ISSN: 2355-9365

Implementasi NLP Pada Chatbot Layanan Akademik Dengan Algoritma Bert

Implementation Of NLP On Academic Service Chatbot With Bertalgorithm

1st Nurul Afifa
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
nurulafifa@student.telkomun
iversity.ac.id

2nd Randy Erfa Saputra
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
resaputra@telkomuniversity.ac.id

3rd Ratna Astuti Nugrahaeni Fakultas Teknik Elektro Universitas Telkom Bandung, Indonesia ratnaan@telkomuniversity.ac.id

Abstrak— Penggunaan Chatbot di era modern seperti saat ini tentunya tidak asing lagi dan sudah banyak dijumpai pada berbagai aplikasi. Dimana terkadang untuk mendapatkan informasi mengenai perkuliahan mahasiswa agak kesulitan dan sering kali datang untuk menanyakan secara langsung kepada petugas namun dengan begitu mahasiswa biasanya hanya dapat menanyakan pertanyaan yang inti saja dikarenakan adanya batas waktu. Oleh karena itu untuk membantu pekerjaan petugas layanan akademik dalam menjawab pertanyaan yang ditanyakan oleh mahasiswa kepada universitas tersebut. dikembangkan sebuah chatbot yang terhubung langsung dengan aplikasi Telegram untuk memudahkan mahasiswa dalam menjawab semua pertanyaan mengenai informasi yang ditanyakan. Di dalam pembuatan Chatbot ini sendiri ada beberapa metode yang diperlukan salah satunya adalah Natural Language Processing (NLP) yang berfokus pada pengolahan bahasa natural yang bertujuan untuk membantu mesin memahami Bahasa manusia secara lebih akurat. Selain menggunakan metode NLP penelitian ini juga menggunakan framework RASA untuk membangun Chatbot menggunakan agoritma BERTuntuk membantu komputer dalam memahami arti Bahasa yang ambigu dalam sebuah teks dengan menggunakan teks sekitarnya untuk membangun konteks. Sistem chatbot layanan akademik ini menggunakan jenis pengujian Non-Response-Rate yaitu dengan hasil akurasi yang didapatkan sebesar 85% sehingga dapat disimpulkan bahwa sistem ini berjalan sesuai dengan tujuannya.

Kata kunci: chatbot, telegram, rasa, BERT

Abstrak— The use of Chatbots in the modern era as it is today is certainly not foreign and has been found in many applications. Where sometimes to get information about student lectures is a bit difficult and often comes to ask directly to the officer but that way students can usually only ask basic questions because there is a

time limit. Therefore, to assist the work of academic service officers in answering questions asked by students to the university. developed a chatbot that is directly connected to the Telegram application to make it easier for students to answer all questions regarding the information asked. In making this Chatbot itself, there are several methods needed, one of which is Natural Language Processing (NLP) which focuses on natural language processing which aims to help machines understand human language more accurately. In addition to using the NLP method, this research also uses the RASA framework to build a chatbot and uses the BERT algorithm to help computers understand the meaning of ambiguous language in a text by using the surrounding text to build context. This academic service chatbot system uses the Non-Response-Rate type of test, with an accuracy of 85%, so it can be concluded that this system is running according to its purpose.

I. PENDAHULUAN

Chatbot adalah sebuah program komputer yang dirancang untuk menyimulasikan percakapan intelektual dengan satu atau lebih manusia baik secara audio maupun teks. Pada mulanya, program komputer (bots) ini diuji dengan merahasiakan identitasnya sebagai mesin sehingga dapat mengelabui orang yang bercakap-cakap dengannya. Jika pengguna tidak dapat mengidentifikasi bots sebagai suatu program komputer, maka chatterbot tersebut dikategorikan sebagai kecerdasan buatan (artificial intelligence). Kemampuan komputer dalam menyimpan banyaknya data tanpa melupakan satu pun informasi yang disimpannya digabungkan dengan kepraktisan bertanya pada sumber informasi langsung dibandingkan dengan mencari

informasi sendiri serta kemampuan learning yang dimilikinya menyebabkan chatbot adalah customer service yang handal. Chatbot juga memiliki fungsi dan peran yang sangat penting dalam bidang akademik karena dapat memudahkan sebuah instansi untuk menjawab semua pertanyaan yang diajukan tanpa harus menjawab secara manual.

Untuk pembuatan chatbot ini sendiri menggunakan sebuah kata kunci yang dimasukkan oleh pengguna dan menghasilkan respon balasan yang dianggap paling tepat, atau kata-kata yang dianggap paling mendekati dan pada umumnya menggunakan pendekatan Natural Language Processing (NLP). Natural Language Processing merupakan salah satu tujuan jangka panjang dari Artficial Intelegence (kecerdasan buatan) yaitu pembuatan program yang memiliki kemampuan untuk memahami bahasa manusia. Pada prinsipnya bahasa alami adalah suatu bentuk representasi dari suatu pesan yang ingin dikomunikasikan antar manusia. Bentuk utama representasinya adalah berupa suara/ucapan, tetapi sering pula dinyatakan dalam bentuk tulisan.

Tidak hanya menggunakan NLP disini juga menggunakan metode NLU atau Natural Language Understanding yang merupakan sub bagian dari NLP dimana setelah komputer mengerti apa maksud dari sebuah dokumen/ kalimat (Document classification task), komputer akan mengelurkan informasi dari hasil sebelumnya, seperti hasil klasifikasi, entitiesnya yang dapat diolah kembali oleh engineer. Natural Language Processing (NLP) digunakan dalam pemrosesan dokumen karena user menentukan relevansi dari dokumen dengan membaca menganalisisnya. Jika sistem dapat melakukan analisis dokumen secara otomatis, maka proses pencarian dokumen yang relevan akan lebih mudah. Sistem ini dirancang untuk memudahkan penggunanya dalam mendapatkan informasi dengan mudah menggunakan Telegram. Metode Languange model (BERTdan RASA), digunakan untuk mendukung kelancaran dan pembuatan sistem chatbot ini.)

II. KAJIAN TEORI

A. Chatbot

Chatbot merupakan sistem perangkat lunak yang dapat berinteraksi dengan manusia dengan bahasa natural layaknya seorang manusia. Sistem ini bersandar pada teknikteknik Natural Language Processing (NLP) baik untuk menginterpretasikan teks masukkan dan untuk memberikan respon teks [6].

Dalam bahasa sehari-hari Chatbot merupakan Aplikasi atau Program computer yang dirancang untuk meniru manusia itu sendiri, batasan yang diambil dari Chatbot adalah mampu meniru komunikasi manusia. Jadi jika manusia sedang bercakap-cakap dengan program ini, seakanakan ada 2 pribadi manusia yang saling berkomunikasi. Bot sudah dirancang untuk merespon segala jenis pertanyaan dan pernyataan yang diinputkan oleh manusia (user) [8].

B. NLP

Natural Language Processing atau pengolahan bahasa alami biasanya disingkat dengan NLP, merupakan bidang kecerdasan buatan dimana komputer didesain untuk dapat berkomunikasi dengan manusia menggunakan bahasa alami, seperti Bahasa Indonesia. NLP tidak bertujuan untuk mentransformasikan bahasa yang diterima dalam bentuk teks atau suara menjadi data digital dan/atau sebaliknya pula; melainkan bertujuan untuk memahami arti dari kalimat yang diberikan dalam bahasa alami dan memberikan respon yang sesuai, misalnya dengan melakukan suatu aksi tertentu atau menampilkan data tertentu (Fendy Ridwan, 2013) [7].

C. NLU

Natural Language Understanding atau NLU adalah cabang dari kecerdasan buatan yang menggunakan perangkat lunak komputer untuk memahami input berupa kalimat menggunakan teks atau ucapan. NLU memungkinkan interaksi manusia-komputer, pemahaman bahasa manusia seperti Inggris, Spanyol dan Perancis, yang memungkinkan komputer untuk memahami perintah tanpa sintaks formal bahasa komputer. NLU juga memungkinkan komputer untuk berkomunikasi kembali ke manusia dalam bahasa mereka sendiri.

Tujuan utama NLU adalah membuat bot yang mendukung obrolan dan suara yang dapat berinteraksi dengan publik tanpa pengawasan. Banyak perusahaan IT besar, seperti Amazon, Apple, Google dan Microsoft, dan perusahaan rintisan memiliki proyek NLU yang sedang berjalan.

Pengenalan entitas adalah jenis NLU spesifik yang berfokus pada pengidentifikasian entitas dalam pesan, kemudian mengekstrak informasi terpenting tentang entitas tersebut. Ada dua jenis entitas: entitas bernama dan entitas numerik. Entitas yang dinamai dikelompokkan ke dalam kategori -- seperti orang, perusahaan, dan lokasi. Entitas numerik diakui sebagai angka, mata uang, dan persentase [4].

D. BERT

BERT adalah teknik pra-pelatihan berbasis jaringan saraf untuk *Natural Language Processing* (NLP) yang dikenal dengan nama *Bidirectional Encoder Representations from Transformers* (BERT). Lahirnya BERTini bertujuan untuk memudahkan komputer untuk memahami bahasa selayaknya manusia.

Pada BERT, input dari suatu teks kalimat akan diubah menjadi token sequence. Dimana terdapat token [CLS] yang merupakan token untuk melambangkan mulainya suatu kalimat dan token [SEP] yang merupakan token untuk memisahkan antar kalimat. Selain itu, juga terdapat token [PAD] untuk melakukan penambahan padding atau token kosong ke dalam kalimat untuk memenuhi *fixed length* dari suatu kalimat yang diinputkan. *Fixed length* pada model BERT adalah 512 [14].

Dalam proses pelatihan BERT, model menerima pasangan kalimat sebagai input dan belajar memprediksi apakah kalimat kedua dalam pasangan tersebut adalah kalimat berikutnya dalam dokumen asli. Selama pelatihan, 50% dari input adalah pasangan di mana kalimat kedua adalah kalimat berikutnya dalam dokumen asli, sedangkan 50% lainnya adalah kalimat acak dari korpus yang dipilih sebagai kalimat kedua. Asumsinya adalah kalimat acak akan

terputus dari kalimat pertama. Untuk membantu model membedakan dua kalimat dalam pelatihan, input diproses dengan cara berikut sebelum masuk ke model [15].

E. RASA

RASA adalah implementasi open source untuk pemrosesan bahasa alami (NLU) dan model Dual Intent and Entity Transformer (DIET). RASA merupakan kombinasi dari dua modul: RASA NLU dan RASA Core. RASA NLU menganalisis input pengguna, lalu mengklasifikasikan maksud pengguna dan mengekstrak entitas. RASA NLU menggabungkan annotator yang berbeda dari spaCy parse untuk menginterpretasikan data input [5].

Umumnya ada dua tahap pemrosesan teks pada sebuah sistem chatbot atau sistem percakapan berbasis kecerdasan buatan. Tahap pertama adalah bagaimana sistem menginterpretasikan teks dari pengguna. Tahap kedua adalah bagaimana sistem menghasilkan teks respon. Framework ini merepresentasikan kedua tahapan tersebut atas dua bagian yaitu RASA NLU dan RASA Core. RASA NLU merupakan bagian yang melakukan proses NLU atau pemahaman teks dan RASA Core melakukan manajemen dialog. Pesan yang diterima oleh sistem percakapan akan diinterpretasikan oleh RASA NLU kemudian diproses oleh RASA Core yang memprediksi dan mengatur pemberian respon ke pengguna [6].

F. Ngrok

Ngrok adalah proxy server untuk membuat/membuka jaringan private melalui NAT atau firewall untuk menghubungkan localhost ke internet dengan tunnel yang aman atau bisa dikatakan sebagai aplikasi yang menghubungkan untuk menerobos system firewall yang ada pada jaringan dan juga bisa digunakan untuk keperluan agar IP lokal pada laptop bisa diubah menjadi IP publik, sehingga alamat IP lokalnya akan diubah menjadi sebuah alamat URL, yang nantinya bisa di akses melalui Internet [3].

G. Telegram

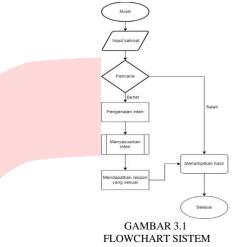
Telegram adalah aplikasi pesan instan berbasis cloud yang fokus pada kecepatan dan keamanan. Telegram dirancang untuk memudahkan pengguna saling berkirim pesan teks, audio, video, gambar dan stiker dengan aman. Secara default, seluruh konten yang ditransfer akan dienskripsi berstandar internasional. Dengan demikian, pesan yang terkirim sepenuhnya aman dari pihak ketiga bahkan dari Telegram sekalipun. Bukan hanya teks, gambar dan video, Telegram juga bisa jadi sarana untuk mengirimkan dokumen, musik, berkas zip, lokasi real-time dan kontak yang tersimpan keperangkat orang lain. Telegram merupakan aplikasi berbasis cloud, yang memudahkan penggunanya dapat mengakses satu account Telegram dari perangkat yang berbeda dan secara bersamaan [2].

III. METODE

A. Gambaran Umum Sistem

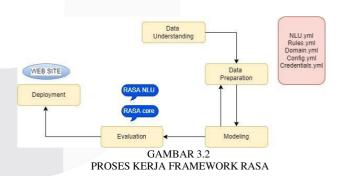
Gambaran umum dari sistem peneletian ini yaitu pengguna atau user akan mendapatkan penjelasan mengenai hal yang ditanyakan atau yang di input pada bot. Alur atau dialog manajemen struktur yang digunakan dalam pendekatan penelitian ini adalah sebagai berikut:

- 1. Tahap pertama yang dilakukan memulai bot
- 2. Langkah kedua yaitu menginput kalimat atau kata
- 3. Ketiga yaitu pencarian pemberitahuan pesan yang diterima.
- 4. Selanjutnya langkah keempat yaitu penerjemahan intent
- 5. langkah kelima yaitu penyesuaian intent yang didapatkan dari hasil input
- 6. langkah keenam menampilkan respon yang sesuai
- 7. Langkah tujuh menampilkan hasil yang telah diperoleh
- 8. Kedelapan proses selesai



B. Sistem Perancangan Chatbot

Penelitian ini disusun menggunakan algoritma BERT, dimana algoritma ini sendiri bertujuan untuk membuat sistem lebih mudah dalam memahami konteks pencarian yang dimaksud oleh user dengan menelaah korelasi dari setiap kata kunci yang diketik. Berikut ilustrasi dari pembuatan chatbot.



C. Implementasi BERT

Dalam penelitian ini digunakan tiga pipeline di RASA yang disimpan didalam dokumen RASA untuk menjalankan BERT. Hal yang pertama dilakukan adalah membuat tiga buah file konfigurasi dalam folder yang diberi nama "config", yang dimana pada masing-masing folder hanya berisi bagian pipa yang relevan untuk pelatihan model NLU dan tidak menyebabkan dialog apapun. Untuk tiga buah pipeline tersebut masing-masing diberi nama:

- a. config-light.yml
- b. config-convert.yml
- c. config-heavy.yml

Dalam konfigurasi ringan RASA memilih CountVectorsFeaturizer yang membuat representasi bag-of-

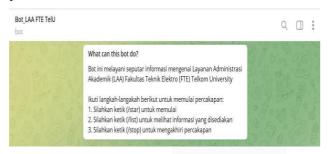
word untuk setiap pesan masuk (pada tingkat kata dan karakter). Namun pada konfigurasi berat akan diganti menggunakan model BERT didalam pipa. HFTransformers NLP adalah komponen utilitas yang melakukan pekerjaan angkat berat dengan memuat model BERT di memori. Selain itu juga RASA menggunakan TransformersHuggingFace untuk menganalisis model Bahasa yang ditentukan dan menambahkan dua komponen tambahan yaitu LanguageModelTokenizer dan LanguageModelFeaturizer yang masing-masing mengambil token dan vector fitur yang dibangun oleh komponen utilitas.

Dalam konfigurasi konversi, komponen ConveRT bekerja dengan cara yang sama. Mereka memiliki tokenizer sendiri serta fitur mereka sendiri. Dengan menggunakan model bahasa ini sebagai fitur, yang berarti bahwa parameternya tidak disesuaikan selama pelatihan model hilir di saluran NLU Anda. Ini menghemat banyak waktu komputasi, dan model machine learning dalam pipeline biasanya dapat mengimbangi kurangnya penyesuaian.

IV. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Pada tahap ini merupakan tahap implementasi sistem yang berisikan tampilan aplikasi dan rancangan proses.



GAMBAR 4.1 TAMPILAN *HOMEPAGE*

Gambar 4.1 merupakan tampilan awal dari sistem chatbot layanan akademik yang menggunakan algoritma BERT. Pada halaman ini menampilkan langkah-langkah untuk memulai bot, disini user dapat mengikuti arahan yang telah dituliskan.



GAMBAR 4.2 TAMPILAN SETELAH /START

Gambar 4.2 merupakan halaman setelah menuliskan /star yang berarti user sudah siap untuk memulai percakapan dengan bot.



GAMBAR 4.3 TAMPILAN /*LIST*

Pada gambar 4.3 merupakan halaman informasi mengenai apa saja topik yang disediakan oleh bot dan dapat ditanyakan oleh user, pada halaman ini juga user dapat memilih topik apa yang akan ditanyakan dan ingin dicari, kemudian setelah itu bot akan menampilkan hasil dari inputan oleh user.

B. Non-Response-Rate

Pengujian Non-Response-Rate yaitu bertujuan untuk mengukur berapakali chatbot gagal dalam menanggapi pertanyaan yang disebabkan karena kurangnya konten atau kesulitan bot dalam memahami pertanyaan pengguna [13]. Didalam penelitian ini dilakukan pengujian non-responserate dengan menggunakan 25 kombinasi pertanyaan untuk setiap masing-masing topik.

TABLE 4.1 PENGUJIAN NON-RESPONSE-RATE

No	Topik	Jumlah respon benar	Jumlah respon eror	Total	Akurasi
1	Registrasi matakuliah	25	0	25	100%
2	Jadwal UTS	20	5	25	80%
3	Jadwal UAS	17	8	25	68%
4	Pembayaran BPP	25	0	25	100%
5	KSM	19	6	25	76%
6	Kartu ujian	21	4	25	84%
7	tugas akhir	23	2	25	92%
8	SK/TA	21	4	25	84%
9	Sidang	21	4	25	84%
10	Bimbingan	21	4	25	84%
11	Kerja praktek	20	5	25	80%
12	Ujian susulan	20	5	25	80%
13	Jadwal kuliah	21	4	25	84%
14	Magang	23	2	25	92%
		RATA- RATA	<u>.</u>		85%

Berdasarkan hasil perhitungan yang dilakukan diatas dengan jumlah topik sebanyak 14 variasi dimana setiap topik diberikan 5 pertanyaan dengan total kombinasi sebesar 25 pertanyaan sehingga diperoleh total sebanyak 70 jenis pertanyaan yang berbeda, didapatkan nilai akhir sebesar 85% sehingga dapat disimpulkan bahwa sistem ini dapat bekerja sesuai tujuan dari tugas akhir ini, yaitu memberikan informasi terkait pertanyaan yang dimasukkan oleh responden.

C. Pengujian Sistem Chatbot

Pengujian ini bertujuan untuk validasi data jika terjadi kesalahan atau interkonsistensi yang muncul pada domain, data NLU, dan data pada rules. Dari hasil pengujian ini dapat

dilihat bahwa data yang telah di input sudah benar sesuai dengan validasi yang dilakukan. Ada dua tahap yang dilakukan dalam validasi ini yaitu rasa train nlu dan cross valitation, untuk contoh validasi yang dilakukan dapat dilihat pada gambar berikut:

GAMBAR 4.8 PENGUJIAN SISTEM CHATBOT

Pada pengujian yang dilakukan diatas ada dua tahap yang pertama adalah rasa test nlu dimana pada tahap ini akan dilakukan validasi pada data yang telah di input di NLU.yml, setelah itu dilakukan validasi silang atau cross-validition yaitu secara otomatis membuat beberapa pemisahan pasa train/test sehingga diperoleh rata-rata hasil pada setiap validasi data.

V. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dalam membuat sistem chatbot, terdapat beberapa kesimpulan yang dapat diambil sebagai berikut:

- 1. Bersadarkan perancangan sistem yang dilakukan menggunakan algoritma BERTdengan framework RASA, chatbot ini mampu memberikan respon sesuai dengan pertanyaan yang di input oleh user.
- 2. Bersadarkaan hasil pengujian Non-Response-Rate, dapat disimpulkan bahwa sistem dapat mengenali berbagai pola kalimat dengan nilai akurasi sebesar 85%. Hal ini dipengaruhi karena adanya proses BERT yang membantu dalam menganalisis kalimat yang di input oleh user.

REFERENSI

- [1] JEISBI: Volume 02 Number 03, 2. (2021). Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT(Bidirectional Encoder Representations from Transformers) . Raden Mas Rizqi Wahyu Panca Kusuma Atmaja1, Wiyli Yustanti2, 55-62.
- [2] Loren Natasya Gunawan1, J. A. (n.d.). Aplikasi Bot Telegram Untuk Media Informasi Perkuliahan Program Studi Informatika-Sistem Informasi Bisnis Universitas Kristen Petra .
- [3] Fifit Fitriansyah1, A. (2020). Penggunaan Telegram Sebagai Media Komunikasi Dalam Pembelajaran Online . Cakrawala-Jurnal Humaniora, Vol 20 No.2 September 2020 , 111-117.
- [4] Khan, W. (2021, February 16). APA ITU NATURAL LANGUAGE PROCESSING? https://algorit.ma/blog/natural-language-processing/, p.

- [5] Nguyen Thi Mai Trang, M. S. (2021). Enhancing RASA NLU model for Vietnamese chatbot. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 9, no.1, 2021, 31-36.
- [6] Dirko G. S. Ruindungan, A. J. (2021). Chatbot Development for an Interactive Academic Information Services using the RASA Open Source Framework . Jurnal Teknik Elektro dan Komputer vol.10 no.1 January-April 2021, pp. 61-68, 61-68.
- [7] SARI, D. W. (2018). IMPLEMENTASI NATURAL LANGUAGE PROCESSING PADA CHATBOT PERIBAHASA INDONESIA. MEDAN.
- [8] Maitri 1, A. L. (2019). RANCANG BANGUN CHATBOT SEBAGAI PUSAT INFORMASI LEMBAGA KURSUS DAN PELATIHAN MENGGUNAKAN PENDEKATAN NATURAL LANGUAGE PROCESSING. eprints Universitas Teknologi Yogyakarta, 1-9.
- [9] Yurio Windiatmoko1, R. R. (2020). Developing Facebook Chatbot Based on Deep Learning Using RASA Framework for University Enquiries. IOP Conf. Series: Materials Science and Engineering 1077 (2021) 012060, 1-11
- [10] Vincentius Riandaru Prasetyo1, N. B. (2021). Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya. TEKNIKA, Volume 10(2), Juli 2021, pp. 114-121, 114-121.
- [11] Rahadi, D. R. (2014). Pengukuran Usability Sistem Menggunakan Use Questionnaire Pada Aplikasi android. Jurnal Sistem Informasi (JSI), VOL. 6, NO. 1, April 2014, , 661-671.
- [12] Yusup, F. (2018). UJI VALIDITAS DAN RELIABILITAS INSTRUMEN PENELITIAN KUANTITATIF. Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan, 17-23.
- [13] RoBERTM. Groves, . J. (2009). Survey Methodology. canada: john wiley& sons, inc. Hoboken, new jersey.
- [14] Franky Halim, L. K. ((2022)). Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT. JURNAL INFRA VOL 10, NO 1 (2022).
- [15] Uszkoreit*, A. V. (2017). Attention Is All You Need. Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA., 1-15.