

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian terdahulu memuat pembahasan tentang penelitian serupa atau mirip yang telah dilakukan terdahulu mengenai penggunaan chatbot maupun kasus cyberbullying dengan tujuan untuk membedakan dengan penelitian terdahulu ataupun menyempurnakan penelitian terdahulu sehingga dapat ditemukan keterbaharuan penelitian.

No	Judul Jurnal	Penulis	Ruang Lingkup Penelitian	Tujuan
1	Optimasi Chatbot Dengan Pemanfaatan Natural Language Processing	Salamun, Aidi Aprialdo, Sukri	Pengembangan chatbot yang menggabungkan Teknik NLP terkini, termasuk pemrosesan Bahasa alami, memberi analisis sentimen, dan pemahaman konteks	Aplikasi Untuk mempermudah dan mendapatkan informasi.
2	Chatbot menggunakan Natural Language Processing untuk pembelajaran bahasa Inggris berbasis android	Yuthsi aprilinda, Tiffany Martavia, Erlangga, Freddy Nur Afandi, Usman Rizal	Aplikasi chatbot berbasis NLP tentang pembelajaran Bahasa Inggris	Untuk Mempermudah pembelajaran Bahasa Inggris
3	Deteksi Cyberbullying dengan Uiva	Yudi Setiawan, Nur	Deteksi menggunakan metode klasifikasi seperti Machine	Pelabelan cyberbullying ing atau

No	Judul Jurnal	Penulis	Ruang Lingkup Penelitian	Tujuan
	mesin pembelajar klasifikasi Peluang dan Tantangan	Maulidevi, Kridantan	Learning, Sentiment analysis, Fuzzy logic, maupun metode deteksi lainnya	non cyberbullying maupun klasifikasi ke dalam jenis jenis cyberbullying
4	Implementasi Chatbot Menggunakan Dialogflow w dan Messenger Muntuk Layanan Customer Service pada E-commerce	Chaulina Alfianti Oktavia	Teknologi chatbot dirancang dengan menggunakan dialogflow yang terdiri dari agent, intent dan training phrase, selanjutnya dilakukan uji coba dengan menggunakan messenger	Teknologi chatbot yang mampu menggantikan layanan customer service dalam memberikan informasi secara responsif
5	Rancang bangun fitur chatbot customer service menggunakan dialogflow w	Jansen wiratama, Samuel ady Sanjaya, victor ilyas sugara	Aplikasi solo yang dapat membantu mengoptimalkan layanan customer service berupa chatbot sehingga dapat meminimalisir pertanyaan pelanggan yang tidak terjawab	Membuat fitur chatbot yang mampu mengatasi permasalahan yang dihadapi oleh

No	Judul Jurnal	Penulis	Ruang Lingkup Penelitian	Tujuan
			customer service untuk membantu menjawab pertanyaan dari pelanggan	

Tabel 2.1 Penelitian Terdahulu

Research Gap:

1. Optimasi Chatbot Dengan Pemanfaatan Natural Language Processing (2024) Salamun, Aldi Aprialdo, & Sukri, 2024, melakukan penelitian yang berjudul "Optimasi Chatbot Dengan Pemanfaatan Natural Language Processing" yang menggunakan flutter untuk web dan android java native untuk mobile, dan firebase sebagai database penyimpanan. Kesimpulan yang dihasilkan dari penelitian tersebut adalah dihasilkannya aplikasi chatbot di Universitas Abdurrahman Wahid yang memberikan informasi yang lebih mudah dan cepat bagi mahasiswa. Pengguna dapat dengan mudah mengajukan pertanyaan atau permintaan, dan chatbot secara akurat memberikan respons yang relevan dan bermanfaat. Kekurangan pada penelitian ini adalah hanya dibuat untuk sebuah universitas saja, dan penggunaan NLP mendasar. Chatbot hanya bisa percakapan antara chatbot dengan satu user secara bersamaan.
2. Chatbot menggunakan Natural Language Processing untuk pembelajaran Bahasa Inggris berbasis Android (2020) Yuthsi Aprilinda, Tifanny Martavia, Erlangga, Freddy Nur Afandi, Usman Riza ,2020, melakukan penelitian yang berjudul "Chatbot menggunakan Natural Language Processing untuk pembelajaran Bahasa Inggris, berbasis Android" yang bertujuan untuk membuat sarana pembelajaran bahasa inggris dengan menggunakan Chatbot. Kesimpulan yang dihasilkan dari

penelitian tersebut adalah dihasilkannya aplikasi Chatbot yang mampu menjawab pertanyaan mengenai 16 tenses dalam bahasa inggris. Kekurangan pada penelitian ini adalah Chatbot hanya bisa menjawab pertanyaan tertentu yang sesuai dengan keyword yang sudah disediakan oleh developernya. Chatbot hanya bisa percakapan antara chatbot dengan satu user secara bersamaan, tidak bisa dalam grup.

3. Deteksi CyberBullying dengan mesin pembelajaran klasifikasi (Supervised Learning): Peluang dan Tantangan (2022) Yudi Setiawan, Nur Ulva Maulidevi, Kridanto Surendro, 2022, melakukan penelitian yang berjudul "Deteksi Cyberbullying dengan mesin pembelajaran klasifikasi (Supervised Learning): Peluang dan Tantangan" yang bertujuan untuk pelabelan cyberbullying atau non-cyberbullying, dan mengklasifikasikannya ke dalam jenis jenis cyberbullying. Kekurangan dari penelitian ini adalah hanya dibuat untuk mengumpulkan data apa dan bagaimana cyerbbullying terjadi, tidak adanya cara pencegahan dan penerapannya menggunakan aplikasi.
4. Implementasi Chatbot Menggunakan Dialogflow dan Messenger Untuk Layanan Customer Services Pada E-commerce (2019) Chaulina Alfiani Oktavianti, 2019, melakukan penelitian yang berjudul "Implementasi Chatbot Menggunakan Dialogflow dan Messenger untuk Layanan Customer Services pada E-commerce" ini bertujuan untuk mengimplementasikan chatbot yang mampu menggantikan layanan customer service dalam memberikan informasi secara responsif. Kekurangan dari penelitian ini yakni kurangnya data pada tranning phase, seharusnya data bisa diambil dari platform yang menyediakan berbagai model training artificial intelligence (AI) contohnya Kata.id atau Huggingface.
5. Rancang Bangun Fitur Chatbot Customer Service Menggunakan Dialogflow (2022) Jansen Wiratama, samuel Ady Sanjaya, Victor Ilyas Sugara, 2022, melakukan penelitian yang berjudul "Rancang Bangun Fitur Chatbot Customer Service Menggunakan Dialogflow" yang bertujuan untuk membuat fitur Chatbot yang mampu mengatasi permasalahan yang dihadapi oleh customer service untuk membantu menjawab pertanyaan dari pelanggan.Kekurangan dari penelitian ini adalah data yang disediakan hanya 6, tidak ada penanganan apabila

yang ditanyakan oleh user berbeda dengan data yang ada di dalam dialogflow.

2.2 Teori Terkait

Bagian ini menjelaskan konsep dan teori yang relevan dengan pengembangan chatbot berbasis Natural Language Processing (NLP) untuk mendeteksi dan menangani cyberbullying di platform grup Discord. Beberapa teori dan konsep penting meliputi chatbot, NLP, cyberbullying, serta metode deteksi berbasis machine learning.

2.2.1 Cyberbullying

Cyberbullying merupakan tindakan agresif atau mempermalukan seseorang secara terus-menerus dalam jangka waktu tertentu dengan memanfaatkan teknologi digital. Aspek keberulangan dan keberlangsungan waktu menjadi faktor penting dalam membedakan cyberbullying dengan tindakan negatif yang hanya terjadi sekali dan tidak terstruktur (Pratiwi, 2022).

2.2.2 Chatbot

Chatbot merupakan program komputer yang dirancang untuk mensimulasikan kecerdasan buatan sehingga dapat berinteraksi dan melakukan percakapan dengan manusia (M. I. Rifaldi, 2023).



Gambar 2.2.2 Chatbot

Chatbot merupakan bagian dari bidang pengolahan bahasa alami, pembelajaran mesin, rekayasa perangkat lunak, dan kecerdasan buatan.

Chatbot dianggap menggunakan kecerdasan buatan karena mampu memahami masukan bahasa alami dengan susunan kata yang berbeda namun memiliki makna serupa dengan masukan sebelumnya, serta dapat memberikan respons yang relevan. Pengembangan chatbot dapat disesuaikan dengan kebutuhan, tetapi setiap chatbot yang dibuat memiliki tipe tertentu dalam memproses masukan dari pengguna.

1. Keyword Recognition-Based Chatbot

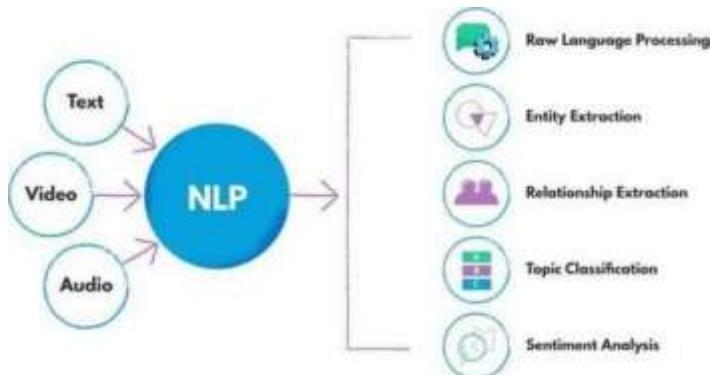
Keyword recognition-based chatbot merupakan jenis chatbot yang merespons masukan pengguna dengan cara mendeteksi kata kunci yang terdapat dalam input yang diberikan. Input yang diterima umumnya berupa susunan kata atau kalimat. Chatbot jenis ini menentukan respons yang tepat dengan memeriksa setiap kata dalam masukan hingga menemukan kata kunci yang sesuai dengan knowledge base yang dimiliki sistem. Namun, chatbot tipe ini memiliki kelemahan ketika harus merespons masukan yang mengandung kata kunci yang sama, sehingga memungkinkan terjadinya kesalahan dalam memberikan balasan yang sesuai.

2. Contextual Chatbot

Contextual chatbot merupakan jenis chatbot yang memanfaatkan pembelajaran mesin dan kecerdasan buatan dalam mengelola percakapan antara sistem dan pengguna. Sama seperti chatbot berbasis pengenalan kata kunci, contextual chatbot juga menerima masukan dalam bentuk kalimat atau susunan kata. Perbedaannya terletak pada kemampuannya dalam mengingat dan mempelajari setiap masukan yang diterima, sehingga apabila terdapat masukan dengan susunan kata yang berbeda namun memiliki makna yang sama, sistem tetap dapat memberikan respons yang relevan. Contextual chatbot memerlukan pengembangan knowledge base secara berkala agar sistem dapat mempelajari setiap input yang diberikan pengguna dengan baik.

2.2.3 Natural Language Processing

Natural Language Processing adalah sebuah teknologi machine learning yang memberi komputer kemampuan untuk menginterpretasikan, memanipulasi, dan memahami bahasa manusia.

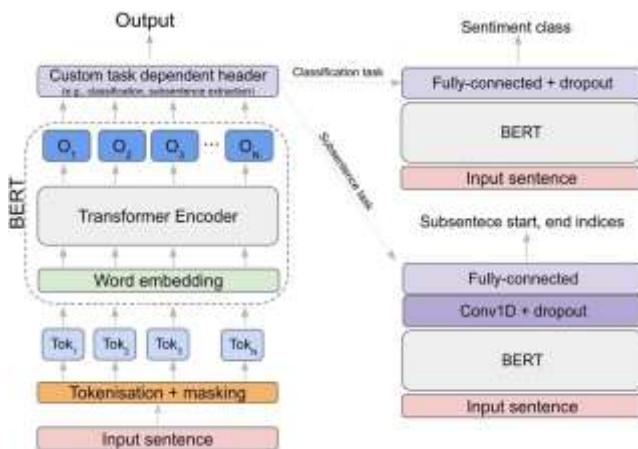


Gambar 2.2.3 Natural Language Processing

Perangkat lunak NLP digunakan untuk memproses data suara dan teks secara otomatis, menganalisis maksud atau sentimen dalam pesan, dan merespons komunikasi manusia dalam waktu nyata (AWS, 2024).

2.2.4 BERT (Bidirectional Encoder Representations from Transformers)

BERT (Bidirectional Encoder Representations from Transformers) adalah model bahasa mutakhir yang dikembangkan oleh Google. Berbeda dengan model sebelumnya, BERT memproses teks secara dua arah—dari kiri ke kanan dan sebaliknya—menggunakan teknik Masked Language Modeling (MLM). Metode ini memungkinkan BERT untuk memahami konteks kata lebih baik dengan memperhatikan kata-kata di sekitarnya secara penuh.



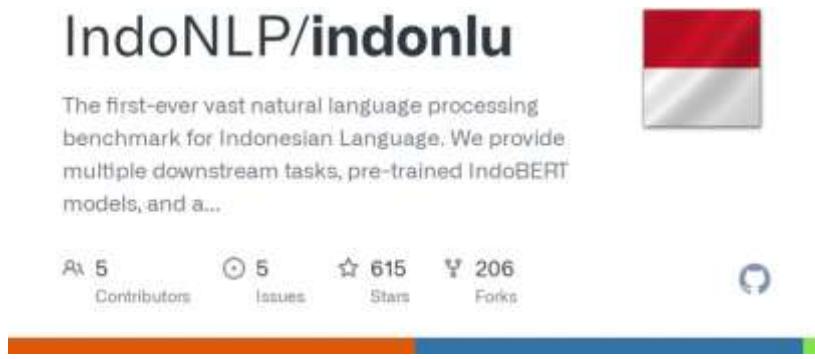
Gambar 2.2.4 BERT

BERT (Bidirectional Encoder Representations from Transformers) adalah model bahasa mutakhir yang dikembangkan oleh Google. Berbeda dengan model sebelumnya, BERT memproses teks secara dua arah—dari kiri ke kanan dan sebaliknya—menggunakan teknik Masked Language Modeling (MLM). Metode ini memungkinkan BERT untuk memahami konteks kata lebih baik dengan memperhatikan kata-kata di sekitarnya secara penuh.

BERT dilatih pada dataset berukuran besar dan kemudian disesuaikan (finetuning) untuk tugas tertentu seperti analisis sentimen atau penjawab pertanyaan. Selain itu, BERT menggunakan teknik Next Sentence Prediction (NSP) untuk meningkatkan kemampuannya dalam memahami hubungan antar kalimat, membuatnya sangat berguna untuk berbagai aplikasi pemrosesan bahasa alami (NLP) (HuggingFace, 2024).

2.2.5 IndoBERT / IndoNLU

IndoBERT atau IndoNLU adalah model berbasis transformer yang dikembangkan untuk pemrosesan bahasa Indonesia, mengambil inspirasi dari model BERT dan dilatih menggunakan sejumlah besar data teks dalam Bahasa Indonesia.

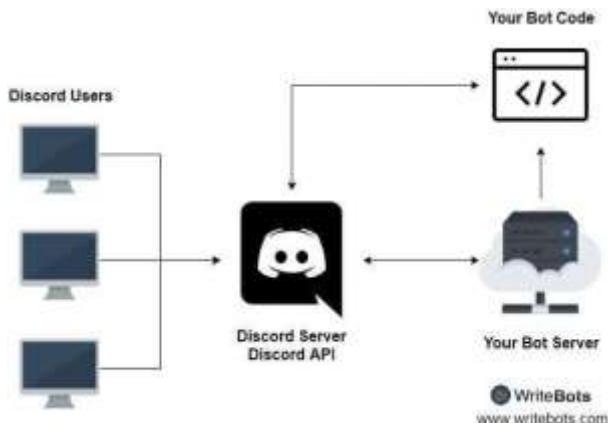


Gambar 2.2.5 IndoBERT/IndoNLU

Model ini mengandalkan lebih dari 220 juta kata yang berasal dari sumber terpercaya, seperti korpus berita daring dan Korpus Web Indonesia. IndoBERT dioptimalkan melalui proses pra-pelatihan yang meliputi 2,4 juta langkah atau sekitar 180 epoch, menghasilkan performa yang baik untuk berbagai tugas NLP. Model ini memanfaatkan teknik masked language modeling (MLM) dan next sentence prediction (NSP) selama pelatihan, dengan arsitektur transformer yang memiliki 12 lapisan dan 768-unit pemrosesan. IndoBERT menggunakan kosakata WordPiece bahasa Indonesia sebanyak 31.923 token. Model ini memiliki beberapa varian, antara lain IndoBERT-base dengan 12 lapisan dan 125 juta parameter, serta IndoBERT-large yang terdiri dari 24 lapisan dan 340 juta parameter (Fajri Koto, 2020).

2.2.6 Discord Bot API

Discord Bot API adalah antarmuka pemrograman aplikasi yang memungkinkan pengembang untuk membuat, menguji, dan mengimplementasikan chatbot pada platform Discord.



Gambar 2.2.6 Discord Bot API

Discord Bot API menyediakan berbagai endpoint dan fitur yang memungkinkan pengembang merancang chatbot tanpa harus memulai dari nol, baik untuk pemrosesan bahasa alami (NLP), integrasi dengan aplikasi pihak ketiga, maupun pengaturan interaksi dengan pengguna Discord secara real-time. Bot dapat diatur untuk merespons pesan, menjalankan perintah tertentu, mengatur peran, hingga melakukan moderasi otomatis dalam server Discord dengan efisien.

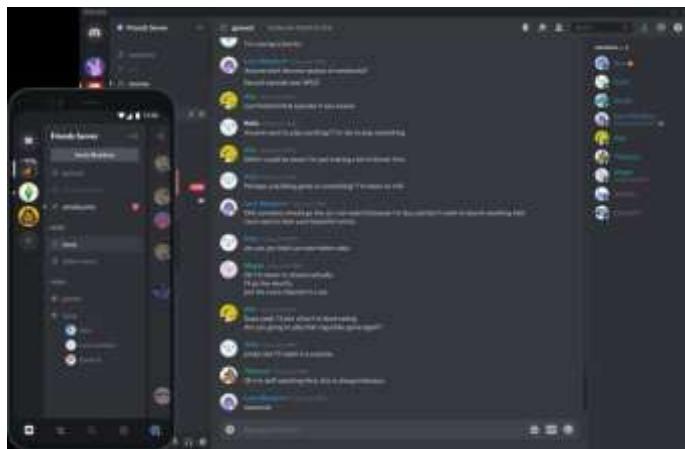
Beberapa fitur utama dari Discord Bot API antara lain:

1. Webhook dan Event Listener untuk mendeteksi pesan atau aktivitas server.
2. Rich Embeds untuk mengirim pesan interaktif dan informatif.
3. Integrasi dengan NLP (misalnya Dialogflow, BERT, atau model NLP lain) untuk memungkinkan chatbot memahami konteks percakapan.

4. Manajemen server dan peran secara otomatis untuk mendukung fungsionalitas komunitas.

2.2.7 Discord

Dalam pengembangan dan penerapan chatbot, Discord memainkan peran penting sebagai platform interaksi antara pengguna dan chatbot.



Gambar 2.2.7 Discord

Discord merupakan salah satu platform populer yang digunakan untuk implementasi chatbot. Platform ini menyediakan API dan fitur khusus yang memungkinkan integrasi chatbot untuk berbagai keperluan, seperti layanan pelanggan, komunikasi komunitas, moderasi server, hingga automasi informasi dalam server Discord. Melalui Discord, chatbot dapat merespons pesan pengguna secara real-time, menjalankan perintah tertentu, mengelola peran dan kanal, serta mendukung interaksi berbasis teks maupun suara dengan efisien.

2.2.8 Python

Python adalah bahasa pemrograman tingkat tinggi yang bersifat interpretatif dan ditujukan untuk berbagai keperluan. Diciptakan oleh Guido van Rossum dan diperkenalkan pada tahun 1991, desain Python

menekankan pentingnya keterbacaan kode melalui penggunaan spasi putih yang signifikan. Struktur bahasa dan pendekatan berorientasi objeknya dirancang untuk membantu programmer menghasilkan kode yang jelas dan terorganisir, baik untuk proyek berskala kecil maupun besar.



Gambar 2.2.8 Python

Python adalah bahasa pemrograman dengan sistem penulisan kode yang bersifat dinamis serta dilengkapi dengan fitur pengelolaan sampah otomatis. Bahasa ini mendukung berbagai paradigma pemrograman seperti prosedural, berorientasi objek, serta fungsional, sehingga fleksibel dalam penggunaannya untuk pengembangan aplikasi (Akbar Nur Syahrudin, 2018).

2.2.9 Akurasi dan F1 Skor

Akurasi dan F1 Skor merupakan dua metrik evaluasi yang digunakan untuk mengukur kinerja model klasifikasi, termasuk dalam

pengembangan chatbot dengan kemampuan deteksi teks seperti deteksi cyberbullying.

Akurasi adalah ukuran yang menentukan seberapa besar proporsi prediksi yang benar dibandingkan dengan keseluruhan data yang diuji. Akurasi memberikan gambaran umum tentang performa model, namun kurang tepat digunakan ketika data tidak seimbang (imbalanced dataset).

F1 Skor merupakan rata-rata harmonik antara presisi dan recall yang berfungsi untuk menilai kinerja model pada dataset yang memiliki distribusi tidak seimbang. Presisi menunjukkan tingkat ketepatan model dalam memprediksi kelas positif, sedangkan recall menunjukkan kemampuan model dalam mendeteksi seluruh data pada kelas positif. F1 Score menjadi indikator penting karena mampu menyeimbangkan pertimbangan antara kesalahan positif palsu (false positive) dan negatif palsu (false negative).

Rumus Akurasi:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Data Total}} \times 100\%$$

$$F1 \text{ Skor} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}$$

Gambar 2.2.9 Rumus F1 Skor

Penggunaan akurasi dan F1 Skor membantu pengembang dalam mengevaluasi performa chatbot secara menyeluruh untuk memastikan hasil deteksi yang tepat dan konsisten, terutama ketika dataset memiliki distribusi yang tidak seimbang antara teks positif (cyberbullying) dan negatif (non-cyberbullying) (Wulandari., 2021).