
Non-Friendly Indonesian In-Game Chat Identifier

William Andrian Dharma T
STEI ITB

Bandung, Jawa Barat
13523006@std.stei.itb.ac.id

Aldoy Fauzan Avanza
STEI ITB

Bandung, Jawa Barat
18223113@std.stei.itb.ac.id

Christopher Brian
STEI ITB

Bandung, Jawa Barat
13522106@std.stei.itb.ac.id

Abstrak

Komunikasi dalam gim daring kerap disisipi percakapan yang tidak ramah seperti ujaran kasar, sarkasme agresif, dan ujaran kebencian, yang dapat mengganggu kenyamanan dan kesejahteraan pemain. Percobaan ini bertujuan mengembangkan sistem deteksi otomatis untuk mengidentifikasi percakapan *non-friendly* dalam gim berbasis Bahasa Indonesia. Dataset dikumpulkan dan dianotasi dari percakapan dalam video gim seperti Roblox dan Minecraft, lalu digunakan untuk melakukan *fine-tuning* model IndoBERT-base-uncased. Proses pelatihan dilakukan dengan pendekatan klasifikasi *single-label* menggunakan metrik evaluasi akurasi, *precision*, *recall*, dan F1-score. Hasil pengujian menunjukkan bahwa model berhasil mencapai akurasi sebesar 84% dan F1-score rata-rata makro sebesar 0.84, dengan performa terbaik pada kategori ujaran *racist*. Hasil ini menunjukkan bahwa pendekatan berbasis model transformer lokal seperti IndoBERT efektif dalam mengenali konten tidak ramah dalam percakapan digital. Percobaan ini diharapkan dapat berkontribusi pada pengembangan sistem moderasi otomatis yang adaptif terhadap konteks lokal dan *slang* pengguna dalam video gim daring.

1 Pendahuluan

Dalam beberapa tahun terakhir, video gim daring telah menjadi media komunikasi sosial yang semakin populer di Indonesia. Video gim seperti Roblox banyak digunakan oleh remaja dan dewasa muda untuk berinteraksi secara *real-time*. Namun, ruang komunikasi ini sering kali ditembus oleh percakapan yang tidak bersahabat, seperti kata-kata kasar, sarkasme agresif, dan ujaran kebencian, yang dapat merugikan kesejahteraan pemain serta dinamika komunitas yang sehat [5].

Non-friendly chat merujuk pada bentuk komunikasi digital yang bersifat menyerang, melecehkan, atau diskriminatif [2]. Fenomena ini dapat menyebabkan pemain merasa tidak nyaman, menurunkan tingkat keterlibatan, bahkan berkembang menjadi *cyberbullying* yang berdampak negatif terhadap kesehatan mental dan kesejahteraan pengguna [3][5].

Untuk mengatasi tantangan ini, dibutuhkan sistem otomatis yang mampu mengenali dan mengkategorikan percakapan *non-friendly* secara cepat dan akurat. Pendekatan Natural Language Processing (NLP), khususnya model transformer seperti BERT, telah terbukti efektif dalam mendeteksi ujaran toxic di berbagai domain, termasuk media sosial [1][4] serta percakapan dalam gim [6][7].

Model khusus Bahasa Indonesia seperti IndoBERT menunjukkan hasil yang sangat baik dalam berbagai tugas klasifikasi teks berbahasa lokal, seperti deteksi ujaran kebencian dan *cyberbully-*

ing [2][3]. Namun, penelitian yang secara khusus mengaplikasikan IndoBERT pada data chat game, yang memiliki karakteristik konteks real-time dan slang khusus pemain, masih sangat terbatas [3].

Dalam ranah gim, dua studi penting yaitu ToxBuster dan Fine-Tuning PreTrained Language Models to Detect In-Game Trash Talks telah menggunakan BERT untuk mengidentifikasi toxic chat pada game seperti Rainbow Six Siege, DOTA2, dan For Honor, dengan hasil F1-score berkisar antara 0.82 hingga 0.90. Penelitian ini menekankan pentingnya konteks riwayat percakapan dan metadata pemain untuk meningkatkan ketahanan deteksi [6][7].

Sementara itu, dataset percakapan dalam gim berbahasa Indonesia yang seimbang dan tersedia secara publik masih sangat terbatas, sehingga memerlukan upaya pengumpulan dan anotasi khusus. Studi “IndoBERT-based Indonesian Cyberbullying Detection” berhasil mencapai akurasi tinggi (96,7%) dalam klasifikasi tweet berbahasa Indonesia, dan dapat dijadikan acuan metodologis untuk domain chat game [3].

Tujuan laporan ini adalah mengembangkan sistem Non-Friendly Indonesian In-Game Chat Identifier dengan langkah-langkah berikut: Mengumpulkan dan menganotasi dataset chat Roblox berbahasa Indonesia untuk mendeteksi percakapan *non-friendly* (*toxic*, kasar, sarkastik, provokasi), melakukan fine-tuning IndoBERT pada dataset tersebut, memperhatikan konteks dialog *real-time* dan metadata sesi permainan, mengevaluasi performa model menggunakan metrik seperti *precision*, *recall*, dan F1-score, dan membandingkan hasil dengan baseline sederhana dan model multilingual lainnya untuk menilai keunggulan IndoBERT.

Diharapkan penelitian ini dapat menghasilkan solusi deteksi otomatis yang efektif dan *scalable*, serta berkontribusi pada pengembangan lingkungan bermain video gim daring yang lebih aman dan ramah pengguna.

2 Kajian Teori

2.1 Natural Language Processing

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan (Artificial Intelligence) yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP mencakup berbagai teknik untuk memungkinkan komputer memahami, menafsirkan, dan menghasilkan bahasa alami secara efektif. Aplikasi umum NLP meliputi klasifikasi teks, analisis sentimen, ekstraksi informasi, terjemahan mesin, dan deteksi ujaran kebencian [10].

Salah satu fondasi penting dalam NLP modern adalah *text mining*, yaitu proses untuk mengekstrak informasi yang bermanfaat dari kumpulan data teks dalam jumlah besar. *Text mining* memanfaatkan teknik NLP untuk mengidentifikasi pola, hubungan, dan makna dari *unstructured text data* [11]. Dalam konteks penelitian ini, *text mining* digunakan untuk menganalisis ribuan chat berbahasa Indonesia yang berasal dari percakapan dalam video gim daring guna mengidentifikasi konten tidak ramah.

Perkembangan besar dalam NLP terjadi setelah diperkenalkannya arsitektur transformer oleh Vaswani et al. pada tahun 2017 melalui makalah “Attention is All You Need” [8]. Transformer merupakan model jaringan saraf yang dirancang khusus untuk memproses data sekuensial secara paralel, dengan menggunakan mekanisme *self-attention*. Mekanisme ini memungkinkan model untuk menentukan kata mana yang relevan dalam memahami makna suatu kata dalam kalimat, terlepas dari posisi kata tersebut.

Model berbasis transformer (transformer-based models) seperti BERT (Bidirectional Encoder Representations from Transformers) dan RoBERTa (Robustly Optimized BERT Pretraining Approach) menunjukkan performa unggul di berbagai tugas NLP karena kemampuannya menangkap konteks secara dua arah (bidirectional) dan fleksibilitas dalam menangani konteks panjang serta struktur bahasa yang kompleks [1][9]. Kekuatan ini menjadikan model transformer sangat relevan dalam mendeteksi ujaran tidak ramah yang sering kali memerlukan pemahaman konteks menyeluruh dan tidak hanya berdasarkan kata kunci eksplisit semata.

Hal ini sangat penting dalam tugas deteksi ujaran tidak ramah, di mana ujaran yang tidak baik bisa tersembunyi dalam kombinasi kata atau konteks yang luas.

2.2 Model BERT dan Adaptasi Untuk Bahasa Indonesia

BERT dilatih melalui dua tugas utama: Masked Language Modeling (MLM) dan Next Sentence Prediction (NSP) [1].

Masked Language Modeling (MLM) adalah proses di mana sebagian kata dalam input teks disembunyikan (masked), dan model dilatih untuk memprediksi kata-kata tersebut berdasarkan konteks di sekitarnya. Sebagai contoh, dalam kalimat "Saya suka [MASK] pagi", model diharapkan dapat menebak bahwa kata yang hilang adalah "kopi" atau kata yang sesuai konteks lainnya. Strategi ini memungkinkan BERT untuk memahami konteks kata dari dua arah sekaligus (*bidirectional*).

Next Sentence Prediction (NSP) bertujuan agar model memahami hubungan antar kalimat. Dalam tugas ini, model diberi dua kalimat dan diminta memprediksi apakah kalimat kedua benar-benar mengikuti kalimat pertama dalam konteks aslinya. Misalnya, diberikan dua kalimat: "Dia membuka laptop." dan "Dia mulai mengerjakan tugas.", model harus menentukan apakah hubungan kedua kalimat itu valid atau tidak. NSP sangat berguna untuk tugas-tugas seperti pertanyaan-jawaban dan inferensi teks.

Untuk menangani keterbatasan model BERT standar dalam memahami Bahasa Indonesia, dikembangkanlah model IndoBERT oleh tim IndoNLU. IndoBERT merupakan versi monolingual dari BERT yang dilatih menggunakan lebih dari 4 miliar kata dari 12 korpus berbeda, termasuk berita, Wikipedia, jurnal, media sosial, dan forum daring berbahasa Indonesia [2].

Model ini dibangun di atas arsitektur BERT-base dengan 12 *encoder layers*, 12 *attention heads*, *hidden size* 768, dan sekitar 110 juta parameter. IndoBERT dilatih dengan menggunakan teknik Masked Language Modeling (MLM) dan Next Sentence Prediction (NSP) seperti BERT asli, namun dengan penyesuaian tokenisasi untuk Bahasa Indonesia menggunakan WordPiece Tokenizer.

IndoBERT juga mendukung berbagai aplikasi NLP seperti klasifikasi teks, ekstraksi entitas, dan analisis sentimen dalam konteks lokal. Keunggulan IndoBERT terletak pada kemampuannya menangkap nuansa gramatikal dan semantik khas Bahasa Indonesia, baik dalam bentuk formal maupun informal, yang tidak dimiliki oleh model multilingual.

Dalam benchmark IndoNLU, IndoBERT konsisten menunjukkan performa unggul dibandingkan model-model lain seperti mBERT dan XLM-R dalam tugas-tugas klasifikasi Bahasa Indonesia. Penelitian menunjukkan bahwa IndoBERT mencapai F1-score hingga 88.97% dalam klasifikasi ujaran *toxic* di media sosial [3], menandakan potensi besar model ini dalam memahami konteks Bahasa Indonesia.

2.3 Fine-Tuning BERT untuk Tugas NLP Spesifik

Model BERT yang telah melalui proses *pre-training* memerlukan proses *fine-tuning* agar dapat disesuaikan untuk tugas-tugas NLP tertentu, seperti klasifikasi *toxic comment*, ekstraksi entitas, atau analisis sentimen. Fine-tuning dilakukan dengan cara menambahkan lapisan keluaran (*output layer*) yang sesuai dengan tugas target, lalu melatih ulang seluruh model atau sebagian besar parameternya pada dataset berlabel yang lebih kecil namun spesifik.

Proses fine-tuning mencakup:

- Penyesuaian Arsitektur: Menambahkan *dense layer* dan fungsi aktivasi (biasanya sigmoid atau softmax) sesuai kebutuhan klasifikasi *binary*, *multi-class*, atau *multi-label*.
- Penentuan Hyperparameter: Seperti *learning rate*, *batch size*, dan jumlah epoch. Kombinasi ini sangat mempengaruhi performa akhir model.
- Tokenisasi dan Padding: Menggunakan tokenizer bawaan dari model yang diadopsi (misalnya BertTokenizer untuk IndoBERT) dengan *padding* dan *truncation* agar input sesuai panjang maksimal.

2.4 Klasifikasi Single-Label untuk Toxicity Detection

Dalam banyak studi terkait deteksi ujaran tidak ramah, pendekatan yang umum digunakan adalah *single-label classification*, di mana setiap entri data diasumsikan hanya memiliki satu label utama.

Meskipun dalam kenyataannya satu kalimat bisa saja mengandung lebih dari satu jenis toxic, pendekatan ini tetap relevan dan efisien dalam skenario di mana sistem hanya perlu mengidentifikasi jenis dominan dari toksisitas dalam satu waktu.

Klasifikasi *single-label* lebih sederhana secara arsitektural dibandingkan *multi-label*. Pada pendekatan ini, model hanya perlu memilih satu kelas dari sekumpulan label yang telah ditentukan sebelumnya, misalnya: *identity attack*, *insult*, *profanity*, *threat*, atau *sexually explicit*. Arsitektur model biasanya menggunakan softmax di *output layer* dan fungsi loss categorical cross-entropy untuk menghitung kesalahan prediksi.

Dalam studi oleh Fesalbon et al. [7], model dilatih dan dievaluasi menggunakan pendekatan single-label classification untuk mendeteksi bentuk percakapan toxic dalam data Twitter berbahasa Indonesia. Meskipun pelabelan awal dilakukan secara *multi-label*, data kemudian disederhanakan dengan mengambil label dominan. Hasilnya menunjukkan bahwa model IndoBERTweet tetap memberikan performa unggul, menegaskan efektivitas pendekatan ini dalam konteks praktis klasifikasi *toxic chat* Bahasa Indonesia.

2.5 Evaluasi Model

Evaluasi yang tepat sangat penting untuk menilai performa model yang telah di-fine-tune, khususnya dalam tugas klasifikasi teks. Metrik evaluasi umum yang digunakan adalah *precision*, *recall*, dan F1-score. Rumus dari masing-masing metrik adalah sebagai berikut [7]:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Untuk klasifikasi multi-label, metrik di atas dapat disesuaikan dengan dua metode agregasi yaitu *macro-averaging* dan *micro-averaging*.

Micro averaging menggabungkan jumlah keseluruhan dari TP, FP, dan FN dari semua label sebelum menghitung metrik. Ini memberikan bobot yang sama untuk setiap instance, cocok untuk dataset yang tidak seimbang.

Sebaliknya, *macro averaging* menghitung metrik untuk setiap label secara terpisah, lalu mengambil rata-ratanya. Ini memberikan bobot yang sama untuk semua label, cocok untuk mengevaluasi performa secara merata di seluruh label.

Rumus evaluasi mikro dan makro adalah sebagai berikut:

$$B_{\text{micro}} = B \left(\sum_{\lambda=1}^q TP_{\lambda}, \sum_{\lambda=1}^q FP_{\lambda}, \sum_{\lambda=1}^q TN_{\lambda}, \sum_{\lambda=1}^q FN_{\lambda} \right) \quad (4)$$

$$B_{\text{macro}} = \frac{1}{q} \sum_{\lambda=1}^q B(TP_{\lambda}, FP_{\lambda}, TN_{\lambda}, FN_{\lambda}) \quad (5)$$

di mana λ adalah label ke- λ dan q adalah jumlah total label.

Dalam studi oleh Fesalbon et al [7], pendekatan macro averaging digunakan karena distribusi label yang cukup seimbang.

3 Solusi Usulan

Kami mengusulkan sebuah *pipeline* berbasis *supervised learning* menggunakan *language model* berbasis *pre-trained transformer*. Kami mengutamakan pengembangan sebuah model kontekstual yang mampu menangani pola-pola bahasa informal dan frasa multibahasa yang umum ditemui pada *in-game chat*.

Pendekatan kami dapat dibagi menjadi beberapa tahap. Pertama, preparasi *dataset* dan *label encoding*. *Dataset* yang kami gunakan berisi pesan-pesan *in-game chat* yang dianotasi dengan kategori yang paling sesuai. Untuk mempermudah proses *training*, kami menghilangkan fitur-fitur yang tidak relevan dan menggunakan *LabelEncoder* untuk mengkonversi label kategorikal menjadi kelas berbentuk numerik.

Selanjutnya, kami melakukan *preprocessing* dan *tokenization*. Mengingat karakteristik *in-game chat* yang didominasi oleh penggunaan bahasa informal, kami bereksperimen dengan melakukan *preprocessing* minimal untuk mempertahankan konten semantik dari tiap pesan, dan baru melakukan normalisasi pada *tokenizer*. *Tokenizer* IndoBERT yang kami gunakan memungkinkan tokenisasi pada tingkat suku kata, cocok untuk penggunaan *slang* dan variasi tipografi yang umum ditemui pada *in-game chat*.

Selanjutnya, kami menggunakan model indolem/indobert-base-uncased untuk memproses *dataset*. Model ini dilatih khusus pada teks berbahasa Indonesia dan sudah melalui proses *fine-tuning* untuk *sequence-classification*. Penggunaan model disesuaikan untuk mengakomodasi jumlah kategori unik dalam *dataset* yang kami gunakan.

Terakhir, *dataset* dibagi menjadi subset untuk *training* dan *testing* dengan rasio 80:20. Kami mengkonversi data menjadi format Dataset dari HuggingFace untuk memungkinkan *batch tokenization* yang efisien pada proses *training*. Kami melakukan *fine-tuning* pada model dan menilai performa model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Pipeline ini memungkinkan kami untuk mengembangkan model klasifikasi yang *contextually-aware* dan *language-specific* yang mampu menggeneralisasikan teks pesan baru dengan *preprocessing* minimal. Dengan memanfaatkan model IndoBERT, kami berharap dapat menghasilkan model yang mampu mempelajari pola-pola linguistik kompleks yang ditemui pada domain gim dan tidak memerlukan pengaturan manual berlebih.

4 Hasil Eksperimen dan Pengujian

Proses *training* kami lakukan pada 1 unit GPU NVIDIA 4060 dengan RAM 20 GB. Parameter yang kami gunakan adalah sebagai berikut.

Parameter	Nilai/Keterangan
<i>Pre-trained model</i>	indolem/indobert-base-uncased
<i>Learning rate</i>	2×10^{-5}
<i>Epochs</i>	3
<i>Batch size (train/eval)</i>	16 / 16
<i>Optimizer</i>	AdamW
Strategi evaluasi	Akhir setiap <i>epoch</i>
<i>Early-stopping / checkpointing</i>	Tidak

Tabel 1: Konfigurasi Proses *Training*

Setelah dilakukan proses *fine-tuning*, kami mengevaluasi hasil pemrosesan *test set* oleh model.

5 Analisis Hasil

Performa model yang kami dapat menunjukkan hasil yang cukup signifikan. Model dapat mengidentifikasi sebagian besar label dengan tepat pada *tepat* pada semua kelas, dengan akurasi keseluruhan sebesar 0,84. Selain itu, performa model juga cukup konsisten di semua kelas.

Label	Precision	Recall	F1-score	Support
0(<i>Harrasment</i>)	0.87	0.91	0.89	47 ⁴
1(<i>Neutral</i>)	0.81	0.75	0.78	547
2(<i>Racist</i>)	0.85	0.90	0.88	495
3(<i>Violence</i>)	0.81	0.81	0.81	628
Accuracy			0.84	2141
Macro avg	0.84	0.84	0.84	2141
Weighted avg	0.83	0.84	0.83	2141

Tabel 2: Evaluasi Performa Model

Performa yang kami dapat juga menunjukkan bahwa parameter yang kami gunakan sudah cukup tepat dalam membantu menangani ketidakseimbangan antarlabeled pada *dataset*. Meskipun persebaran label pada *dataset* tidak merata, model masih dapat meraih performa yang terbilang baik bahkan untuk label dengan jumlah data paling sedikit. Proses *pretraining* model pada korpus yang berukuran besar dan spesifik untuk bahasa Indonesia memberikan pemahaman yang menyeluruh, bahkan untuk pola dan ekspresi linguistik yang *under-represented*. Selain itu, penggunaan metrik evaluasi *macro avg* juga memastikan kelas minoritas tetap mendapat pembobotan yang sepadan saat proses evaluasi.

6 Kesimpulan dan Saran

Berdasarkan hasil eksperimen yang telah dilakukan, model IndoBERT-base-uncased yang di-*fine-tune* untuk tugas klasifikasi percakapan *non-friendly* dalam video gim daring berbahasa Indonesia mampu mencapai performa yang baik. Dengan konfigurasi sederhana dan *preprocessing* minimal, model memperoleh **skor akurasi sebesar 84%**, serta **skor F1** secara *macro average* dan *weighted average* masing-masing sebesar **0.84 dan 0.83**. Hasil ini menunjukkan bahwa pendekatan *context-aware* berbasis transformer **sangat efektif** dalam menangkap nuansa bahasa informal, sarkasme, serta bentuk toksisitas yang sering muncul dalam percakapan pemain gim daring.

Model menunjukkan performa yang stabil di berbagai label kategori toksik seperti *harassment*, *racist*, dan *violence*, dengan F1-score tertinggi tercapai pada label *racist* (0.88). Hal ini menunjukkan potensi besar penggunaan IndoBERT dalam skenario nyata moderasi percakapan digital berbahasa Indonesia.

Berikut adalah **saran** untuk pengembangan lebih lanjut:

- **Pengayaan Dataset:** Perlu dilakukan perluasan dataset dengan variasi jenis gim dan platform untuk meningkatkan kemampuan generalisasi model terhadap konteks percakapan yang lebih luas.
- **Peningkatan Preprocessing:** Meskipun *preprocessing* minimal cukup efektif, eksplorasi teknik normalisasi *slang*, deteksi *typo*, dan *context window* yang lebih panjang dapat meningkatkan pemahaman konteks pesan.
- **Evaluasi Konteks Multisentensial:** Penambahan informasi kontekstual seperti riwayat percakapan sebelumnya atau metadata sesi dapat membantu model dalam mendeteksi ujaran ambigu yang tergantung konteks.
- **Penggunaan Model Lightweight:** Untuk aplikasi real-time, perlu dipertimbangkan versi model lightweight seperti DistilBERT atau ALBERT agar deteksi dapat dilakukan dengan latensi rendah di perangkat terbatas.
- **Integrasi Sistem Moderasi:** Implementasi hasil penelitian ini dapat diarahkan ke dalam sistem moderasi otomatis pada gim daring lokal dengan dukungan *dashboard* visualisasi untuk moderator manusia.

7 Referensi

- [1] J. Devlin, M.-W. Chang, K. Lee, dan K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. of NAACL-HLT, Stroudsburg, PA, USA: ACL, 2019, hlm. 4171–4186. doi:10.18653/v1/N19-1423.
- [2] B. Wilie, T. Vincentio, F. Winata, dkk., "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in Proc. of the 1st Asia-Pacific ACL IJCNLP, Suzhou, China, Des. 2020, hlm. 843–857.
- [3] G. Z. Nabillah, S. Y. Prasetyo, Z. N. Izdihar, dan A. S. Girsang, "BERT base model for toxic comment analysis on Indonesian social media," *Procedia Computer Science*, vol. 216, hlm. 714–721, 2023. doi:10.1016/j.procs.2022.12.188.
- [4] K. M. Hana, Adiwijaya, S. Al Faraby, dan A. Bramantoro, "Multi-label Classification of Indonesian Hate Speech on Twitter Using Support Vector Machines," in 2020 Int. Conf. on Data Science and Its Applications (ICoDSA), IEEE, Agt. 2020, hlm. 1–7. doi:10.1109/ICoDSA50139.2020.9212992.
- [5] F. Hauser, J. Hautz, K. Hutter, dan J. Füller, "Firestorms: Modeling conflict diffusion and management strategies in online communities," *The Journal of Strategic Information Systems*, vol. 26, no. 4, hlm. 285–321, Des. 2017. doi:10.1016/j.jsis.2017.01.002.
- [6] Y. Yang, K. Kim, H. Nam, dan J. Park, "ToxBuster: In-game Chat Toxicity Buster with BERT," in Findings of EMNLP 2023, Singapura, Des. 2023, hlm. 1267–1279.
- [7] S. Fesalbon, D. Serafi, R. H. Rachmadi, dan M. F. Pasha, "Fine-Tuning Pre-trained Language Models to Detect In-Game Trash Talks," arXiv preprint, arXiv:2403.15458, Mar. 2024.
- [8] A. Vaswani et al., "Attention is All You Need," in Proc. of NeurIPS, 2017.
- [9] Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv:1907.11692, 2019.
- [10] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, dan M. R. Yeganegi, "Text Mining in Big Data Analytics," *Big Data and Cognitive Computing*, vol. 4, no. 1, p. 1, Jan. 2020. doi:10.3390/bdcc4010001.
- [11] M. O. Topal, A. Bas, dan I. van Heerden, "Exploring Transformers in Natural Language Generation: GPT, BERT, and XLNet," Feb. 2021.