

**ANALISIS PERBANDINGAN KLASIFIKASI *INTENT*
CHATBOT MENGGUNAKAN *DEEP LEARNING*
BERT, ROBERTA, DAN INDOBERT**



OLEH:

ASWIN DWIYONO

09012682125012

PROGRAM STUDI MAGISTER ILMU KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2024

**ANALISIS PERBANDINGAN KLASIFIKASI *INTENT*
CHATBOT MENGGUNAKAN *DEEP LEARNING*
BERT, ROBERTA, DAN INDOBERT**

TESIS

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister**



OLEH:

ASWIN DWIYONO

09012682125012

**PROGRAM STUDI MAGISTER ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN

**ANALISIS PERBANDINGAN KLASIFIKASI *INTENT*
CHATBOT MENGGUNAKAN *DEEP LEARNING*
BERT, ROBERTA, DAN INDOBERT**

TESIS

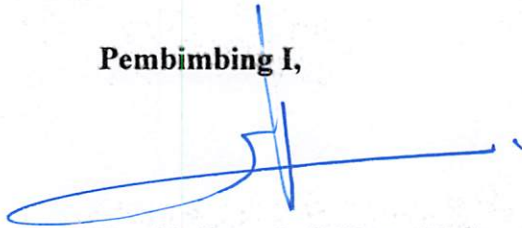
Diajukan untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

Oleh:

Aswin Dwiyono

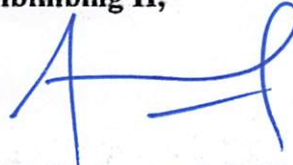
09012682125012

Pembimbing I,



**Dr. Abdiansah, S.Kom, M.Cs.
NIP. 198410012009121005**

**Palembang, Desember 2024
Pembimbing II,**



**Dr. M. Fachrurrozi, M.T.
NIP. 198005222008121002**

**Mengetahui,
Koordinator Program Studi Magister Ilmu Komputer**



**Dr. Firdaus, M.Kom
NIP. 197801212008121003**

HALAMAN PERSETUJUAN

Pada hari Senin tanggal 30 Desember 2024 telah dilaksanakan ujian sidang tesis oleh Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Aswin Dwiyono
NIM : 09012682125012
Judul : Analisis Perbandingan Klasifikasi *Intent Chatbot* Menggunakan *Deep Learning* BERT, RoBERTa, dan IndoBERT

1. Ketua Penguji



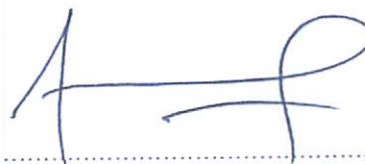
Prof. Dr. Ir. Bambang Tutuko, M.T.
NIP. 196001121989031002

2. Pembimbing I



Dr. Abdiansah, S.Kom, M.Cs.
NIP. 198410012009121005

3. Pembimbing II



Dr. M. Fachrurrozi, M.T.
NIP. 198005222008121002

4. Penguji I



Dian Palupi Rini, M.Kom., Ph.D.
NIP. 198005222008121002

5. Penguji II



Julian Supardi, S.Pd., M.T., Ph.D.
NIP. 197207102010121001

Mengetahui,
Koordinator Program Studi Magister Ilmu Komputer



Dr. Firdaus, M.Kom
NIP. 197801212008121003

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Aswin Dwiyono
NIM : 09012682125012
Program Studi : Magister Ilmu Komputer
Judul : Analisis Perbandingan Klasifikasi *Intent Chatbot*
Menggunakan *Deep Learning* BERT, RoBERTa, Dan
IndoBERT

Hasil Pengecekan Software iThenticate/Turnitin : 14 %

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, Desember 2024



Aswin Dwiyono
NIM. 09012682125012

KATA PENGANTAR

Puji syukur atas Rahmat dan Karunia yang telah Allah SWT berikan sehingga penulis dapat menyelesaikan tesis yang berjudul “Analisis Perbandingan Klasifikasi *Intent Chatbot* Menggunakan *Deep Learning* BERT, RoBERTa, Dan IndoBERT”. Tesis ini disusun untuk memenuhi persyaratan kelulusan tingkat S2 pada Program Studi Magister Ilmu Komputer Universitas Sriwijaya.

Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada pihak yang telah memberikan dukungan, bimbingan, motivasi, dan kemauan kepada penulis untuk menyelesaikan tesis ini:

1. Kedua orang tua serta saudari saya yang tercinta, yang tak henti-hentinya memberikan saya dukungan serta motivasi untuk terus gigih dalam menyelesaikan tesis ini.
2. Bapak Dr. Abdiansah, S.Kom, M.Cs., selaku dosen pembimbing I yang telah sabar memberikan bimbingan dan saran sehingga penulis dapat menyelesaikan tesis ini.
3. Bapak Dr. M. Fachrurrozi, M.T., selaku dosen pembimbing I yang telah sabar memberikan bimbingan dan saran sehingga penulis dapat menyelesaikan tesis ini.
4. Bapak dan Ibu Dosen yang selama ini telah melimpahkan ilmunya kepada penulis selama proses belajar mengajar di Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Staf administrasi yang telah membantu proses administrasi dan akademik selama masa perkuliahan.
6. Teman-teman seperjuangan yang terus menyemangati dan saling mengingatkan untuk menyelesaikan tesis.
7. Semua pihak yang telah membantu dalam penyelesaian tesis ini dan tak dapat disebutkan satu persatu.

Akhir kata, penulis menyadari bahwa tesis ini jauh dari kata sempurna. Untuk itu, penulis mengharapkan kritik dan saran yang membangun dari semua

pihak untuk penyempurnaan penelitian ini. Semoga tesis ini dapat bermanfaat bagi pihak yang membutuhkan.

Palembang, Desember 2024



Penulis

Analisis Perbandingan Klasifikasi *Intent Chatbot* Menggunakan *Deep Learning* BERT, RoBERTa, dan IndoBERT

Aswin Dwiyono

Abstrak

Chatbot adalah aplikasi percakapan yang dirancang untuk menangani masukan dari pengguna dan menghasilkan balasan yang sesuai berdasarkan masukan tersebut, yang kemudian dikomunikasikan kembali kepada pengguna. Agar dapat memberikan balasan yang akurat, *chatbot* harus dapat memahami maksud dari pengguna dengan benar. Salah satu permasalahan dalam pengembangan *chatbot* adalah bagaimana cara melakukan klasifikasi *intent* dari pengguna secara akurat. Kesalahan dalam memahami maksud pengguna dapat menyebabkan balasan yang tidak relevan. Untuk melakukan percakapan dengan pengguna, keinginan dari pengguna perlu diklasifikasikan dengan baik. Klasifikasi dilakukan untuk menentukan maksud dari teks yang dimasukkan pengguna agar sistem *chatbot* dapat memberikan jawaban yang sesuai. Penelitian ini membandingkan tiga model berbasis transformer yang canggih yaitu BERT (*Bidirectional Encoder Representations from Transformers*), RoBERTa (*Robustly Optimized BERT Pretraining Approach*), dan IndoBERT (*Indonesia Bidirectional Encoder Representations from Transformer*) untuk melakukan klasifikasi *intent* dalam sistem *chatbot*. Penelitian ini menggunakan dataset *University Chatbot* yang berisikan 38 *intent* dan telah diterjemahkan ke dalam Bahasa Indonesia. Berbagai metrik performa, termasuk akurasi, F1-score, presisi, dan recall, dianalisis untuk mengevaluasi dan menentukan model yang menghasilkan kinerja paling efektif dengan menggunakan parameter dan dataset yang sama. Pada akhir penelitian, model BERT mencapai akurasi 0,89, model RoBERTa mencapai akurasi 0,84, sedangkan model IndoBERT mencapai akurasi 0,94. Performa IndoBERT lebih baik dibandingkan dengan BERT dan RoBERTa disebabkan oleh pelatihan yang lebih spesifik untuk bahasa Indonesia, *pretraining* yang lebih relevan, dan adaptasi yang lebih efektif terhadap konteks dan struktur bahasa Indonesia.

Kata Kunci: Klasifikasi *Intent*; Chatbot; BERT; RoBERTa; IndoBERT; Transformer

Comparative Analysis of Chatbot Intent Classification Using Deep Learning BERT, RoBERTa, and IndoBERT

Aswin Dwiyono

Abstract

A chatbot is a software application to designed handle user inputs and generate appropriate replies based on those inputs, which are then communicated back to the user. In able to provide accurate responses, the chatbot must be able to understand the intent of the user accurately. An issue in the development of chatbots is how to accurate classify user intent. Incorrectly understanding user intent can result in irrelevant responses. In order to have a conversation with the user, the intent of the user needs to be classified correctly. This paper compares three state-of-the-art transformer-based models BERT (Bidirectional Encoder Representations from Transformers), RoBERTa (Robustly Optimized BERT Pretraining Approach), and IndoBERT (Indonesia Bidirectional Encoder Representations from Transformer) for the task of intent classification in chatbot systems. Various performance metrics, including accuracy, F1-score, precision, and recall, were analyzed to determine which model performs more effectively in the same parameter conditions. Performance metrics like accuracy and F1-score were compared to assess model BERT, RoBERTa and IndoBERT performs better in a University Chatbot Dataset in Indonesian language. The BERT model achieved an accuracy of 0.89, RoBERTa model achieved 0.84 and IndoBERT model achieved an accuracy of 0.94. The better performance of IndoBERT compared to BERT and RoBERTa is caused by more language-specific training, more relevant pretraining, and more effective adaptation to Indonesian context and structure.

Keywords: Intent classification; Chatbot; BERT; RoBERTa; IndoBERT; Transformer

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN.....	iii
LEMBAR PERNYATAAN	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I. PENDAHULUAN	1
1.1. Latar Belakang	4
1.2. Perumusan Masalah	4
1.3. Tujuan Masalah	4
1.4. Manfaat Penelitian	4
1.5. Batasan Penelitian	5
1.6. Metodologi Penulisan	5
BAB II. TINJAUAN PUSTAKA	7
2.1. <i>Chatbot</i>	7
2.2. <i>Intent</i>	8
2.3. Klasifikasi <i>Intent</i>	9
2.4. <i>Large Language Model (LLM)</i>	10
2.4.1. <i>Bidirectional Encoder Representations from Transformers</i>	11
2.4.2. <i>Robustly Optimized BERT Pretraining Approach</i>	14
2.4.3. <i>Indonesian BERT Pretraining Approach</i>	16
2.4.4 Perbandingan BERT, RoBERTa, dan IndoBERT.....	18

2.5. Pra-pengolahan Teks	20
2.6. Alat Ukur Percobaan	21
2.6.1. K-Fold Cross Validation	21
2.6.2. <i>Confusion Matrix</i>	21
2.6.3. Akurasi	22
2.6.4. <i>Precision</i>	22
2.6.5. <i>Recall</i>	23
2.6.6. <i>F1 – Score</i>	23
2.7. Penelitian lain yang relevan	23
BAB III. METODOLOGI PENELITIAN	27
3.1. Tahapan Penelitian	27
3.2. Persiapan dan Pengumpulan Data	29
3.3. Alat dan Bahan Penelitian	39
3.4. Skenario Penelitian	40
3.5. Analisis Hasil	46
3.6. Kesimpulan dan Saran	46
BAB IV. HASIL DAN ANALISIS	47
4.1. Dataset Percobaan.....	47
4.2. Hasil Model BERT.....	49
4.3. Hasil Model RoBERTa.....	53
4.4. Hasil Model IndoBERT.....	56
4.5. Perbandingan Hasil	59
4.6. Perbandingan Waktu	60
BAB V. KESIMPULAN DAN SARAN	63
5.1. Kesimpulan	63
5.2. Saran	64
DAFTAR PUSTAKA	65

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Arsitektur Dasar Model BERT	11
Gambar 2.2 Proses Tokenisasi dalam Model BERT.....	13
Gambar 2.3 Arsitektur Model RoBERTa	15
Gambar 3.1 Diagram Alir Tahapan Penelitian	27
Gambar 3.2 Sampel Data <i>Intent</i>	29
Gambar 3.3 Pra-Pengolahan Data	41
Gambar 4.1 Sampel Dataset	49
Gambar 4.2 Hasil <i>Confusion Matrix</i> Percobaan Model BERT.....	51
Gambar 4.3 Metric Pengukuran tiap <i>Intent</i> Model BERT	52
Gambar 4.4 Hasil <i>Confusion Matrix</i> Percobaan Model RoBERTa.....	54
Gambar 4.5 Metric Pengukuran tiap <i>Intent</i> Model RoBERTa.....	55
Gambar 4.6 Hasil <i>Confusion Matrix</i> Percobaan Model IndoBERT.....	57
Gambar 4.7 Metric Pengukuran tiap <i>Intent</i> Model IndoBERT.....	58
Gambar 4.8 Perbandingan Waktu	61

DAFTAR TABEL

	Halaman
Tabel 2.1 Tabel Varian Model IndoBERT	17
Tabel 2.2 Tabel Perbandingan Model BERT, RoBERTa, dan IndoBERT..	18
Tabel 2.3 Tabel <i>Confusion Matrix</i>	22
Tabel 2.4 Tabel Penelitian Klasifikasi <i>Intent</i>	21
Tabel 3.1 Tabel Dataset <i>Chatbot</i> dalam Bahasa Indonesia	30
Tabel 3.2 Tabel Contoh Data <i>Intent</i>	41
Tabel 3.3 Tabel Contoh <i>Stopword Removal</i>	41
Tabel 3.4 Tabel Contoh <i>Cleaning</i>	42
Tabel 3.5 Tabel Contoh <i>Case Folding</i>	42
Tabel 3.6 Tabel Contoh Change Word.....	42
Tabel 3.7 Tabel Contoh <i>Stemming</i>	42
Tabel 3.8 Tabel Pembagian Data Latih dan Data Uji	43
Tabel 3.9 Tabel Konfigurasi Parameter	44
Tabel 4.1 Tabel Bentuk Dataset Penelitian	47
Tabel 4.2 Tabel Parameter BERT yang digunakan dalam Percobaan.....	49
Tabel 4.3 Tabel Hasil Evaluasi <i>Train</i> dan <i>Test</i> Model BERT.....	50
Tabel 4.4 Tabel Parameter RoBERTa yang digunakan dalam Percobaan..	53
Tabel 4.5 Tabel Hasil Evaluasi <i>Train</i> dan <i>Test</i> Model RoBERTa.....	56
Tabel 4.6 Tabel Parameter IndoBERT yang digunakan dalam Percobaan..	57
Tabel 4.7 Tabel Hasil Evaluasi <i>Train</i> dan <i>Test</i> Model IndoBERT.....	58
Tabel 4.8 Tabel Hasil Perbandingan	59

BAB I

PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah yang menjadi dasar dalam penelitian ini.

1.1 Latar Belakang

Chatbot adalah program yang mampu memproses masukan dari pengguna dan menghasilkan tanggapan atas jawaban yang sesuai dengan masukan dari pengguna yang akan dikirim kembali ke pengguna (Rohim & Zuliarso, 2022). Tujuan utama dari *chatbot* adalah untuk berkomunikasi dengan manusia dan dengan bantuan itu membuat banyak pekerjaan yang berlebihan menjadi lebih mudah bagi manusia (Hutama et al., 2021). *Chatbot* tidak dibatasi waktu dan karena dapat melayani pengguna kapan saja sehingga dapat membuat pekerjaan menjadi efisien. Salah satu elemen kunci dalam mengembangkan *chatbot* yang efektif adalah kemampuan *chatbot* untuk memahami dan mengklasifikasikan *intent* pengguna secara akurat. Untuk melakukan percakapan dengan pengguna, keinginan atau *intent* dari pengguna perlu diklasifikasikan dengan baik. Klasifikasi *intent* merupakan langkah awal yang penting dalam menentukan alur percakapan dan memahami konteksnya guna memberikan respons yang tepat dan relevan.

Ada beberapa metode dan penelitian sebelumnya dalam melakukan klasifikasi *intent* dalam sistem *chatbot* seperti penelitian *intent detection* yang menggunakan *machine learning tradisional* seperti *Naive Bayes*, *Adaboost*, *Support Vector Machine (SVM)* dan *Logistic Regression*, namun teknik ini tidak dapat memahami informasi dalam teks secara menyeluruh sehingga sulit untuk mengetahui intensi pengguna secara akurat (Liu et al., 2019). Penelitian aplikasi *chatbot* berbasis teks menggunakan metode *Naive Bayes Classifier* dengan dataset kumpulan pertanyaan yang sering timbul (FAQ) pada perusahaan GRABADS

menghasilkan nilai akurasi sebesar 93,33% dan nilai kesalahan sebesar 6,66% (Hutama et al., 2021). Penggunaan metode *Naive Bayes Classifier* dikarenakan komputasi yang lebih sederhana namun pada implementasi *chatbot* tidak dapat mengenali persamaan kata dari pertanyaan yang diajukan sehingga mengakibatkan *chatbot* memberikan jawaban yang salah (Hutama et al., 2021). Selanjutnya, terdapat penelitian pengembangan *chatbot* menggunakan metode *Natural Language Processing* berbasis dialogflow yang menghasilkan akurasi sebesar 92.3% (Sanjaya & Winarno, 2024). Metode *dialogflow* ini mampu membuat komputer mampu mengerti bahasa manusia dalam bentuk teks namun tidak dapat memproses pertanyaan yang kompleks sehingga perlu dilakukan pembaruan secara berkala pada data *training phrases chatbot* yang bertujuan untuk meningkatkan pemahaman pada sistem *chatbot* (Sanjaya & Winarno, 2024). Kemudian terdapat juga penelitian deteksi intensi *chatbot* berbahasa Indonesia dari percakapan pelanggan PT. Kaze dengan menggunakan *Metode Capsule Network* atau CapsNet yang menghasilkan nilai akurasi 70% (Fatharani et al., 2022). CapsNet adalah model yang terdiri dari sejumlah kapsul yang merupakan kumpulan *neuron* dan dirancang untuk memproses informasi yang diberikan sehingga dapat mengidentifikasi intensi sebuah teks berdasarkan perhitungan yang dilakukan (Fatharani et al., 2022). Hasil model CapsNet 6 intensi sudah cukup baik namun mengalami penurunan nilai akurasi untuk 18 intensi dan jika metode ini digunakan dalam *chatbot* maka *chatbot* mampu merespon pengguna, akan tetapi respon yang diberikan masih banyak yang tidak sesuai (Fatharani et al., 2022).

Untuk mengatasi tantangan ini, maka muncul teknik *deep learning*, terutama model berbasis transformer seperti BERT (*Bidirectional Encoder Representations from Transformers*), RoBERTa (*Robustly Optimized BERT Pretraining Approach*) dan IndoBERT (*Indonesia Bidirectional Encoder Representations from Transformer*). Model-model ini menggunakan pembelajaran mendalam untuk menangkap hubungan semantik dalam teks, meningkatkan kemampuan *chatbot* dalam memahami *intent* pengguna. BERT telah diadopsi secara luas untuk klasifikasi *intent* karena kemampuannya untuk memahami konteks dengan memproses teks dua arah, menawarkan akurasi yang tinggi dalam tugas klasifikasi

intent (Boudjani et al., 2023). Namun, meskipun BERT memberikan kemampuan yang canggih, BERT bukannya tanpa keterbatasan, seperti bias yang dapat mendukung kelas tertentu di atas yang lain (Sayenju et al., 2022), dan masalah yang terkait dengan sub-tokenisasi, yang dapat mengakibatkan ketidakselarasan yang mempengaruhi keakuratan prediksi label (Guo et al., 2023).

Meskipun BERT efektif, pengembangan RoBERTa oleh Facebook AI menunjukkan peningkatan signifikan dalam prapelatihan dan pemrosesan data BERT. Tidak seperti BERT, RoBERTa menghilangkan tujuan Prediksi Kalimat Berikutnya (NSP), menyederhanakan proses pelatihan dan memungkinkan model untuk lebih fokus pada pemahaman konteks dalam kalimat tunggal, yang mengarah pada peningkatan akurasi dalam klasifikasi *intent* (Liu et al., 2019; Souha et al., 2023). Arsitektur RoBERTa juga memungkinkannya untuk memahami pola dan hubungan yang detail dalam data, yang sangat penting untuk pengenalan *intent* yang efektif, terutama dalam dataset multibahasa (Jayanth et al., 2024). Namun dalam implementasi menggunakan bahasa Indonesia, munculnya IndoBERT telah menjadi terobosan yang lebih signifikan. IndoBERT adalah varian dari BERT yang secara khusus dilatih pada korpus bahasa Indonesia yang dikumpulkan dari sumber-sumber umum seperti media sosial teks, blog, berita, dan situs web sehingga mampu menangkap nuansa bahasa yang unik, termasuk struktur morfologis dan sintaksis yang khas (Wilie et al., 2020).

Dalam membandingkan kinerja BERT, RoBERTa dan IndoBERT untuk klasifikasi *intent chatbot*, ketiga model tersebut memiliki kekuatan yang unik. Model BERT dengan pendekatan dua arah dan pemodelan kontekstual yang baik memberikan hasil yang memuaskan dalam tugas-tugas bahasa alami. Model RoBERTa dengan proses prapelatihan yang telah dioptimalkan membuktikan bahwa pengembangan lebih lanjut dapat meningkatkan akurasi dan efisiensi klasifikasi *intent*. Sementara itu, IndoBERT yang dikembangkan secara khusus untuk bahasa Indonesia menawarkan kemampuan pemodelan yang lebih relevan dengan konteks linguistik lokal. Dengan menggunakan IndoBERT, diharapkan dapat mengetahui performa model dalam memahami nuansa bahasa Indonesia terutama dalam tugas-tugas klasifikasi *intent* yang melibatkan percakapan alami.

Berdasarkan latar belakang sebelumnya, penelitian ini bertujuan untuk melakukan analisis perbandingan klasifikasi *intent chatbot* menggunakan *deep learning* BERT, RoBERTa, dan IndoBERT. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi wawasan yang bermanfaat dalam penerapan model berbasis *deep learning* untuk sistem percakapan otomatis atau *chatbot* yang lebih cerdas dan akurat.

1.2 Perumusan Masalah

Pada latar belakang yang telah dijelaskan sebelumnya terdapat beberapa isu yang akan dibahas dalam penelitian sesuai dengan penjelasan dari latar belakang di atas. Maka dari itu perlu perumusan beberapa masalah dalam penelitian kali ini, yaitu:

1. Bagaimana melakukan klasifikasi *intent* pada *chatbot* menggunakan metode BERT, RoBERTa dan IndoBERT?
2. Bagaimana hasil perbandingan kinerja klasifikasi *intent* pada *chatbot* menggunakan metode BERT, RoBERTa, dan IndoBERT?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai pada penelitian ini adalah:

1. Melakukan klasifikasi *intent chatbot* menggunakan metode BERT, RoBERTa, dan IndoBERT.
2. Mengetahui hasil perbandingan kinerja antara metode BERT, RoBERTa, dan IndoBERT pada klasifikasi *intent chatbot*.

1.4 Manfaat Penelitian

Hasil yang akan didapatkan dari penelitian kali ini adalah:

1. Dapat mengetahui informasi hasil klasifikasi *intent chatbot* menggunakan metode BERT, RoBERTa, dan IndoBERT.
2. Hasil penelitian ini dapat dimanfaatkan sebagai acuan atau dasar untuk pengembangan penelitian lebih lanjut di masa mendatang.

1.5 Batasan Masalah

Hasil yang akan didapatkan dari penelitian kali ini adalah:

1. Data yang akan digunakan merupakan *dataset* dari Kaggle, yaitu *University Chatbot Dataset* dan diterjemahkan ke Bahasa Indonesia.
2. Acuan yang menjadi perbandingan kinerja hanya akurasi, *recall*, *precision*, dan *F1-score*.

1.6 Metodologi Penulisan

Agar memperoleh gambaran jelas mengenai penelitian ini, maka dibuatlah suatu sistematika penulisan yang berisi gambaran dalam tiap bab penelitian ini, yaitu:

1. BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat dari topik yang dipilih berupa analisis perbandingan klasifikasi *intent chatbot* dengan menggunakan *deep learning* BERT, RoBERTa, dan IndoBERT.

2. BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan mengenai literatur revidu klasifikasi *intent chatbot* dengan menggunakan *deep learning* BERT, RoBERTa, dan IndoBERT serta menggunakan *University Chatbot Dataset* yang diterjemahkan ke dalam Bahasa Indonesia. Algoritma yang akan digunakan dalam penelitian, dan penambahan beberapa metode yang bisa mendukung peningkatan nilai akurasi.

3. BAB III METODE PENELITIAN

Pada bab ini penulis melakukan pembahasan secara bertahap dan terperinci langkah-langkah yang akan digunakan pada penelitian mengenai analisis perbandingan klasifikasi *intent chatbot* dengan menggunakan *deep*

learning BERT, RoBERTa, dan IndoBERT. Parameter yang akan digunakan untuk menentukan hasil pengujian yaitu akurasi, *recall*, *precision*, dan *F1-score*.

4. BAB IV HASIL DAN ANALISA

Pada bab ini berisi tentang hasil dari penelitian dan analisa hasil penelitian yang telah dilakukan terhadap data-data yang telah dikumpulkan. Hasil penelitian berupa nilai dari evaluasi kinerja model-model. Dari hasil penelitian akan dianalisa dan dibahas masing-masing evaluasi model.

5. BAB V KESIMPULAN DAN SARAN

Pada bab ini berisi tentang kesimpulan dari hasil dan analisa pengolahan data yang telah dilakukan sebelumnya. Bab ini menjelaskan detail mengenai kesimpulan dari penelitian yang telah dilakukan, serta saran untuk penelitian mendatang.

DAFTAR PUSTAKA

- Aljabar, A. (2024). *Mengungkap Opini Publik: Pendekatan BERT-based- caused untuk Analisis Sentimen pada Komentar Film*. 5(1).
- Anugerah Simanjuntak, Rosni Lumbantoruan, Kartika Sianipar, Rut Gultom, Mario Simaremare, Samuel Situmeang, & Erwin Panggabean. (2024). Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 13(1), 60–67. <https://doi.org/10.22146/jnteti.v13i1.8532>
- Arisya, Nur Sanifa.(2021). Klasifikasi Intent Chatbot Terapi Menggunakan Multinomial Naive Bayes dan Mutual Information. Sriwijaya University Repository.<http://repository.unsri.ac.id/54113/>
- DeepL. (2024). *DeepL Translator*. <https://www.deepl.com/>
- Fatharani, F., Kania, K. P., Hutahaeen, J., & Wulan, S. R. (2022). Deteksi Intensi Chatbot Berbahasa Indonesia dengan Menggunakan Metode Capsule Network. *Journal of Information System Research (JOSH)*, 3(4), 590–596. <https://doi.org/10.47065/josh.v3i4.1821>
- Ferlin, J., Bachtiar, F. A., & Rusydi, A. N. (2009). Klasifikasi Customer Intent untuk Mengetahui Tingkat Kepuasan Pelanggan menggunakan Metode Support Vector Machine pada Restoran Bakso President. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Fudholi, D. H. (2023). *Mi-Botway Chatbot FAQs Universitas Berbasis Deep Learning—Yurio Windiatmoko—19917020 (1)*.
- Guo, Y., Xie, Z., Chen, X., Chen, H., Wang, L., Du, H., Wei, S., Zhao, Y., Li, Q., & Wu, G. (2023). *ESIE-BERT: Enriching Sub-words Information Explicitly with BERT for Joint Intent Classification and SlotFilling* (arXiv:2211.14829). arXiv. <http://arxiv.org/abs/2211.14829>
- Harahap, D. W., & Fitria, L. (2020). Aplikasi Chatbot Berbasis Web Menggunakan Metode Dialogflow. *J-ICOM - Jurnal Informatika dan Teknologi Komputer*, 1(1), 6–13. <https://doi.org/10.33059/j-icom.v1i1.2796>
- Hefny, A. H., Dafoulas, G. A., & Ismail, M. A. (2020). Intent Classification for a Management Conversational Assistant. *2020 15th International Conference on Computer Engineering and Systems (ICCES)*, 1–6. <https://doi.org/10.1109/ICCES51560.2020.9334685>
- Setyawan H., M. Y., Awangga, R. M., & Efendi, S. R. (2018). Comparison Of Multinomial Naive Bayes Algorithm And Logistic Regression For Intent Classification In Chatbot. *2018 International Conference on Applied Engineering (ICAE)*, 1–5. <https://doi.org/10.1109/INCAE.2018.8579372>
- Hutama, R. C., Fauziah, F., & Komalasari, R. T. (2021). Aplikasi Chatbot Berbasis Teks Menggunakan Algoritma Naive Bayes Classifier FAQ GrabAds. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 6(1), 90. <https://doi.org/10.30998/string.v6i1.9919>
- Jayanth, K. K., Bharathi Mohan, G., Kumar, R. P., & Rithani, M. (2024). Intent Recognition Leveraging XLM-RoBERTa for Effective NLU. *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 877–882. <https://doi.org/10.1109/ICAAIC60222.2024.10575275>

- Khusuma, R., Maharani, W., & Gani, P. H. (2023). Personality Detection On Twitter User With RoBERTa. *Jurnal Media Informatika Budidarma*, 7.
- Liu, J., Li, Y., & Lin, M. (2019). Review of Intent Detection Methods in the Human-Machine Dialogue System. *Journal of Physics: Conference Series*, 1267(1), 012059. <https://doi.org/10.1088/1742-6596/1267/1/012059>
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (arXiv:1907.11692). arXiv. <http://arxiv.org/abs/1907.11692>
- Menda, M., & Keerthi, G. S. (2022). Intent Classification in Conversational System using Machine Learning Techniques. *International Journal of Computer Applications*, 183(51), 6–11. <https://doi.org/10.5120/ijca2022921913>
- N, D. G., G, M. V., & T A, V. (2021). Intent Classification using BERT for Chatbot application pertaining to Customer Oriented Services. *Proceedings of the First International Conference on Combinatorial and Optimization, ICCAP 2021, December 7-8 2021, Chennai, India*. Proceedings of the First International Conference on Combinatorial and Optimization, ICCAP 2021, December 7-8 2021, Chennai, India, Chennai, India. <https://doi.org/10.4108/eai.7-12-2021.2314563>
- Nirali Vaghani. (2024). *Chatbot dataset* [Dataset]. Kaggle. <https://doi.org/10.34740/KAGGLE/DSV/5024271>
- Oktavia, C. A. (2020). Implementasi Chatbot Menggunakan Dialogflow dan Messenger Untuk Layanan Customer Service Pada E-Commerce. *J I M P - Jurnal Informatika Merdeka Pasuruan*, 4(3). <https://doi.org/10.37438/jimp.v4i3.230>
- Rezkanintio, A. N. (2021). *Klasifikasi Intent Pada Chatbot Menggunakan Metode Query Expansion dan Neural Network*. 103.
- Rizqie, Q., Afifah, N., & Bardadi, A. (2023). *Eksplorasi Penggunaan Large Language Model (LLM) dalam Pembangunan Permainan Minesweeper dengan Python Programming*. 2(3).
- Rohim, & Zuliarso. (2022). Penerapan Algoritma Deep Learning Untuk Pengembangan Chatbot Yang Digunakan Untuk Konsultasi Dan Pengenalan Tentang Virus Covid-19. *Pixel :Jurnal Ilmiah Komputer Grafis*, 15(2), 267–278. <https://doi.org/10.51903/pixel.v15i2.777>
- Sanjaya, R. A., & Winarno, E. (2024). Pengembangan Chatbot Informasi Pariwisata di Kabupaten Pati Menggunakan Metode Natural Language Processing Berbasis Dialogflow. *Jutisi : Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 13(1), 368. <https://doi.org/10.35889/jutisi.v13i1.1828>
- Sayenju, S., Aygun, R., Boardman, J., Don, D. P. R., Zhang, Y., Franks, B., Johnston, S., Lee, G., Sullivan, D., & Modgil, G. (2022). Quantification and Mitigation of Directional Pairwise Class Confusion Bias in a Chatbot Intent Classification Model. *International Journal of Semantic Computing*, 16(04), 497–520. <https://doi.org/10.1142/S1793351X22500040>
- Situmorang, G. F., & Purba, R. (2024). Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan Teknik NLP dan Model IndoBERT. *Building of Informatics, Technology and Science (BITS)*, 6(2).

- Souha, A., Ouaddi, C., Benaddi, L., & Jakimi, A. (2023). Pre-Trained Models for Intent Classification in Chatbot: Comparative Study and Critical Analysis. *2023 6th International Conference on Advanced Communication Technologies and Networking (CommNet)*, 1–6. <https://doi.org/10.1109/CommNet60167.2023.10365312>
- Suwarningsih, W., Pramata, R. A., Rahadika, F. Y., & Purnomo, M. H. A. (2022). RoBERTa: Language modelling in building Indonesian question-answering systems. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 20(6), 1248. <https://doi.org/10.12928/telkomnika.v20i6.24248>
- Tandijaya, J. H., & Sugiarto, I. (2021). *Klasifikasi dalam Pembuatan Portal Berita Online dengan Menggunakan Metode BERT. Vol 9, No 2 (2021)*.
- 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* (arXiv:2009.05387). arXiv. <http://arxiv.org/abs/2009.05387>
- Zain, M. M., Simbolon, R. N., & Sulung, H. (2021). Analisis Sentimen Pendapat Masyarakat Mengenai Vaksin Covid-19 Pada Media Sosial Twitter dengan Robustly Optimized BERT Pretraining Approach. *Jurnal Komputer Terapan, Vol 7, No 2*.