

ANALISIS SENTIMEN KOMENTAR CYBERBULLYING PADA TIKTOK MENGGUNAKAN ARSITEKTUR TEXT CNN

TUGAS AKHIR

Diajukan sebagai syarat menyelesaikan jenjang strata Satu (S-1) di
Program Studi Teknik Informatika, Fakultas Teknologi Industri,
Institut Teknologi Sumatera

Oleh:

Nikola Arinanda

121140202



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SUMATERA
LAMPUNG SELATAN
2025**

LEMBAR PENGESAHAN

Saya menyatakan bahwa Tugas Akhir berjudul “Analisis Sentimen Komentar Cyberbullying pada TikTok Menggunakan Arsitektur Text CNN” merupakan hasil karya saya sendiri dan belum pernah diajukan, baik sebagian maupun seluruhnya, di Institut Teknologi Sumatera atau institusi pendidikan lain oleh saya maupun pihak lain.

Lampung Selatan, 26-05-2025

Penulis,

Nikola Arinanda
NIM. 121140202

Diperiksa dan disetujui oleh,

Pembimbing

1. Martin Clinton Tosima Manullang, Ph.D.
NIP. 19930109 2019 03 1 017
2. Leslie Anggraini, S.Kom., M.Cs.
NRK. 19970817 2024 2294

Penguji

1. Andika Setiawan, S.Kom., M.Cs.
NIP. 19911127 2022 03 1 007
2. Eko Dwi Nugroho, S.Kom., M.Cs.
NIP. 19910209 2024 06 1 001

Disahkan oleh,

Koordinator Program Studi Teknik Informatika
Fakultas Teknologi Industri
Institut Teknologi Sumatera

Andika Setiawan, S.Kom., M.Cs.
NIP. 19911127 2022 03 1 007

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir dengan judul “Analisis Sentimen Komentar Cyberbullying pada TikTok Menggunakan Arsitektur Text CNN” adalah karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan benar.

Nama : Nikola Arinanda

NIM : 121140202

Tanda Tangan :

Tanggal :

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Institut Teknologi Sumatera, saya yang bertanda tangan di bawah ini:

Nama : Nikola Arinanda

NIM : 121140202

Program Studi : Teknik Informatika

Fakultas : Teknologi Industri

Jenis Karya : Tugas Akhir

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Institut Teknologi Sumatera **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**Analisis Sentimen Komentar Cyberbullying pada TikTok Menggunakan
Arsitektur Text CNN**

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Institut Teknologi Sumatera berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Lampung Selatan

Pada tanggal : 31 Juli 2025

Yang menyatakan

Nikola Arinanda

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa atas limpahan rahmat, karunia, serta petunjuk-Nya sehingga penyusunan tugas akhir ini telah terselesaikan dengan baik. Dalam penyusunan tugas akhir ini penulis telah banyak mendapatkan arahan, bantuan, serta dukungan dari berbagai pihak. Oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Kedua Orang Tua yang selalu memberikan dukungan dan doa sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Bapak Martin Clinton Tosima Manullang, Ph.D. selaku Dosen Pembimbing atas ide, waktu, tenaga, perhatian, dan masukan yang telah disumbangsihkan kepada penulis.
3. Ibu Leslie Anggraini, S.Kom., M.Cs. selaku Dosen Pembimbing atas ide, waktu, tenaga, perhatian, dan masukan yang telah disumbangsihkan kepada penulis.
4. Teman-teman penulis yang membantu selama masa perkuliahan dan bekerja sama dalam melakukan penelitian tugas akhir yang tidak bisa disebutkan satu persatu.

Akhir kata penulis berharap semoga tugas akhir ini dapat memberikan manfaat bagi kita semua. Penulis menyadari bahwa tugas akhir ini tidak luput dari kekurangan dan kelemahan, dan penulis terbuka untuk menerima saran, kritik, dan masukan.

RINGKASAN

Analisis Sentimen Komentar Cyberbullying pada TikTok Menggunakan

Arsitektur Text CNN

Nikola Arinanda

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

ABSTRAK

Analisis Sentimen Komentar Cyberbullying pada TikTok Menggunakan

Arsitektur Text CNN

Nikola Arinanda

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Kata Kunci: kunci1, kunci2

ABSTRACT

Comparison

Nikola Arinanda

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Keywords: keywords1, keywords2

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR RUMUS

DAFTAR KODE

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital yang pesat telah membawa perubahan signifikan dalam cara manusia berinteraksi. Namun, di balik kemudahan yang dihadirkan seperti media sosial, permainan video, dan internet, muncul pula tantangan serius, terutama terkait dampaknya pada kesehatan mental. Teknologi digital dapat membawa dampak negatif yang signifikan untuk kesehatan mental seperti kecanduan, gangguan tidur, depresi, dan kecemasan, terutama pada kelompok rentan seperti anak muda [1]. Fenomena ini menjadi perhatian global seiring dengan meningkatnya jumlah pengguna digital dari berbagai usia.

Salah satu ancaman terbesar terhadap kesehatan mental di dunia maya adalah *cyberbullying*. Ini merupakan tindakan perundungan yang terjadi di ranah digital, seringkali dalam bentuk komentar bernada negatif, hinaan, maupun pelecehan verbal [2]. Dampak *cyberbullying* sangat serius, tidak hanya menyebabkan tekanan psikologis seperti stres dan gangguan tidur, tetapi juga dapat meningkatkan risiko masalah kesehatan mental yang lebih parah, termasuk depresi, kecemasan, dan, dalam kasus ekstrem, pikiran untuk bunuh diri pada korbannya [3].

Kemajuan teknologi digital telah menghadirkan berbagai platform komunikasi yang memungkinkan interaksi antarindividu dari seluruh penjuru dunia. Platform-platform ini menjadi ruang utama bagi pengguna untuk berbagi informasi, mengekspresikan diri, serta menjalin relasi sosial secara luas [4]. Salah satu platform yang mengalami pertumbuhan sangat pesat dalam beberapa tahun terakhir adalah TikTok. Aplikasi ini digunakan oleh jutaan orang dari berbagai usia dan latar belakang, menjadikannya fenomena global

yang mengubah pola konsumsi serta produksi konten video pendek. Popularitas TikTok yang terus meningkat membuatnya menjadi salah satu aplikasi paling banyak diunduh di dunia pada tahun 2023 [5].

Maraknya komentar bernada negatif, hinaan, dan pelecehan verbal menjadi indikasi kuat adanya *cyberbullying* di TikTok. Tindakan semacam ini tidak hanya mencemari ruang digital, tetapi juga dapat berdampak buruk pada kondisi psikologis korban. Korban dapat merasa malu, takut, dan terintimidasi oleh komentar yang bersifat menyerang [6]. Karakteristik komentar yang singkat, tidak baku, serta informal membuat proses deteksi *cyberbullying* sangat sulit jika dilakukan secara manual, karena membutuhkan banyak sumber daya dan cenderung subjektif.

Mengatasi permasalahan ini, diperlukan solusi analisis sentimen yang mampu mengklasifikasikan komentar secara otomatis. Analisis sentimen adalah proses identifikasi opini atau emosi dalam teks, yang dapat dikategorikan menjadi sentimen positif, netral, atau negatif. Proses analisis yang optimal membutuhkan tahapan text processing yang sistematis, seperti pembersihan teks dari noise, tokenisasi, penghapusan stopword yang tidak relevan, serta konversi teks ke representasi numerik yang dapat diproses oleh sistem komputasi [7].

Pendekatan modern dalam analisis sentimen banyak mengandalkan teknik Deep Learning. Metode ini, yang merupakan bagian dari pembelajaran mesin dengan jaringan saraf tiruan berlapis-lapis, memiliki kemampuan luar biasa dalam mempelajari pola kompleks dari data mentah, termasuk data teks, tanpa perlu fitur yang dirancang secara manual [8]. Ini membuka jalan bagi sistem otomatis yang lebih canggih untuk memahami nuansa bahasa.

Di bawah payung deep learning, bidang natural language processing (NLP) memegang peranan krusial. NLP berfokus pada interaksi antara komputer dan bahasa manusia, memungkinkan mesin untuk memproses, menganalisis, memahami dan menghasilkan bahasa alami. Hal ini sangat esensial untuk

mengurai makna di balik komentar-komentar pengguna dan mengidentifikasi sentimen yang terkandung di dalamnya [7].

Beberapa penelitian terdahulu telah mencoba mengatasi masalah *cyberbullying* dengan menggunakan arsitektur canggih seperti *Bidirectional Encoder Representations from Transformers* (BERT) melalui proses fine-tuning pada dataset spesifik. Model-model berbasis transformer seperti BERT memang menunjukkan performa tinggi dalam banyak tugas *Natural Text Processing* (NLP). Namun, hasil dari studi-studi tersebut seringkali menunjukkan kecenderungan overfitting, di mana model terlalu spesifik terhadap data pelatihan sehingga performanya buruk pada data baru yang belum pernah dilihat sebelumnya, terutama dengan ukuran dataset yang terbatas atau distribusi data yang tidak seimbang [9].

Melihat tantangan tersebut, arsitektur TextCNN muncul sebagai alternatif yang menjanjikan. Penelitian terdahulu terkait TextCNN telah membuktikan keunggulannya dalam mengekstraksi fitur penting dari rangkaian kata dan telah banyak digunakan dalam berbagai tugas klasifikasi teks, termasuk analisis sentimen dan deteksi spam atau misinformasi. Keunggulan utamanya adalah kemampuannya dalam mempelajari representasi fitur secara otomatis melalui filter konvolusi, yang efektif menangkap pola lokal dan global dalam teks **kim2014convolutional**.

Berdasarkan uraian di atas, penelitian ini berfokus pada penerapan arsitektur TextCNN untuk membangun sistem klasifikasi sentimen komentar TikTok, khususnya yang mengandung indikasi *cyberbullying*. Sistem ini akan menggunakan pendekatan binary classification untuk membedakan antara komentar *cyberbullying* dan *non-cyberbullying*, dengan harapan dapat mencapai generalisasi yang baik. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi signifikan dalam mendukung sistem moderasi konten serta menciptakan lingkungan media sosial yang lebih sehat, aman, dan ramah bagi seluruh pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka permasalahan penelitian dirumuskan sebagai berikut:

1. Bagaimana mengklasifikasi komentar yang termasuk dan tidak termasuk *cyberbullying* pada platform TikTok dengan arsitektur TextCNN
2. bagaimana menguji performa model TextCNN dalam mengklasifikasikan komentar *cyberbullying* dan yang tidak.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, maka tujuan dari penelitian ini adalah:

1. Mengembangkan model klasifikasi untuk mendeteksi komentar yang termasuk dan tidak termasuk *cyberbullying* pada platform TikTok menggunakan arsitektur TextCNN.
2. Mengevaluasi performa model TextCNN dalam mengklasifikasikan komentar *cyberbullying* dan komentar non-*cyberbullying* berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.

1.4 Batasan Masalah

Adapun batasan masalah dari penelitian ini agar sesuai dengan yang diharapkan adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data komentar yang berasal dari platform TikTok.
2. Data komentar TikTok yang digunakan dalam penelitian ini diambil dari dataset yang telah dikumpulkan dan dipublikasikan oleh Bunga Aura Prameswari et al [9].
3. Klasifikasi dilakukan secara biner, yaitu dengan fokus pada pendeteksian komentar yang mengandung indikasi *cyberbullying* dan yang tidak.
4. Data yang dianalisis berjumlah 1.508 komentar dan seluruhnya

menggunakan bahasa Indonesia.

1.5 Manfaat Penelitian

Adapun manfaat yang diperoleh dari hasil penelitian ini adalah sebagai berikut:

1. Memberikan kontribusi dalam pengembangan sistem otomatis untuk mendeteksi komentar negatif dan *cyberbullying* di media sosial, khususnya pada platform TikTok.
2. Menjadi acuan atau referensi dalam penerapan arsitektur TextCNN untuk klasifikasi teks pendek, terutama pada komentar media sosial yang memiliki karakteristik bahasa informal dan ringkas.
3. Membantu mengurangi penyebaran komentar berunsur *cyberbullying* di platform digital melalui pengembangan sistem analisis teks berbasis pembelajaran mesin.

1.6 Sistematika Penulisan

Sistematika penulisan berisi pembahasan apa yang akan ditulis disetiap Bab. Sistematika pada umumnya berupa paragraf yang setiap paragraf mencerminkan bahasan setiap Bab.

Bab I

Bab ini membahas latar belakang yang melandasi penelitian, perumusan masalah yang ingin diselesaikan, serta tujuan yang ingin dicapai. Selain itu, dijelaskan juga batasan masalah, manfaat dari penelitian, dan sistematika penulisan sebagai panduan struktur laporan.

Bab II

Bab ini menguraikan teori-teori yang menjadi dasar penelitian, seperti konsep cyberbullying, analisis sentimen, dan text processing. Juga dijelaskan arsitektur model Text CNN yang digunakan dalam penelitian ini.

Bab III

Bab ini menjelaskan metode yang digunakan dalam penelitian, mulai dari teknik pengumpulan data hingga tahapan text processing. Selain itu, dibahas pula perancangan model Text CNN serta metode evaluasi performa model.

Bab IV

Bab ini menyajikan hasil pelatihan dan pengujian model yang telah dibangun. Dilengkapi dengan analisis performa dan interpretasi hasil klasifikasi sentimen untuk menilai keberhasilan model.

Bab V

Bab terakhir ini berisi kesimpulan dari hasil penelitian yang telah dilakukan. Penulis juga memberikan saran sebagai masukan untuk pengembangan penelitian di masa mendatang.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Penulis melakukan pencarian referensi terkait beberapa penelitian serupa yang pernah dilakukan sebagai dasar penelitian. Penelitian-penelitian yang menjadi referensi penulis dijabarkan pada Tabel ??.

1. Prameswari et al. (2023) melakukan penelitian dengan tujuan mengembangkan model prediksi untuk mendeteksi komentar bernuansa cyberbullying di platform TikTok. Data yang digunakan berjumlah 1.508 komentar TikTok yang dikumpulkan dan diberi label secara manual menjadi dua kelas, yaitu *cyberbullying* dan *non-cyberbullying*. Model yang digunakan dalam penelitian ini adalah arsitektur BERT (Bidirectional Encoder Representations from Transformers), sebuah model deep learning berbasis representasi kontekstual kata. Model BERT ini mampu menangkap makna dan konteks dalam komentar, dan hasil eksperimen menunjukkan bahwa model mampu mencapai akurasi validasi sebesar 63%. Meskipun demikian, terdapat indikasi *overfitting* dalam pelatihan yang membuat performa pada data validasi tidak meningkat secara signifikan. Penelitian ini menunjukkan pentingnya pengembangan model prediktif untuk mendeteksi perilaku cyberbullying secara otomatis, khususnya pada platform yang memiliki tingkat interaksi tinggi seperti TikTok. Temuan ini menjadi dasar bahwa analisis sentimen dapat dijadikan pendekatan awal dalam memahami dan mengklasifikasikan konten berbahaya di media sosial.
2. Yoon Kim (2014) melakukan penelitian untuk memperkenalkan arsitektur *TextCNN*, yaitu model konvolusional sederhana yang digunakan untuk klasifikasi teks pada tingkat kalimat. Model ini

bekerja dengan menerapkan beberapa filter konvolusi pada representasi vektor kata dari kalimat untuk mengekstrak fitur penting secara lokal. Selanjutnya, fitur terbaik dari masing-masing filter diambil melalui teknik *max-over-time pooling* dan diproses ke dalam lapisan *fully connected* untuk klasifikasi.

Penelitian ini menunjukkan bahwa meskipun arsitekturnya sederhana dan minim *hyperparameter tuning*, model TextCNN mampu memberikan hasil yang kompetitif pada berbagai *benchmark* NLP seperti analisis sentimen dan klasifikasi pertanyaan. Keunggulan utamanya terletak pada efisiensi dan kemampuannya dalam menangkap informasi semantik lokal tanpa memerlukan struktur kalimat yang kompleks seperti *parsing*. Temuan dari Kim ini menunjukkan bahwa TextCNN layak dijadikan alternatif model yang ringan namun tetap akurat untuk tugas klasifikasi komentar, seperti dalam konteks deteksi cyberbullying di media sosial seperti TikTok.

Tabel 2.1 Literasi Penelitian Terdahulu

No.	Peneliti (tahun)	Judul Penelitian [sitasi]	Permasalahan	Ekstraksi Fitur	Metode Klasifikasi	Hasil Penelitian
1.	Yoon Kim (2014)	Convolutional Neural Networks for Sentence Classification	Permasalahan: Penelitian ini berfokus pada tugas klasifikasi tingkat kalimat (sentence-level classification tasks), termasuk analisis sentimen dan klasifikasi pertanyaan.	Fitur diekstrak menggunakan vektor kata pra-terlatih (pre-trained word vectors) yang diperoleh dari model bahasa saraf tak terawasi (unsupervised neural language model). Vektor kata yang digunakan dilatih pada 100 miliar kata dari Google News.	Penelitian ini menggunakan Jaringan Saraf Tiruan Konvolusional (CNN) yang sederhana dengan satu lapisan konvolusi di atas vektor kata. Arsitektur model ini menggunakan lapisan konvolusi, operasi max-over-time pooling, dan lapisan fully connected softmax dengan dropout.	Model CNN yang sederhana dengan vektor statis menunjukkan hasil yang sangat baik pada beberapa tolak ukur, dan kinerjanya ditingkatkan lebih lanjut melalui fine-tuning. Secara keseluruhan, model ini mengungguli model
2.	Bunga Prameswari et al. (2023)	Building Prediction Model for Detecting Cyberbullying using TikTok Comments	Deteksi cyberbullying pada komentar TikTok.	Analisis konten tekstual dari komentar TikTok. Komentar diklasifikasikan berdasarkan kriteria seperti bahasa yang merendahkan, ancaman, atau body-shaming.	Model deep learning BERT (Bidirectional Encoder Representations from Transformers)	Akurasi validasi model mencapai 0.63 pada epoch kesembilan

2.2 Dasar Teori

2.2.1 Media Sosial

Media sosial telah menjadi fenomena global yang mengubah cara manusia berinteraksi dan berkomunikasi. Platform-platform seperti Facebook, Twitter, Instagram, dan TikTok menyediakan ruang virtual yang memungkinkan penggunanya untuk saling terhubung tanpa batasan geografis. Kehadiran fitur interaktif seperti komentar, like, dan berbagi konten mendorong terciptanya budaya partisipatif yang tinggi dalam masyarakat digital **sari2019literasi**.

Namun, seiring dengan pertumbuhan pengguna yang sangat pesat, muncul pula tantangan serius. Salah satunya adalah potensi penyalahgunaan media sosial sebagai sarana untuk menyebarkan ujaran kebencian, pelecehan, dan bahkan perundungan digital atau cyberbullying. Komentar-komentar yang awalnya ditujukan sebagai bentuk ekspresi sering kali berubah menjadi alat untuk menyerang dan menjatuhkan orang lain secara psikologis.

2.2.2 Cyberbullying

Cyberbullying merupakan bentuk perundungan yang terjadi di ranah digital, dengan pelaku biasanya menyembunyikan identitas di balik akun anonim. Tindakan ini dapat berupa hinaan, ancaman, pelecehan verbal, atau penyebaran informasi yang merugikan korban melalui berbagai media sosial dan platform digital **qolbya2023empati**. Berbeda dengan perundungan tradisional yang terjadi secara fisik, cyberbullying bersifat persisten, karena jejak digital yang ditinggalkan dapat bertahan dalam waktu lama dan menyebar lebih cepat.

Korban cyberbullying sering mengalami tekanan emosional dan psikologis yang berat, termasuk depresi, rasa takut berlebihan, isolasi sosial, bahkan dorongan untuk menyakiti diri sendiri atau bunuh diri **hafifah2024perubahan**. Oleh karena itu, identifikasi dan deteksi otomatis terhadap komentar bernuansa cyberbullying menjadi upaya penting dalam mendukung lingkungan digital yang sehat dan aman.

2.2.3 Analisis Sentimen

Analisis sentimen atau sentiment analysis adalah suatu teknik dalam pemrosesan bahasa alami yang bertujuan untuk mengetahui emosi atau opini dalam suatu teks. Dengan menganalisis komentar atau teks yang ditulis oleh pengguna, sistem dapat mengkategorikan apakah teks tersebut bersifat positif, negatif, atau netral **pambudi2021effect**.

Pada konteks penelitian ini, analisis sentimen digunakan untuk mengidentifikasi komentar negatif yang mengandung unsur cyberbullying. Untuk itu, digunakan pendekatan klasifikasi biner (binary classification), yaitu pengelompokan komentar menjadi dua kategori: komentar yang mengandung cyberbullying dan komentar yang tidak mengandungnya. Teknik klasifikasi ini bertujuan menghasilkan sistem yang dapat mengotomatisasi proses identifikasi komentar secara efisien, mengingat komentar di media sosial jumlahnya sangat besar dan terus bertambah setiap detik.

2.2.4 Natural Language Processing (NLP)

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP memungkinkan sistem untuk membaca, memahami, dan menganalisis teks atau ucapan dengan cara yang menyerupai cara kerja otak manusia **liao2018some**. Dalam konteks analisis komentar di media sosial, NLP memainkan peran krusial dalam memproses data mentah berupa teks informal agar dapat diolah secara sistematis oleh model klasifikasi.

Proses dalam NLP mencakup berbagai tahap seperti tokenisasi (memecah kalimat menjadi kata-kata), normalisasi (standarisasi teks), penghapusan kata-kata tidak penting (stopword removal), dan transformasi teks ke dalam bentuk numerik. Setiap tahap ini diperlukan untuk memastikan bahwa model pembelajaran mesin menerima data yang bersih dan relevan.

2.2.5 Text Pre-processing

Sebelum teks dianalisis oleh model pembelajaran mesin atau deep learning, diperlukan tahap pemrosesan awal atau text pre-processing. Tahapan ini bertujuan untuk mengubah komentar yang tidak terstruktur menjadi bentuk yang lebih teratur dan konsisten.

Salah satu tahapan utama dalam pre-processing adalah normalisasi, yaitu proses menyederhanakan teks agar sesuai dengan kaidah bahasa standar. Normalisasi mencakup pengubahan huruf kapital menjadi huruf kecil, penghapusan tanda baca atau karakter khusus, serta konversi kata tidak baku menjadi bentuk baku **wang2019convolutional**. Hal ini sangat penting mengingat komentar di media sosial seringkali menggunakan bahasa yang singkat, penuh dengan singkatan, dan tidak mengikuti struktur kalimat yang formal.

Selain normalisasi, teknik augmentasi teks juga sering digunakan, terutama ketika jumlah data yang tersedia terbatas. Augmentasi memungkinkan peningkatan variasi data pelatihan dengan menghasilkan data baru dari data yang sudah ada. Beberapa teknik augmentasi yang umum digunakan antara lain penggantian kata dengan sinonim (synonym replacement), pertukaran posisi kata, atau penerjemahan bolak-balik (back translation) untuk menciptakan versi lain dari kalimat asli **wei2019eda**. Augmentasi sangat membantu dalam mencegah overfitting pada model pembelajaran, terutama ketika dataset tidak seimbang atau jumlah data yang tersedia tidak mencukupi.

2.2.6 Word Embedding

Setelah teks diproses, langkah selanjutnya adalah mengubah kata-kata menjadi representasi numerik agar bisa dipahami oleh komputer. Proses ini dikenal sebagai word embedding. Word embedding mengubah kata menjadi vektor berdimensi tetap yang dapat menangkap makna semantik serta hubungan antar kata dalam suatu ruang vektor **mikolov2013distributed**.

Berbeda dengan metode lama seperti one-hot encoding yang menghasilkan vektor biner yang sangat jarang (sparse), embedding memungkinkan sistem mengenali bahwa kata “marah” memiliki makna yang lebih dekat dengan “kesal” dibandingkan dengan “makan”. Beberapa metode populer dalam word embedding adalah Word2Vec, GloVe, dan FastText.

2.2.7 Deep Learning Dalam Pemrosesan Teks

Deep learning adalah pendekatan modern dalam pembelajaran mesin yang menggunakan jaringan saraf tiruan dengan banyak lapisan untuk mempelajari representasi dari data. Dalam bidang pemrosesan bahasa alami, deep learning memberikan hasil yang sangat baik karena kemampuannya dalam mengenali pola kompleks dalam teks, bahkan tanpa perlu fitur yang dirancang secara manual **goodfellow2016deep**.

Model deep learning seperti Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), dan Transformer telah banyak digunakan dalam tugas-tugas NLP. Namun, untuk teks pendek seperti komentar media sosial, arsitektur yang lebih sederhana namun efektif seperti CNN untuk teks atau TextCNN sering kali menjadi pilihan yang lebih tepat.

2.2.8 TextCNN

TextCNN adalah varian dari Convolutional Neural Network (CNN) yang dikembangkan secara khusus untuk klasifikasi teks. Berbeda dengan CNN pada gambar, TextCNN bekerja dengan cara menerapkan filter konvolusi pada urutan kata (yang telah diubah menjadi vektor embedding) untuk menangkap pola lokal seperti n-gram yang mengandung informasi penting **kim2014convolutional**.

Arsitektur TextCNN umumnya terdiri atas lapisan embedding sebagai input, dilanjutkan dengan beberapa filter konvolusi berukuran berbed, kemudian diteruskan ke lapisan max-pooling untuk mengekstraksi fitur paling penting dari setiap filter. Hasil pooling kemudian digabungkan dan dilewatkan ke fully

connected layer untuk menghasilkan output berupa klasifikasi biner.

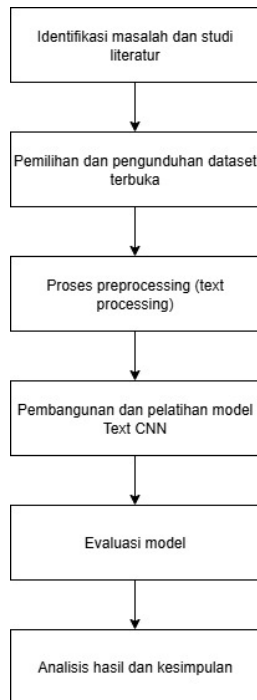
Keunggulan utama TextCNN terletak pada kesederhanaan dan efisiensinya. Ia tidak hanya mampu menangkap pola-pola lokal dalam teks secara efektif, tetapi juga lebih ringan dibandingkan model berbasis transformer seperti BERT, serta tidak memerlukan data pelatihan dalam jumlah besar. Hal ini menjadikan TextCNN sebagai pilihan ideal untuk klasifikasi komentar pendek di media sosial seperti TikTok.

BAB III

METODE PENELITIAN

3.1 Alur Penelitian

Penelitian ini dilakukan dengan pendekatan kuantitatif eksperimental yang berfokus pada penerapan metode deep learning, khususnya arsitektur Text CNN, untuk klasifikasi sentimen komentar cyberbullying. Secara umum, alur penelitian ini terdiri dari beberapa tahapan utama, yaitu:



Gambar 3.1 Alur Penelitian

3.2 Penjabaran Langkah Penelitian

Untuk memperjelas setiap langkah-langkah yang telah didefinisikan pada Gambar ??, berikut ini akan dijelaskan secara rinci tahapan-tahapan yang

dilakukan dalam penelitian ini.

3.2.1 Identifikasi Masalah

3.3 Alat dan Bahan Tugas Akhir

Dalam menjalani penelitian, beberapa alat dan bahan digunakan untuk memastikan penelitian berjalan dengan baik.

3.3.1 Alat

Dalam menganalisis sentimen komentar cyberbullying pada TikTok menggunakan arsitektur Text CNN, berikut adalah alat-alat yang digunakan:

3.3.1.1 Perangkat Lunak

1. *Visual Studio Code* sebagai *text editor*.
2. Torch versi 2.3.1
3. transformers versi 4.51.2
4. NumPy versi 1.26.2
5. Pandas versi 2.2.3
6. Scikit-learn versi 1.7.0
7. Matplotlib versi 3.9.0
8. Wandb versi 0.19.11
9. Tqdm versi 4.66.5
10. Muon-optimizer versi 0.1.0

3.3.1.2 Perangkat Keras

- Asus Vivobook 14x m1403qa
- Processor AMD Ryzen 5 5600H with Radeon Graphics
- RAM 24 GB
- AMD Radeon (TM) Graphics
- Sistem operasi Windows 11 Home 64-bit

3.3.2 Bahan

Dataset yang digunakan pada penelitian ini merupakan dataset komentar TikTok yang diperoleh dari repositori GitHub penelitian Prameswari et al [9]. Dataset ini berisi komentar-komentar TikTok yang telah dilabeli dengan kategori sentimen positif dan negatif. Dataset ini berisi 1.508 data komentar TikTok yang telah dilabeli. Dataset ini digunakan untuk melatih dan menguji model Text CNN dalam mengklasifikasikan sentimen komentar TikTok yang mengandung unsur cyberbullying.

3.4 Ilustrasi Perhitungan Metode

Metode yang digunakan dalam penelitian ini adalah Text CNN (Text Convolutional Neural Network), yaitu arsitektur CNN yang diadaptasi untuk klasifikasi teks. Berikut tahapan dan ilustrasi metode:

3.4.1 Embedding Layer

Komentar yang telah diproses diubah menjadi vektor representasi kata (menggunakan Word2Vec, GloVe, atau embedding trainable).

Misal: "kamu jelek" \rightarrow $[[0.2, -0.4], [0.5, 0.7]]$

3.4.2 Convolutional Layer

Filter konvolusi diterapkan untuk mengekstrak pola n-gram. Misal filter size: 2, 3, 4.

3.4.3 ReLU Activation Function

Fungsi aktivasi ReLU digunakan untuk menambah non-linearitas.

3.4.4 Max Pooling

Memilih nilai fitur tertinggi dari hasil konvolusi untuk mempertahankan fitur paling signifikan.

3.4.5 Fully Connected Layer

Output pooling digabung dan diteruskan ke dense layer.

3.4.6 Output Layer

Fungsi aktivasi softmax digunakan untuk mengklasifikasikan komentar ke dalam salah satu kelas: positif, netral, atau negatif.

DAFTAR PUSTAKA

- [1] Queene Br Sembiring. “Dampak Teknologi Digital Terhadap Kesehatan Mental Generasi Muda”. *Circle Archive* 1.4 (2024).
- [2] Yessi Mareta Andari Putri et al. “Cyberbullying di media sosial tiktok terhadap remaja sekolah menengah pertama”. *Jurnal common* 7.1 (2023), pp. 33–44.
- [3] Kurniawan Kurniawan et al. “Dampak Cyberbullying terhadap Risiko Bunuh Diri pada Remaja: Narrative Review”. *JPP (Jurnal Kesehatan Poltekkes Palembang)* 19.1 (2024), pp. 8–16.
- [4] Astari Clara Sari et al. “Komunikasi dan media sosial”. *Jurnal The Messenger* 3.2 (2018), p. 69.
- [5] Puput Silva Rosiana et al. “Analisis aplikasi tiktok berdasarkan prinsip dan paradigma interaksi manusia dan komputer menggunakan evaluasi heuristic”. *Jurnal Informatika dan Teknik Elektro Terapan* 11.3 (2023).
- [6] Fabiola Greselda Aser, Sinta Paramita, et al. “Fenomena Cyberbullying di Media Sosial TikTok”. *Kiwari* 1.3 (2022), pp. 449–453.
- [7] AR Razi. “Klasifikasi Artikel Berita Berbahasa Indonesia Menggunakan Convolutional Neural Network”. PhD thesis. Universitas Gadjah Mada, 2017.
- [8] Shiyang Liao et al. “CNN for situations understanding based on sentiment analysis of twitter data”. *Procedia computer science* 111 (2017), pp. 376–381.
- [9] Bunga Aura Prameswari et al. “Building Prediction Model for Detecting Cyberbullying using TikTok Comments”. *2023 IEEE 8th International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*. 2023, pp. 1–7.

LAMPIRAN