v7i2.3547

Suryono, S., Utami, E., & Luthfi, E. T. (2018). Analisis Sentiment Pada Twitter Dengan Menggunakan Metode *Naive Bayes* Classifer. *Proceedings SEMINAR NASIONAL GEOTIK 2018*, 9–15. http://hdl.handle.net/11617/9777

Yuniar, E., Utsalinah, D. S., & Wahyuningsih, D. (2022). Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompet Digital dengan Menggunakan Algoritma Machine Learning. *Jurnal Janitra Informatika Dan Sistem Informasi*, *2*(1), 35–42. https://doi.org/10.25008/janitra.v2i1.145