



FORM PENGAJUAN JUDUL



Nama : Adelweys Margaretha Sijabat

NIM : 211402103

Judul diajukan oleh* : ☐ Dosen
☒ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) :

1. Data Science and Intelligent System
2. Computer Graphics and Vision

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Ivan Jaya S.Si., M.Kom.
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Ulfi Andayani S.Kom., M.Kom

Paraf Calon Dosen Pembimbing I

Medan, 27 Februari 2025

Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

** Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT.)

NIP. 198908172019032023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

RINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	PENDETEKSIAN BAHASA KASAR DI MEDIA SOSIAL DENGAN MENGGUNAKAN INDOBERT
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Saat ini, internet telah menjadi bagian tak terpisahkan dari kehidupan sehari-hari. Banyak aktivitas masyarakat yang sangat bergantung pada internet. Penggunaannya membawa dampak positif maupun negatif. Internet dapat mempermudah pekerjaan manusia, namun di sisi lain, jika tidak dimanfaatkan secara bijak, dapat berdampak buruk pada kesehatan mental.</p> <p>Menurut data dari Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), jumlah pengguna internet di Indonesia pada tahun 2024 mencapai 221.563.479 jiwa dari total populasi sebanyak 278.696.200 jiwa pada tahun 2023. Dengan angka tersebut, Indonesia menempati posisi keempat sebagai negara dengan jumlah pengguna internet terbanyak di dunia.</p> <p>Studi UNICEF pada tahun 2024 menemukan bahwa 50,3% anak telah terpapar konten bermuatan seksual melalui media sosial, sementara 48% anak pernah mengalami perundungan oleh sesama anak. Badan Pusat Statistik (BPS) mencatat bahwa 88,99% anak usia 5 tahun ke atas di Indonesia mengakses internet untuk media sosial. Hasil survei APJII menunjukkan bahwa tingkat penetrasi internet di Indonesia mencapai 79,5% dari total populasi. Menteri Komunikasi dan Digital (Menkomdigi), Meutya Hafid, menyatakan bahwa sekitar 9,17% pengguna internet di Indonesia merupakan anak-anak berusia di bawah 12 tahun. Meskipun para orang tua berupaya untuk membatasi penggunaan internet bagi anak-anak mereka, data menunjukkan bahwa 22% anak-anak tetap mengabaikan aturan yang telah ditetapkan oleh orang tua saat mengakses dunia maya.</p> <p>Penggunaan internet di kalangan anak-anak Indonesia terus meningkat, membawa berbagai dampak negatif yang perlu mendapat perhatian serius. Survei yang dilakukan pada tahun 2021 mengungkapkan bahwa 19,3% anak Indonesia mengalami kecanduan internet selama masa pandemi. Sebagian besar waktu mereka dihabiskan untuk bermain gim daring dan mengakses media sosial.</p> <p>Indonesia dikenal sebagai negara dengan masyarakat yang ramah, sopan, dan santun, baik terhadap sesama warga negara maupun orang asing. Hal ini telah dibuktikan oleh berbagai lembaga melalui riset dan survei. Salah satu lembaga yang melakukan jajak pendapat adalah InterNation, yang melibatkan 20.259 responden dari berbagai negara. Survei tersebut menilai keramahan penduduk serta kenyamanan mereka terhadap budaya lokal, termasuk aspek bahasa dan kemudahan dalam menjalin pertemanan. Dari 64 negara yang disurvei, Indonesia berhasil masuk dalam daftar sepuluh negara paling ramah di dunia.</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Namun, perilaku dan komentar masyarakat Indonesia di media sosial sangat bertolak belakang dari budaya Indonesia yang terkenal akan keramahannya. Pada awal Maret 2021, Microsoft merilis laporan *Digital Civility Index* (DCI) yang menilai tingkat kesopanan digital pengguna internet di seluruh dunia saat berkomunikasi di dunia maya. Dalam penelitian ini, Indonesia menempati peringkat ke-29 dari 32 negara yang disurvei, menunjukkan bahwa tingkat kesopanan digital di Indonesia tergolong rendah. Bahkan, di kawasan Asia Tenggara, Indonesia menduduki peringkat terbawah. Indeks kesopanan digital warganet Indonesia mengalami penurunan sebesar delapan poin, mencapai angka 76, di mana semakin tinggi skor tersebut, semakin buruk tingkat kesopannya.

Unggahan Microsoft mengenai hasil survei tersebut di Instagram mendapat respon negatif dari warganet Indonesia. Mayoritas warganet Indonesia berkomentar dengan kata-kata kasar sebagai tanda tidak terima dengan hasil survei yang dinyatakan. Banyaknya komentar negatif di akun perusahaannya membuat Microsoft menonaktifkan komentar di akun Instagram-nya. Komentar negatif warganet Indonesia yang membanjiri akun microsoft seolah semakin membenarkan hasil survei yang telah dilakukan.

Penggunaan bahasa kasar di internet, termasuk ujaran kebencian dan pelecehan verbal, dapat berdampak negatif pada kesehatan mental pengguna. Bahasa kasar di internet dapat menyebabkan kecemasan, depresi, dan ketidaknyamanan sosial, terutama pada anak muda (Sarfraiz et al., 2022).

Pemerintah melalui Kementerian Komunikasi dan Informatika (Kemenkominfo), menciptakan sebuah program yang bertujuan untuk meningkatkan pengetahuan dan pemahaman masyarakat, terutama remaja, dalam menggunakan internet dengan bijak. Program ini dikenal dengan nama Internet Sehat dan Aman (INSAN). INSAN bertujuan untuk mengedukasi masyarakat mengenai penggunaan internet yang aman dan bertanggung jawab. Program ini mengajarkan etika berinternet yang baik dengan melibatkan berbagai elemen masyarakat.

Berbagai metode berbasis kecerdasan buatan telah dikembangkan untuk mendeteksi ujaran kebencian dan bahasa kasar di internet. Penelitian oleh Chatterjee (2021) mengenai deteksi bahasa kasar di media sosial. Dataset yang digunakan berasal dari Twitter dan Wikipedia, dengan teks yang telah diberi label oleh manusia. Pendekatan yang digunakan melibatkan Bag-of-Words (BoW) untuk vektorisasi teks dan Support Vector Machine (SVM) sebagai model klasifikasi utama dan mendapatkan akurasi pengujian sebesar 91.8%.

Penelitian lain oleh Kaur et al. (2024) mendeteksi dan mengklasifikasikan bahasa kasar dan ofensif dengan berbagai model. Hasilnya adalah Model Bidirectional LSTM dan Multi-Dense LSTM menunjukkan hasil terbaik dalam mendeteksi dan mengklasifikasikan bahasa kasar dengan akurasi mencapai 99.9% dalam beberapa kategori. Penelitian oleh Hussain et al., (2025) mendeteksi bahasa ofensif dalam Roman Urdu menggunakan kombinasi teknik Machine Learning (ML) dan Deep Learning (DL). Penelitian ini



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI**
UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

menyimpulkan bahwa Bi-LSTM dengan Word2Vec embeddings adalah model terbaik dengan F1-score 98.00%.

Selain itu, ada beberapa penelitian yang menggunakan IndoBERT. Penelitian oleh Sagama & Alamsyah (2023) bertujuan untuk mengevaluasi efektivitas model berbasis BERT dan RoBERTa dalam mendeteksi konten beracun (toxicity) dalam bahasa Indonesia di media sosial, khususnya twitter. Dalam penelitian ini IndoBERTweet mendapatkan hasil paling tinggi dengan F1-score 0.89. Penelitian lain oleh ilie et al. (2020) menunjukkan bahwa IndoBERT adalah model terbaik untuk pemrosesan bahasa Indonesia, dengan performa superior dibandingkan model multibahasa seperti XLM-R dan mBERT. IndoBERTLARGE memiliki performa terbaik dalam klasifikasi kalimat (F1-score 88.97%), menunjukkan bahwa model monolingual lebih unggul dibanding model multibahasa seperti XLM-R dan mBERT dalam memahami bahasa Indonesia.

Berdasarkan dari latar belakang dan penelitian terdahulu, maka penulis mengusulkan untuk melakukan penelitian dengan judul “Implementasi Indobert Dalam Pendeteksian Bahasa Kasar Indonesia”. Sistem akan mendeteksi bahasa kasar dalam bahasa Indonesia dan menyensor setiap huruf menjadi karakter “*” untuk meminimalisir dampak negatif dari bahasa kasar kepada pengguna, khususnya bagi anak-anak.

Penelitian terdahulu

No.	Judul	Keterangan
1	Profanity detection in social media text using a hybrid approach of NLP and machine learning (Chatterjee, 2021)	Penelitian ini mengenai deteksi bahasa kasar di media sosial. Dataset yang digunakan berasal dari Twitter dan Wikipedia, dengan teks yang telah diberi label oleh manusia. Pendekatan yang digunakan melibatkan Bag-of-Words (BoW) untuk vektorisasi teks dan Support Vector Machine (SVM) sebagai model klasifikasi utama dan mendapatkan akurasi pengujian sebesar 91.8%.
2	Deep learning-based approaches for abusive content detection and classification for multi-class online user-generated data(Kaur et al., 2024)	Penelitian ini menggunakan deep learning dengan berbagai model seperti Bidirectional LSTM, Multi-Dense LSTM, Gated Recurrent Unit (GRU), dan Long Short-Term Memory (LSTM) untuk mendeteksi dan mengklasifikasikan bahasa kasar dan ofensif. Model Bidirectional LSTM dan Multi-Dense LSTM menunjukkan hasil terbaik dalam mendeteksi dan mengklasifikasikan bahasa kasar dengan akurasi mencapai 99.9% dalam beberapa kategori.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	3	ORUD-Detect: A Comprehensive Approach to Offensive Language Detection in Roman Urdu Using Hybrid Machine Learning–Deep Learning Models with Embedding Techniques(Hussain et al., 2025)	Penelitian ini bertujuan untuk mendeteksi bahasa ofensif dalam Roman Urdu menggunakan kombinasi teknik Machine Learning (ML) dan Deep Learning (DL). Bi-LSTM dengan Word2Vec embeddings adalah model terbaik dengan F1-score 98.00%, menunjukkan bahwa model ini paling baik dalam menangkap konteks bahasa kasar dalam Roman Urdu.
	4	ORUD-Detect: A Comprehensive Approach to Offensive Language Detection in Roman Urdu Using Hybrid Machine Learning–Deep Learning Models with mbedding Techniques(Hussain et al., 2025)	Penelitian ini menguji berbagai dataset dari platform media sosial, mencakup ujaran kebencian eksplisit dan implisit menggunakan metode machine learning tradisional dan Deep Learning. Kesimpulan dari penelitian ini adalah metode Deep Learning mampu menangkap konteks ujaran kebencian lebih baik dibandingkan dengan model Machine Learning tradisional. Model Bi-LSTM dengan CNN memiliki performa terbaik, dengan akurasi 90.2% dan F1-score 89.9%
	5	Multi-Label Classification of Indonesian Online Toxicity using BERT and RoBERTa (Sagama & Alamsyah, 2023)	Penelitian ini bertujuan untuk mengevaluasi efektivitas model berbasis BERT dan RoBERTa dalam mendeteksi konten beracun (toxicity) dalam bahasa Indonesia di media sosial, terutama Twitter. IndoBERTweet memiliki performa terbaik dengan F1-score 0.89, karena model ini dilatih khusus untuk memahami bahasa gaul dan slang di Twitter.
	6	IndoNLU: Benchmark and Resources for Evaluating Indonesia Natural Language Understanding(Wilie et al., 2020)	Penelitian ini menunjukkan bahwa IndoBERT adalah model terbaik untuk pemrosesan bahasa Indonesia, dengan performa superior dibandingkan model multibahasa seperti XLM-R dan mBERT. IndoBERTLARGE memiliki performa terbaik dalam klasifikasi kalimat (F1-score 88.97%), menunjukkan bahwa model monolingual lebih unggul dibanding model multibahasa seperti XLM-R dan mBERT dalam memahami bahasa Indonesia.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

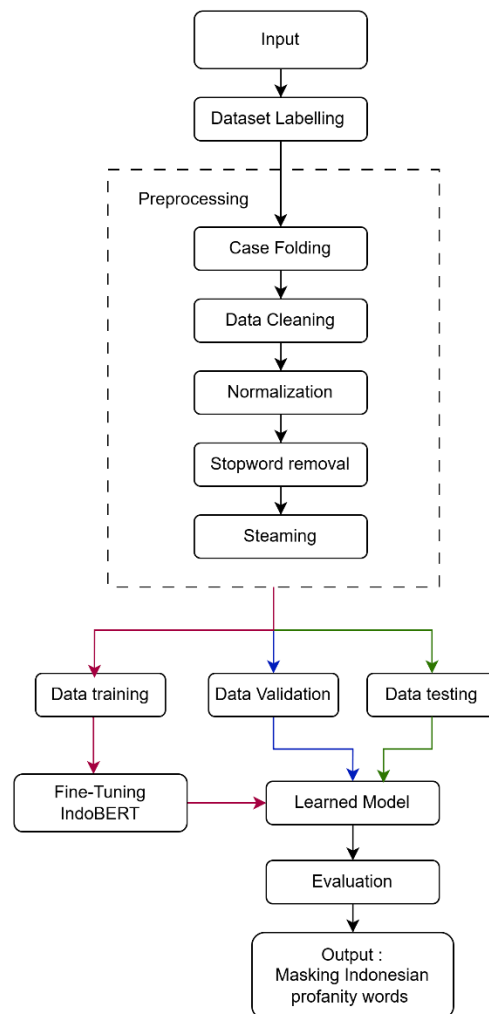
Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Rumusan Masalah

Dalam era digital, internet telah menjadi bagian penting dalam kehidupan sehari-hari, namun penggunaannya yang tidak bijak dapat berdampak negatif, terutama pada kesehatan mental. Indonesia merupakan salah satu negara dengan jumlah pengguna internet terbanyak di dunia, dengan penetrasi tinggi di kalangan anak-anak. Sayangnya, kesopanan digital warganet Indonesia tergolong rendah, terbukti dari laporan Digital Civility Index (DCI) 2021 oleh Microsoft, Indonesia menempati peringkat ke-29 dari 32 negara yang disurvei. Fenomena ini semakin diperkuat dengan banyaknya ujaran kebencian dan bahasa kasar yang beredar di media sosial, yang dapat menyebabkan kecemasan, depresi, dan ketidaknyamanan sosial. Pemerintah melalui **Internet Sehat dan Aman (INSAN)** telah berupaya meningkatkan literasi digital masyarakat, namun tantangan dalam menanggulangi bahasa kasar di internet masih menjadi permasalahan. Oleh karena itu, diperlukan teknologi yang dapat menyaring, mendeteksi, serta menyensor bahasa kasar tersebut kata-kata untuk meminimalisir dampak negatif yang ditimbulkan, khususnya bagi anak-anak.

Metodologi

Arsitektur umum





KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

1. Dataset labelling
Dataset labeling adalah proses memberikan label atau kategori pada data, sehingga bisa digunakan untuk melatih model. Label ini membantu model memahami pola dalam data dan membuat prediksi yang lebih akurat.
2. Pre-processing
Preprocessing adalah proses persiapan dan pembersihan teks sebelum digunakan dalam model Natural Language Processing (NLP). Tahapan ini penting agar data lebih terstruktur dan mudah dipahami oleh model machine learning atau deep learning.
3. Split data
Proses membagi dataset menjadi beberapa bagian untuk melatih, menguji, dan mengevaluasi model. Pembagian ini penting untuk memastikan model tidak hanya menghafal data tetapi juga dapat melakukan generalisasi dengan baik terhadap data baru.
4. Learned Model
Learned Model adalah model yang telah dilatih menggunakan data untuk mengenali pola dan membuat prediksi berdasarkan pengalaman yang diperoleh selama proses pelatihan. Sistem belajar dari data historis untuk membuat keputusan atau klasifikasi baru.
5. Evaluation
Proses mengukur kinerja model pemrosesan bahasa alami (Natural Language Processing) berdasarkan berbagai metrik evaluasi. Evaluasi ini bertujuan untuk menilai seberapa baik model memahami, mengolah, dan menghasilkan teks yang sesuai dengan tugas yang diberikan.
6. Output
Output merupakan hasil dari proses menggunakan *learned model*. Output yang dihasilkan berupa *masking* alfabet yang membentuk kata kasar.

Contoh dataset kata kasar :





Referensi

- Chatterjee, R. (2021). Profanity detection in social media text using a hybrid approach of NLP and machine learning. In *International Journal of Advance Research*.
- Hussain, N., Qasim, A., Mehak, G., Kolesnikova, O., Gelbukh, A., & Sidorov, G. (2025). ORUD-Detect: A Comprehensive Approach to Offensive Language Detection in Roman Urdu Using Hybrid Machine Learning-Deep Learning Models with Embedding Techniques. *Information*, 16(2), 139. <https://doi.org/10.3390/info16020139>
- Kaur, S., Singh, S., & Kaushal, S. (2024). Deep learning-based approaches for abusive content detection and classification for multi-class online user-generated data. *International Journal of Cognitive Computing in Engineering*, 5, 104–122. <https://doi.org/10.1016/j.ijcce.2024.02.002>
- Sagama, Y., & Alamsyah, A. (2023). Multi-Label Classification of Indonesian Online Toxicity using BERT and RoBERTa. *Proceedings of the 2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, IAICT 2023*, 143–149. <https://doi.org/10.1109/IAICT59002.2023.10205892>
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). *IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*. <http://arxiv.org/abs/2009.05387>

Medan, 27 Februari 2025
Mahasiswa yang mengajukan,

(Adelweys Margaretha Sijabat)

NIM. 211402103