Implementasi Rasa Pada Chatbot Layanan Akademik

Rasa's Implementation On The Academic Service Chatbot

1st Mutiara Zaafira Itonsaputri
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
mutiarazaa@student.telkomuniversity.ac.id

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

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

Abstrak— Chatbot adalah program perangkat lunak yang dibangun untuk berinteraksi melalui kata-kata tertulis atau lisan dengan pengguna. *Chatbot* telah digunakan di banyak bidang, namun masih sedikit penerapannya di lingkungan akademik Pembuatan *chatbot* inimenggunakan RASA sebagai *chatbot development* platformnya.Dengan mengambil data dari BAA dalam lingkup universitas dan LAA FTE dalam lingkup fakultas, pengguna dapat memberikan 29 topik pertanyaan mengenai informasi seputar layanan akademik. *Chatbot* layanan akademik ini melewati pengujian alfa dengan akurasi sebesar 100%, pengujian akurasisistem yang dilakukan oleh 5 pengguna sebesar 90%, dan pengujian beta yang didapatkan dari 30 responden sebesar 93% memilih setuju dan sangat setuju. Hasil ini didapatkan dengan melakukan pengujian apakah *chatbot* dapat memberikanjawaban yang sesuai dengan yang diajukan pengguna. Selain pengujian *chatbot*, pengujian kuesioner juga dilakukan kepada 30 responden untuk mengetahui apakah data yang digunakan dapat dipertanggung jawabkan atau tidak. Pengujian tersebut adalah uji reliabilitas kuesioner dengan hasil 0.955 (sangat tinggi) dan uji validitas kuesioner dengan hasil kuesioner yang digunakan pada tugas akhir ini dinyatakan valid. Dengan hasilpengujian yang sudah dilakukan, dapat diambil kesimpulan bahwa sistem ini dapat berjalan sesuai dengan tujuannya.

Kata kunci— Chatbot, RASA, Layanan Akademik.

Abstract* — Chatbots are software programmes created to engage via speaking or writing with users. Chatbots have been employed in various disciplines, although there is currently minimal use in the academic context. The building of this chatbot employs RASA as its chatbot development platform. By using data from BAA within the university scope and LAA FTE insidethe faculty, users may submit 28 topics of inquiries for information about the academic services they seek. This academic service chatbot completed alpha testing with 100% accuracy, 90% system accuracy testing by 5 users, and beta testing with 30 responses, of whom 93% picked agree and highly agree. These findings are achieved by determining if the chatbot can deliver replies that are consistent with the information provided by the user. In addition to chatbot testing, 30 respondents were subjected to questionnaire testing to see if the data utilised could be justified. The test is a questionnaire reliability test with a score of 0.955 (excellent) and a questionnaire validity test with the findings of the questionnaire utilized in this final project deemed legitimate. Based on the results of the testing, it is possible to infer that this system can function as intended.*Keywords — Chatbot, RASA, Layanan Akademik.***

I. PENDAHULUAN

Agen percakapan atau *chatbot* adalah program perangkat lunak yang berinteraksi melalui kata-kata tertulis atau lisan dengan orang atau sekelompok orang. *Chatbot* dapat ditemukan saat ini di banyak lingkungan seperti halaman web, aplikasi dan di perangkat 'pintar'. Perangkat 'pintar' tersebut meliputi ponsel pintar, TV pintar, jam tangan pintar, dan bahkan cincin pintar [1].

Chatbot telah menjadi alat yang sangat berguna di banyak domain dan aplikasi, dengan asisten virtual yang dibuat oleh Google, Apple (Siri) atau Microsoft (Cortana), baik dalam bentuk teks atau suara. Pada dasarnya, *chatbot* adalah agen otonom yang menggunakan banyak teknik kecerdasan buatan, seperti pemrosesan bahasa alami, pengenalan ucapan otomatis, penambangan data, dan pembelajaran mesin[2].

Sebagian besar dari *chatbot* melewati tes yang digunakan untuk melihat kepintaran mesin tersebut. *Chatbot* mampu menjawab hampir semua pertanyaan yang diberikan dengan lancar, dan mampu mengajukan pertanyaan lain kepada pengguna [2]. *Chatbot* menggunakan pemrosesan bahasa alami untuk memahami percakapan dan untuk mempermudah penggunaan, *chatbot* dapat digunakan dalam aplikasi pesan yang digunakan oleh banyak orang.

Chatbot sudah banyak digunakan di berbagai bidang seperti kesehatan, pemasaran, hiburan, dan juga ada yang menggunakan chatbot dalam lingkup akademik [3]. Implementasi bot dalam bidang yang sudah disebutkan dapat terbukti bahwa *chabot* dapat membantu pengguna.

Chatbot yang dikembangkan pada tugas akhir ini berinteraksi (dalam mode teks) dengan mahasiswa/I Universitas Telkom dengan menggunakan bahasa Indonesia. Untuk mengetahui kebutuhan yang diperlukan dalam pengembangan *chatbot* ini, maka survei awal disebarkan kepada mahasiswa/I Universitas Telkom. Tujuan utama pembuatan *chatbot* ini adalah agar pengguna mendapatkan informasi yang mereka inginkan dengan cepat tanpa perlu mencari informasi tersebut di halaman *website*, ataupun media sosial layanan akademik. Pada *chatbot* ini terdapat 30 topik pertanyaan yang dapat diajukan oleh pengguna.

Setelah melihat dari beberapa referensi riset ilmiah, selain itu dilakukan survei dan dari hasil tersebut terlihat bahwa mahasiswa/I Universitas Telkom membutuhkan *chatbot* yang dapat membantu dalam bidang layanan akademik. *Chatbot* ini dibuat menggunakan RASA sebagai chatbot development platformnya. RASA merupakan *open-source machine learning framework* untuk percakapan secara teks atau lisan. Dengan adanya *chatbot* ini pengguna dapat menemukan informasi yang mereka inginkan dengan mudah.

II. KAJIAN TEORI

A. *Chatbot*

Chatbot adalah pesan instan yang menyediakan layanan percakapan yang efisien kepada pengguna. *Chatbot* dilengkapi dengan kecerdasan buatan dan pemrosesan bahasa alami yang memungkinkan mereka menjadi sangat cerdas dan mampu menjawab pertanyaan yang diajukan manusia. *Chatbot* tersebut dibuat dengan sebuah kata kunci yang dapat dipahami oleh mesin untuk mendapatkan jawaban yang seharusnya di dapatkan. Untuk beberapa individu, *chatbot* hanyalah metode untuk terhubung dengan mesin intelijen. Namun, perlu dibangun kesadaran bahwa tidak ada yang artifisial tentang AI; karena itu, terdiri dari pembelajaran mesin dan algoritma *deep learning* yang dibangun. Sebuah mesin belum bisa mencapai tingkat kecerdasan di mana mereka dapat berpikir sama dengan manusia. Apa yang dilakukan sistem adalah hasil dari metode *training* yang dilakukan [4].

B. NLP (*Natural Language Processing*)

Kecerdasan buatan telah melahirkan atau membuatcabang baru dalam ilmu komputer dan teknik yang disebut dengan NLP (*Natural Language Processing*). Tujuan dari NLP adalah untuk membuat dan membangun aplikasi yang memungkinkan manusia untuk berkomunikasi dengan mesin dan perangkat lain dengan lebih mudah menggunakan bahasa alami. Hal ini adalah salah satu bagian terpenting dari kecerdasan buatan, maka dari itu digunakanlah sebuah skenario seperti penerjemah, pengenalan suara, dan lain-lain. Untuk memahami bahasa manusia, komputer harus memecah sebuah teks menjadi paragraf, frasa, dan kata-kata individual. Selain itu, komputer harus memahami bagaimana untuk mendeteksi hubungan antar satu kata dengan kata yang lain, dan memahami kalimat dalam konteks yang berbeda-beda. NLP terdiri dari beberapa aspek yang mendasar, aspek tersebut termasuk fonetik, morfologi, sintaksis, semantik, dan pragmatika [5].

Fonetik, morfologi, sintaksis, semantik, dan pragmatika merupakan cabang ilmu lingustik dari cabang linguistik mikro. Fonetik adalah studi untuk mempelajari bunyi bahasa tanpa memperhatikan fungsinya sebagai faktor pembeda makna. Morfologi adalah studi untuk mempelajari kata-kata dan bagaimana mereka terbentuk. Sintaksis adalah studi tentang hubungan antara kata dan kalimat. Semantik adalah studi untuk mengeksplorasi bagaimana bahasa dapat dipahami. Pragmatik adalah studi untuk menganalisis makna kata dalam kaitannya dengan lingkungannya [6].

Tujuan penelitian linguistik adalah untuk dapat mendefinisikan dan menjelaskan kebanyakan pengamatan linguistik di sekitar kita dalam percakapan, tulisan, dan media lainnya. Sebagian dari ini berkaitan dengan ukuran kognitif tentang bagaimana orang belajar, menghasilkan, dan memahami bahasa, sebagian dengan memahami hubungan antara ujaran linguistik dan lingkungan, dan sebagian lagi dengan memahami struktur linguistik yang melaluinya bahasa berkomunikasi. Kemudian, dengan berkonsentrasi pada inferensi dengan menggunakan pendekatan statistik dalam pemrosesan bahasa alami. Tujuan dari NLP statistik adalah untuk menerapkan inferensi statistik ke ranah bahasa alami. Secara umum, inferensi statistik mengumpulkan beberapa data (menurut distribusi probabilitas yang tidak diketahui) dan kemudian menarik kesimpulan tentang distribusi ini [7].

C. RASA.AI

Ada beberapa opsi ketika ingin membuat *chatbot*. Solusi ini dapat diklasifikasikan ke dalam dua kategori: *closed source* dan *open source*. Kelemahan dari sistem *closed source* adalah memiliki biaya tinggi, vendor *lock-in*, bahaya apabila terdapat kebocoran data, dan kesulitan untuk mengembangkan fitur baru. Berbeda dengan *closed source, open source* memiliki kelemahan sebagai berikut tidak adanya dukungan dalam pendanaan dan memiliki tampilan antar muka yang berbeda. Rasa terdiri dari dua komponen utama: rasa dan perangkat pengembangan perangkat lunak Rasa (Rasa SDK). Di dalam Rasa terdapat NLU dan Core. Rasa NLU dapat mengubah *input* pengguna menjadi *intents* dan entitas. Istilah untuk ini adalah NLU. Sedangkan, Rasa Core menentukan tindakan selanjutnya berdasarkan rekaman percakapan yang sudah dilakukan (termasuk *output* dari Rasa NLU) [8]. Pesan diproses dalam dua tahap: modul Core memeriksa dialog, dan modul NLU menyediakan layanan bahasa utama. Dengan dukungan pelacak inti RASA,platform RASA mengelola panggilan API eksternal dan hubungan CRM [9].

D. GPT-2

Model GPT-2 sudah termasuk ke dalam kumpulan model transformasi dalam berbagai bahasa. Ini menunjukkan bahwa GPT-2 secara eksklusif diajarkan pada teks mentah, tanpa label yang ditambahkan oleh pengguna, dan menggunakan teknik otomatis untuk menghasilkan *input* dan label berdasarkan teks-teks yang diberikan ketika telah menyelesaikan pelatihannya. Tujuan dari pembuatan model ini salah satunya mempermudah pengguna untuk memprediksi kata atau kalimat yang berikutnya akan muncul [14]. Model bahasa ini telah dilatih menggunakan kumpulan data yang luas dan terdiversifikasi dapat berkinerja baik di berbagai domain dan kumpulan data. GPT-2 mampu bekerja hingga 7 dari 8 set data permodelan. Kemampuan ini digunakan untuk melakukan berbagai tugas dalam situasi *zero-shot* yang menunjukkan bahwa model berkapasitas tinggi yang dilatih untuk mengoptimalkan kemampuan yang beragam. Dimulai dengan belajar bagaimana menyelesaikan sejumlah tugas tanpa pengawasan eksplisit [15]

GPT-2, hanya di *training* untuk mengantisipasi kata berikutnya dalam 40GB teks. GPT-2 adalah model bahasa berbasis transformator besar dengan 1,5 miliar parameter, dilatih pada 8 juta *dataset* halaman web. GPT-2 dilatih dengan tujuan: memprediksi kata berikutnya dalam teks yang diberikan kata-kata sebelumnya. Tujuan dasar ini memiliki contoh yang terjadi secara alami dari beberapa tugas di berbagai disiplin ilmu karena beragamnya kumpulan data.

GPT-2 adalah peningkatan langsung dari GPT, dengan lebihdari 10 kali lebih banyak parameter dan dilatih di lebih dari 10 kali lebih banyak data [16].

E. Ngrok

Ngrok adalah layanan yang mirip dengan Bitnami ProcessMaker. Layanan ini memungkinkan untuk membuat server web atau aplikasi lokal yang dapat beroperasi di komputer lokal. Dengan menggunakan ngrok, pengguna dapat mengonfigurasi aplikasi seluler untuk mengakses semua proses yang telah dibuat di ProcessMaker lokal [17].

F. Telegram

Telegram adalah layanan pesan yang berfokus pada kecepatan dan keamanan. Telegram dapat digunakan di semua perangkat secara bersamaan, pesan akan disinkronkan dengan seluruh perangkat yang tersambung. Telegrammemiliki lebih dari 500 juta pengguna bulanan aktif dan salah satu dari sepuluh

aplikasi yang paling banyak diunduh di dunia. Telegram dapat membagikan pesan, gambar, video, dan data dalam bentuk apa saja (pdf, mp3, zip, doc, ppt, dan lain-lain). Telegram memiliki bot yang beroperasi langsung di dalam telegram. Pengguna dapat berkomunikasi dengan bot melalui pesan, instruksi dan *inline request*. Dengan menggunakan panggilan HTTPS ke Bot API Telegram, pengguna data mengontrol bot yang dibuat [18].

G. BotFather

BotFather adalah sebuah bot yang dibuat apabila pengguna ini membuat bot baru dan melakukan perubahan terhadap bot yang sudah ada. Telegram menyimpan BotFather dengan id '@botfather' yang dapat ditemukan dengan mudah oleh pengguna. Untuk membuat bot baru, gunakan perintah '/newbot'. BotFather akan meminta nama bot dan nama pengguna sebelum membuat token otentikasi. Nama bot akan dipublikasikan di informasi kontak, dan nama pengguna adalah pengidentifikasi singkat yang digunakan dalam komentar dan tautan t.me. Nama pengguna terdiri dari 5-32 karakter dan dapat menggunakan huruf besar/kecil; namun, hanya karakter latin, angka, dan garis bawah yang diizinkan. Nama pengguna bot Anda harus diakhiri dengan 'bot', seperti 'aera bot' atau 'AeraBot'. Token adalah string yang mirip dengan 110201543:AAHdqTcvCH1vGWJxfSeofSA0P5PALD yang diperlukan untuk mengotorisasi bot dan mengirimkan permintaan ke Bot API. Jaga agar token tetap aman dan terlindungi, karena token dapat digunakan untuk mengoperasikan bot oleh siapa saja. [19].

H. Usability Testing

Usability berasal dari kata ussable yang memiliki arti dapat digunakan dengan baik. Terdapat cara yang dapat digunakan untuk menentukan di mana sistem dapat berjalan dengan baik atau tidak, serta menentukan penyesuaian apa yang harus dilakukan agar sistem atau produk dapat digunakan dengan lancar [20]. Cara tersebut adalah dengan memberikan pertanyaan apakah sistem cukup baik untuk memenuhi kebutuhan pengguna. Menurut Jacob Nielson, *usability testing* diukur dengan lima kriteria, yaitu: [21]

1. *Learnability*, mengacu kepada kemudahan dan kecepatan sistem dapat dipelajari oleh pengguna untuk pertama kali.
2. *Efficiency of Use*, berperan setelah pengguna menguasai sistem dan berhubungan dengan kemudahan dan kecepatan penggunaan sistem untuk mendapatkan tujuan awal dari penggunaan sistem tersebut.
3. *Memorability*, mengacu kepada kemudahan pengguna dapat kembali ke sistem setelah jangka waktu tertentu dan masih ingat bagaimana menggunakannya.
4. *Error rates*, mengacu kepada kesalahan yang dilakukan pengguna dan cara yang dilakukan pengguna untuk mengatasinya.
5. *Satisfacation*, mengacu kepada kriteria subjektif pengguna mengenai kepuasan dalam penggunaan sistem.

Teknik pengumpulan data dilakukan dengan menggunakan kuesioner untuk memperoleh data responden. Penghitungan ini menggunakan *skala likert*. Untuk mengukur bagaimana persepsi atau pendapat individual atau kelompok terhadap sebuah peristiwa atau fenomena sosial, digunakan *skala likert.* Berikut adalah rumus yang digunakan dalam *skala likert* [22]:

$$\text{Total skor} = T \times Pn \quad (2.1)$$

$$\text{Nilai Indeks} = \frac{Total\ skor}{Y} \times 100\% \quad (2.2)$$

$$\text{Rumus Interval} = \frac{100}{Jumlahskor(likert)} \quad (2.3)$$

Keterangan:
T = Total responden yang memilih
Pn = Pilihan angka skor likert
Y = Nilai likert tertinggi × jumlah responden

I. Uji Validitas

Uji Validitas adalah uji yang digunakan sebagai alat ukur dan biasanya digunakan untuk mengukur apakah sistem tersebut valid atau tidak. Sebuah instrumen dapat dikatakan valid apabila dapat memberikan data dari variabel secara tepat sesuai dengan keadaan yang sebenarnya [23]. Sebuah penelitian dapat dikatakan valid apabila [24]:

1. Memiliki koefisien korelasi *product moment* yang melebihi 0.3
2. Memiliki koefisien korelasi *produce moment* yang lebih besar besar '>' dari r-tabel $(\alpha; Df = n - 2)$
3. Memiliki nilai $sig \leq \alpha$

Teknik uji validitas yang digunakan adalah teknik korelasi *product moment*, dengan rumus:

$$r_{xy} = \frac{n(\sum xy) - (\sum x \times \sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}} \quad (2.4)$$

Keterangan:

r_{xy}= Korelasi product moment
n = Jumlah responden
x = Skor variabel (jawaban responden)
y = Skor total dari variabel untuk responden ke-n

J. *Uji Realibilitas*

Uji reliabilitas adalah pengujian yang menentukan apakah hasil pengukuran telah bebas dari kesalahan pengukuran dan dapat diandalkan [23]. Sedangkan uji reliabilitas instrumen adalah pengujian untuk mengetahui apakah data yang dimiliki dapat dipercaya atau tidak. Metode uji reliabilitas yang digunakan adalah Cronbach's alpha. Pengujian ini dilakukan dengan cara membandingkan nilai Cronbach's alpha dengan tingkat yang digunakan. Tingkat signifikan yang digunakan adalah 0.5, 0.6, dan 0.7. Kriteria

yang digunakan untuk melakukan pengujian adalah sebagai berikut [25]:

1. Instrumen dapat dikatakan reliabel apabila nilai Cronbach's alpha lebih besar '>' tingkat signifikan.
2. Instrumen dapat dikatakan tidak reliabel apabila nilai Cronbach's alpha lebih besar '<' tingkat signifikan.

Rumus yang digunakan untuk menentukan reliabilitas instrumen dengan menggunakan metode Cronbach's alpha adalah sebagai berikut:

$$r_{11} = [\frac{k}{k-1}][1 - \frac{\sum \sigma_b^2}{\sigma_t^2}] \quad (2.5)$$

Keterangan:

r_{11} = Koefisien reliabilitas instrumen
k = Jumlah butir pertanyaan
$\sum \sigma_b^2$ = Jumlah varian butir
σ_t^2 = Varian total

Secara umum, para ahli telah menggambarkan koefisien reliabilitas sebagai sangat tinggi, tinggi, sedang, rendah, dan sangat rendah. Kategori tersebut berdasarkan koefisien yang telah dihitung dengan menggunakan rumus yang dibuat oleh beberapa ahli, seperti rumus Kuder-Richardson 20 (KR-20), Kuder-Richardson 21 (KR-21), rumus Hoyt, dan sebagainya. Guildford mengategorikan koefisien reliabilitas sebagai berikut[26] :

TABEL 2.1
KATEGORI KOEFISIEN RELIABILITAS

Interval	Tingkat koefisien reliabilitas
0.80 – 1.00	Sangat tinggi
0.60 – 0.79	Tinggi
0.40 – 0.59	Sedang
0.20 – 0.39	Rendah
0.00 – 0.19	Sangat rendah (tidak reliabel)

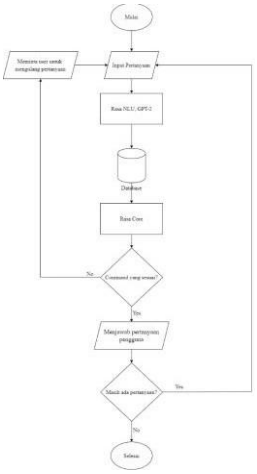
III. METODE

A. Gambaran Umum Sistem

Gambaran umum pengembangan *chatbot* dibangun dengan basis pengetahuan yang mencerminkan domain *chatbot*. Domain tersebut diperlakukan sebagai bagian dari lingkup *chatbot*. Dengan membuat sebuah data yang berisikan *intent, template*, ataupun sebuah jalan cerita bagaimana sistem ini berjalan. Pengguna dapat menggunakan *chatbot* yang memberikan jawaban seputar informasi yang didapatkan dari layanan akademik. *Chatbot* ini menggunakan Bahasa Indonesia. *Chabot* dapat diakses melalui telegram sebagai salah satu alternatif antarmuka pengguna atau *user interface*.

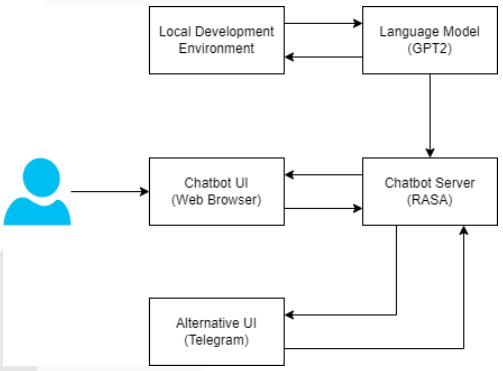
B. Perancangan Sistem

Berikut adalah *flowchart* sistem percakapan yang digunakan dalam *chatbot* ini



GAMBAR 3.1
FLOWCHART SISTEM PERCAKAPAN

Pada gambar 3.1 dapat dilihat bahwa pengguna akan memasukkan pertanyaan yang diinginkan dan dilanjutkan dengan mengirimkan pertanyaan kepada Rasa NLU dengan bantuan GPT-2 untuk mencari tahu '*intent'* atau maksud pertanyaan pengguna. Rasa NLU akan mengirimkan ke Rasa core akan menentukan keputusan berdasarkan hasil yang sudah diberikan. Apabila tidak ditemukannya kata kunci yang sesuai, *chatbot* akan meminta kepada pengguna untuk mengulang pertanyaan yang ingin diajukan. Apabila ditemukan kata kunci yang sesuai, *chatbot* akan memberikan jawaban sesuai dengan pertanyaan pengguna. Apabila pengguna ingin kembali menggunakan *chatbot*, pengguna dapat langsung mengajukan pertanyaan yang diinginkan. Apabila tidak, maka pengguna dapat mengucapkan selamat tinggal untuk mengakhirinya.



GAMBAR 3.2
IMPLEMENTASI SISTEM

Pada gambar 3.2 dapat dilihat bahwa pada pengembangan *chatbot* ini Rasa digunakan sebagai dan menggunakan GPT2 sebagai model bahasa yang akan digunakan. Pada RASA terdapat proses yang berlangsung, proses tersebut seperti *Inisialisasi*, konfigurasi *framework*, NLU, respons *chatbot*, data dialog, pelatihan, dan yang terakhir pengujian. Setelah memastikan bahwa *chatbot* dapat berjalan dengan baik, tahap selanjutnya menyambungkan RASA dengan alternatif antar muka pengguna yaitu Telegram untuk mempermudah penggunaannya.

IV. HASIL DAN PEMBAHASAN

A. Implementasi Antar Muka

1. Tampilan profil *chatbot*

GAMBAR 4.1 TAMPILAN PROFIL *CHATBOT* DI TELEGRAM.

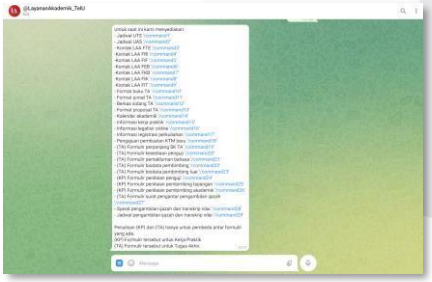
Tampilan berikut merupakan profil *chatbot* di dalam Telegram. Pada gambar 4.1 ini terlihat foto profil *chatbot*, nama pengguna, dan deskripsi dari *chatbot* yang sudah dibuat.

2. Tampilan *chatbot* dalam aplikasi Telegram

GAMBAR 4.2
TAMPILAN CHATBOT DI TELEGRAM.

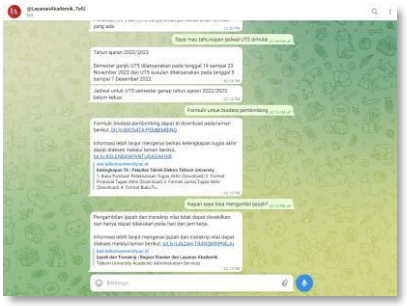
Gambar 4.2 merupakan halaman *chatbot* yang menggunakan Telegram sebagai alternatif antar muka. Halaman ini akan muncul apabila pengguna ingin menggunakan *chatbot*. Apabila pengguna menekan tombol ‘start’ maka *chatbot* dapat digunakan

3. Tampilan *chatbot* saat digunakan pengguna



GAMBAR 4.3
TAMPILAN CHATBOT DI TELEGRAM.

Gambar 4.3 merupakan tampilan ketika pengguna sedang berinteraksi dengan *chatbot*. Apabila pengguna sudah pernah menggunakan *chatbot*, maka pengguna dapat langsung menanyakan informasi yang dia inginkan. Namun, apabila pengguna tidak pernah menggunakan *chatbot*, maka pengguna dapat menekan atau mengetik ‘/list agar *chatbot* dapat memberikan daftar pertanyaan yang dapat diajukan oleh pengguna. Apabila pengguna memberikan pertanyaan seusai dengan daftar, maka *chatbot* akan menampilkan jawaban yang sesuai. Gambar 4.4 merupakan tampilan ketika *chatbot* memberikan jawaban sesuai dengan pertanyaan yang diajukan.



GAMBAR 4.4
MEMBERIKAN PERTANYAAN YANG SESUAI

Apabila pengguna memberikan pertanyaan di luar dari daftar yang sudah diberikan, maka chatbot akan mengeluarkan perintah untuk mengulang pertanyaan menyarankan pengguna untuk melihat daftar pertanyaan yang ada. Gambar 4.5 merupakan tampilan ketika pengguna memberikan pertanyaan yang tidak sesuai dengan daftar pertanyaan yang sudah diberikan.

GAMBAR 4.5
MEMBERIKAN PERTANYAAN YANG TIDAK SESUAI

B. Pengujian Alfa

Pengujian alfa merupakan pengujian yang dilakukan untuk mengetahui apakah sistem yang sudah dibuat berjalan dengan baik atau tidak secara internal. Pengujian alfa dilakukan dengan 2 aspek. Aspek tersebut adalah aspek fungsionalitas dan aspek dalam penanganan masalah. Dengan adanya pengujian ini, maka pembuat dapat segera memperbaiki kesalahan yang ada dengan cepat. Menggunakan metode *black box texting* untuk memperhatikan ketepatan respons dalam memberikan jawaban pada *chatbot*. Pengujian alfa dilakukan oleh penulis pada tanggal 8 sampai 10 Juli 2022. Pengujian alfa dilakukan secara daring. Berdasarkan hasil pengujian tersebut, 100% pengujian berhasil diterima.

$$\text{Uji alfa} = \frac{\text{Total berhasil}}{\text{Total pengujian}} \times 100\%$$

$$\text{Uji alfa} = \frac{35}{35} \times 100\%$$

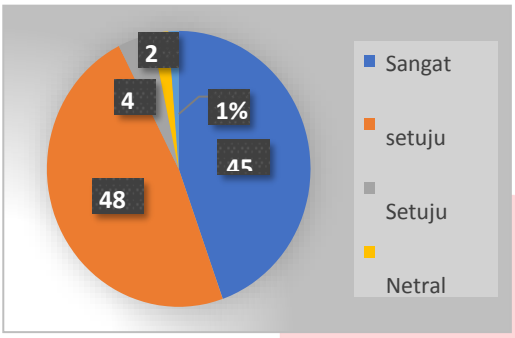
$$\text{Uji alfa} = 100\%$$

C. Pengujian Beta

Pengujian beta merupakan pengujian yang dilakukan oleh pengguna. Pengguna yang terlibat pada pengujian ini adalah alumni dan mahasiswa/I fakultas teknik elektro di Universitas Telkom. Pengujian tersebut menggunakan kuesioner mengenai beberapa pertanyaan seputar sistem

chatbot yang kemudian akan dihitung dengan *skala likert*. *Skala likert* digunakan karena mudah untuk menyusun pernyataan dan membandingkan skor yang lebih tinggi dan lebih rendah.

Pengujian beta dilakukan oleh 30 responden yang merupakan alumni dan mahasiswa/I fakultas teknik elektro di Universitas Telkom pada tanggal 12 sampai 16 Juli 2022. Pengujian beta dilakukan secara daring. Hasil dari jawaban responden untuk keseluruhan pertanyaan dapat dilihat di gambar 4.6.



GAMBAR 4.6
HASIL PENGUJIAN BETA, KESELURUHAN

Pada gambar 4.6 terlihat bahwa dari keseluruhan pertanyaan, didapatkan hasil 45% memilih sangat setuju dan 48% setuju.

D. *Usability Testing*

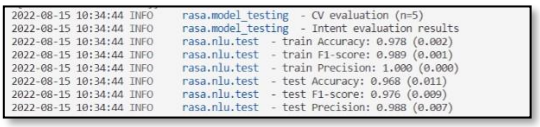
Usability Testing dilakukan apabila sudah mendapatkan respons mengenai beberapa pertanyaan yang diajukan dalam kuesioner. Menggunakan data yang didapatkan dari 30 responden yang merupakan alumni dan mahasiswa/I fakultas teknik elektro di Universitas Telkom. Berikut adalah perhitungan *usability testing* pada *chatbot* layanan akademik.

TABEL 4.1
USABILITY TESTING

Pertanyaan	*Usability testing*					Nilai	Ket
	LR	EF	MR	ER	SF		
Saya akan sering menggunakan *chatbot ini*			✓		✓	80.7%	SS
Saya dapat mengakses *chatbot* ini dengan mudah	✓	✓	✓	✓	✓	87.3%	SS
Memilliki fitur yang dapat mempermudah dalam penggunaan *chatbot*	✓	✓		✓	✓	85.3%	SS
Saya mendapatkan informasi yang saya inginkan memalui *chatbot* dengan mudah	✓	✓		✓	✓	86.7%	SS
Saya merasa *chatbot* ini praktis ketika digunakan							
Apakah *chatbot* memberikan jawaban sesuai dengan pertanyaan yang diajukan		✓		✓	✓	87.3%	SS
Saya merasa *chatbot* ini	✓	✓	✓		✓	87.3%	SS

E. Pengujian Sistem *Chatbot*

Apabila sudah membuat modifikasi besar pada data pelatihan NLU dengan memecah satu *intent* menjadi dua *intent* atau menambahkan sejumlah besar contoh pelatihan. Tinjauan validasi silang terhadap NLU harus dilakukan. Validasi silang menghasilkan beberapa pembagian pengujian dan rata-rata temuan penilaian untuk setiap pemisahan pengujian secara otomatis. Ini menyiratkan bahwa data yang dimiliki dinilai selama validasi silang, menjadikannya metode paling lengkap untuk menguji model NLU secara otomatis. Untuk pengujian sistem *chatbot* didapatkan hasil sebagai berikut:



GAMBAR 4.7
HASIL PENGUJIAN SISTEM *CHABOT*

Pada gambar 4.7 terlihat bahwa pengujian sistem yang dilakukan oleh rasa mendapatkan nilai akurasi sebesar 97% dan 96%. Sedangkan, nilai presisi sistem ini adalah 100% dan 98%. Sehingga, dapat dikatakan bahwa sistem pada *chatbot* ini berjalan dengan baik.

F. Perform *Chatbot*

Pengujian performa *chatbot* dilakukan dengan cara melihat berapa rata-rata kecepatan yang diperlukan untuk menjawab pertanyaan pengguna. Rumus yang digunakan untuk mengetahui rata-rata kecepatan tersebut adalah sebagai berikut:

$$\text{Performa chatbot} = \frac{\text{Total Waktu}}{\text{Jumlah Pertanyaan}}$$

$$\text{Performa chatbot} = \frac{193899.36}{320} = 605.9355$$

Berdasarkan hasil penghitungan tersebut, dapat diambil kesimpulan bahwa *chabot* ini dapat merespons pertanyaan pengguna dengan rata-rata 605.9355 mili detik.

V. KESIMPULAN

Berdasarkan hasil pengujian dan pengamatan yang telah dibahas pada bab sebelumnya, maka dapat disimpulkan bahwa:

1. Berdasarkan dengan hasil pengujian alfa dengan akurasi sebesar 100%, dan pengujian beta yang

didapatkan dari 30 responden sebesar 93% memilih setuju dan sangat setuju. *Chatbot* layanan akademik ini dapat berjalan dengan lancar sesuai dengan tujuan awalnya yaitu membantu mahasiswa/I yang sedang mencari informasi seputar layanan akademik.

2. Berdasarkan dengan hasil pengujian validitas bahwa seluruh pertanyaan kuesioner memiliki nilai $r_{xy} > r_{tabel}$ dan pengujian reliabilitas dengan hasil 0.955 yang menunjukkan bahwa kuesioner yang digunakan pada tugas akhir ini dapat dinyatakan valid dan sangat reliabel.
3. Dengan *dataset* yang terdiri dari 36 *rules*, 35 *intent*, 36 respons dan 1,485 pertanyaan yang sudah disusun. Pengujian akurasi sistem yang dilakukan oleh 5 pengguna mendapatkan hasil sebesar 90%.

REFERENSI

[1] Panayiotis Zaphiris, "Learning and Collaboration Technologies Games and Virtual Environments for Learning", doi: https://doi.org/10.1007/978-3-030-77943-6.

[2] Agrawal Ritu and Wadhwa Mani, "Review of State of the Art Design Techniques for Chatbots", doi: 10.1007/s42979-020-00255-3.

[3] Arcangelo Castiglione, Florin Pop, Massimo Ficco, and Francesco Palmieri, "Cyberspace Safety and Security," 10th International Symposium, CSS 2018, Amalfi, Italy, October 29–31, 2018, Proceedings, Sep. 2018.

[4] S. Raj, Building Chatbots with Python. Apress, 2019. doi: 10.1007/978-1-4842-4096-0.

[5] A. Jiao, "An Intelligent Chatbot System Based on Entity Extraction Using RASA NLU and Neural Network," in Journal of Physics: Conference Series, Apr. 2020, vol. 1487, no. 1. doi: 10.1088/1742-6596/1487/1/012014.

[6] Unsiah Frida and Yuliati Ria, "Pengantar Ilmu Linguistik," Universitas Brawijaya Press, pp. 2–5, Jan. 2018.

[7] J. Brownlee Disclaimer, "Deep Learning for Natural Language Processing Develop Deep Learning Models for Natural Language in Python Acknowledgements Copyright Deep Learning for Natural Language Processing," 2017.

[8] X. Kong and G. Wang, Conversational AI with Rasa : build, automate, and deploy AI-powered text and voice-based assistants and chatbots.

[9] A. Singh, K. Ramasubramanian, and S. Shivam, Building an Enterprise Chatbot. Apress, 2019. doi: 10.1007/978-1-4842-5034-1.

[10] GitHub CI, "Writing Conversation Data," RASA, Jul. 21, 2022. https://rasa.com/docs/rasa/writing-stories/ (accessed Jun. 04, 2022).

[11] GitHub CI, "NLU Training Data," RASA, Jul. 21, 2022. https://rasa.com/docs/rasa/nlu-training-data/(accessed Jun. 04, 2022).

[12] Github CI, "Model Configuration," RASA, 2022. https://rasa.com/docs/rasa/model-configuration/ (accessed Jun. 04, 2022).

[13] GitHub CI, "Domain," RASA, 2022. https://rasa.com/docs/rasa/domain/ (accessed Jun. 04, 2022).

[14] Patrick von Platen, "GPT-2," huggingface. https://huggingface.co/gpt2 (accessed Jun. 04, 2022).

[15] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners." [Online]. Available: https://github.com/codelucas/newspaper

[16] Alec Radford et al., "Better Language Models and Their Implications," OPEN AI, 2019.

[17] D. Majekodunmi, Business process automation with ProcessMaker 3.1 : A Beginner's Guide.

[18] "Telegram FAQ," Telegram. https://telegram.org/faq (accessed Jun. 04, 2022).

[19] "BotFather," Telegram. https://core.telegram.org/bots#6-botfather (accessed Jun. 04, 2022).

[20] Jeffery Rubin and Dana Chisnell, "Handbook of Usability Testing," Wiley, May 2008.

[21] J. Nielsen, "Usability Engineering," Elsavier Science, 1994.

[22] Lailatus Sa'adah, "Metode Penelitian Ekonomi dan Bisnis," LPPM Universitas KH. A. Wahab Hasbullah, pp. 101–104, 2021.

[23] Suharsimi Arikunto, "Prosedur penelitian: suatu pendekatan praktik," Rineka Cipta, 2010.

[24] Ferry and D. Deddy Prasetya Kristadi, "Metode Kuantitatif Pengambilan Keputusan Mengukur Kepuasan Pengguna Web Pada Perguruan Tinggi," Penerbit Lakeisha, Mar. 2021.

[25] Budi Darma, "STATISTIKA PENELITIAN MENGGUNAKAN SPSS (Uji Validitas, Uji Reliabilitas, Regresi Linier Sederhana, Regresi Linier Berganda, Uji t, Uji F, R2)," Guepedia, 2021.

[26] Dr. Sumardi, "Teknik Pengukuran Dan Penilaian Hasil Belajar," Deepublish, Dec. 2020.