

5. PENGUJIAN

5.1. Pengujian Model Pre-Trained Word2Vec

Pengujian ini dilakukan untuk menemukan model Word2Vec dengan hasil intent understanding terbaik, dari dataset wikidump. Pengujian dilakukan untuk menentukan dimensi vector dan metode untuk menghasilkan embedding secara optimal dan akurat. Pengujian dilakukan dengan mengganti dimensi vector menggunakan beberapa metode berbeda.

Pengujian dilakukan dengan melakukan tes analogi dan tes kedekatan similarity dari 2 kata. Tes analogy dilakukan dengan melakukan analogi laki : raja, maka perempuan : ?. Terdapat tiga pasang kata yang digunakan sebagai metrik pengukuran untuk tes kedekatan similarity, karena kata-kata tersebut memiliki nilai semantik yang serupa. Hasil similarity 1 didapatkan dengan membandingkan antara kata 'belajar' dan 'membaca'. Sedangkan hasil similarity 2 didapatkan dengan membandingkan antara kata 'submit' dan 'kumpul'. Hasil similarity 3 didapatkan dengan membandingkan antara kata 'dosen' dan 'dosbing'. Penghitungan similaritynya sendiri dilakukan menggunakan cosine similarity dan dijalankan dengan function `wv.similarity` yang disediakan oleh gensim.

5.1.1. Pengujian menggunakan Referensi Dieka Nugraha (2018)

Model referensi Dieka Nugraha menggunakan dataset latest pages article milik wikidump yang berukuran 923,9 MB dan berisi data artikel di wikidump berbahasa Indonesia.

Tabel 5.1 Hasil Pengujian Menggunakan Metode Dieka Nugraha

Pengujian	Nama model	Vector Size	Hasil analogy test	Hasil similarity test 1	Hasil similarity test 2	Hasil similarity test 3
1	idwiki_word2vec_100_new_lower	100	[('ratu', 0.71826159 95407104)]	0.5789078	-0.031480443	NULL
2	idwiki_word2vec_200_new_lower	200	[('ratu', 0.63475555 1815033)]	0.49985608	0.030349568	NULL

3	idwiki_word 2vec_300_n ew_lower	300	[('ratu', 0.57432264 08958435)]	0.4608853	-0.01972603	NULL
----------	---------------------------------------	-----	---------------------------------------	-----------	-------------	------

5.1.2. Pengujian menggunakan Referensi Eko Setiawan (2023)

Model referensi Eko Setiawan menggunakan dataset indonesian.pkl yang memiliki ukuran 138MB dan berisi artikel wikihow berbahasa indonesia. Parameter yang digunakan pada percobaan ini adalah 10 window, 5 min_count, dengan 10 epoch.

Tabel 5.2 Hasil Pengujian Menggunakan Metode Eko Setiawan

No	Nama model	Dimensi	Hasil analogy test	Hasil similarity test 1	Hasil similarity test 2	Hasil similarity test 3
1	word2vec-wiki how-indonesia .	100 vector size (default)	[('ratu', 0.7252589464 187622)]	0.5991431	0.023927497	NULL
2	word2vec-wiki how-indonesia _200	200 vector size	[('ratu', 0.6037102341 651917)]	0.50546986	0.0040785572	NULL
3	word2vec-wiki how-indonesia _300	300 vector size	[('ratu', 0.5442094802 856445)]	0.47535247	0.012273572	NULL

5.1.3. Pengujian menggunakan Metode Gabungan (Dieka Nugraha dan Eko Setiawan)

Model gabungan ini akan menggunakan dataset milik Dieka Nugraha, namun menggunakan preprocess data milik Eko Setiawan.

Tabel 5.3 Hasil Pengujian Menggunakan Metode Gabungan

No	Nama model	Parameter	Hasil analogy test	Hasil similarity test 1	Hasil similarity test 2	Hasil similarity test 3
1	w2v_try	10 epoch, 10 window, 5 min_count, 100 vector_size (default)	[('ratu', 0.6951261758 804321)]	0.56981796	0.13857403	NULL
2	w2v_try_30	30 epoch, 10 window, 5 min_count, 100 vector_size (default)	[('ratu', 0.7199980020 523071)]	0.5684663	0.058498345	NULL
3	w2v_try_50	50 epoch, 10 window, 5 min_count, 100 vector_size (default)	[('ratu', 0.6944590806 96106)]	0.5716902	0.11395024	NULL
4	w2v_try_300	10 epoch, 10 window, 15 min_count, 300 vector_size	[('ratu', 0.5818066596 984863)]	0.45038253	0.06493273	NULL
5	w2v_try_300_10	10 epoch, 10 window, 10 min_count, 300 vector_size	[('ratu', 0.5567870736 122131)]	0.45933282	0.015918497	NULL
6	w2v_try_v2.0	30 epoch, 10 window, 10 min_count, 300 vector_size	[('ratu', 0.56177812814 71252)]	0.46590608	0.055648055	NULL
7	w2v_try_v2.1	30 epoch, 10 window, 10 min_count, 100 vector_size	[('ratu', 0.6923102140 426636)]	0.5726459	0.070136376	NULL
8	w2v_try_v2.2	10 epoch, 10 window, 10 min_count, 100	[('ratu', 0.6870281696 31958)]	0.56892294	0.013394177	NULL

No	Nama model	Parameter	Hasil analogy test	Hasil similarity test 1	Hasil similarity test 2	Hasil similarity test 3
		vector_size				
9	w2v_try_v2.3	30 epoch, 10 window, 5 min_count, 100 vector_size	[('ratu', 0.6914869546890259)]	0.56835395	0.055110704	NULL

Berdasarkan percobaan yang dilakukan, dapat dilihat bahwa model w2v_try_v2.1.model, dengan parameter 30 epoch dan 100 vector size dengan 10 min_count dan 10 window menghasilkan hasil analogy dan similarity terbaik. Model inilah yang akan digunakan sebagai model dasar untuk fine-tuning.

Hasil similarity 3 menghasilkan *NULL* pada model di semua metode dikarenakan tidak ditemukan vector pada kata 'dosbing'. Vector kata 'dosbing' ini nantinya akan muncul jika model Word2Vec telah melalui fine-tuning sesuai dataset chat yang telah di preprocess sebelumnya.

5.2. Pengujian Fine-Tuning Word2Vec

Pengujian ini dilakukan untuk menentukan model Word2Vec yang telah disesuaikan dengan dataset chatbot dengan hasil terbaik diantara 2 metode *concatenation intent* yang dilakukan.

Pengujian dilakukan dengan mengganti jenis pre-trained word2vec yang digunakan, serta parameter yang oleh model. Perbandingan hasil akurasi dilakukan diantara penghitungan hasil similarity dari dua kata. Terdapat tiga pasang kata yang digunakan sebagai metrik pengukuran untuk tes kedekatan similarity, karena kata-kata tersebut memiliki nilai semantik yang serupa. Hasil similarity 1 didapatkan dengan membandingkan antara kata 'belajar' dan 'membaca'. Sedangkan hasil similarity 2 didapatkan dengan membandingkan antara kata 'submit' dan 'kumpul'. Hasil similarity 3 didapatkan dengan membandingkan antara kata 'dosen' dan 'dosbing'. Penghitungan similaritynya sendiri dilakukan menggunakan cosine similarity dan dijalankan dengan function `wv.similarity` yang disediakan oleh gensim.

5.2.1. Fine-Tuning Percobaan 1

Kondisi awal saat memulai pengujian percobaan 1 adalah dengan melakukan *training* menggunakan data berupa dictionary intent, yang di tiap intentnya berisi list of *pattern*(pertanyaan) untuk tiap intent. Program kemudian mengakses tiap pattern dengan melakukan iterasi secara bertingkat (*nested loop*) sesuai epoch yang telah ditentukan. Model word2vec kemudian akan diberikan *list of sentence* dari *pattern* pada *list of pattern*.

Pada percobaan pertama, concat akan dilakukan dengan menggabungkan tiap intent dengan semua pattern pertanyaan yang dimiliki oleh masing-masing intent tersebut. Berikut ini pengujian yang dilakukan menggunakan percobaan 1.

Tabel 5.4 Hasil Pengujian Fine Tuning Menggunakan Percobaan 1

Pengujian	Nama model	Pre-Trained Word2Vec	Parameter	Hasil Similarity 1	Hasil Similarity 2	Hasil Similarity 3
1	model_finetuned_1.0	w2v_try_v2.0	10 epoch, inner epoch 1 untuk tiap pattern, start alpha = 0.01, end alpha = 0.01, total_example per pattern = 1.	0.46590608	0.05564805	0.068741776
2	model_finetuned_2.0	w2v_try_v2.0	30 epoch, inner epoch 1 untuk tiap pattern, start alpha = 0.01, end alpha = 0.01, total_example per pattern = 1.	0.46590608	0.05564805	0.068741776
3	model_finetuned_2.1	w2v_try_v2.0	30 epoch, inner epoch 5 untuk tiap pattern, start alpha = 0.01, end alpha =	0.46590608	0.05564805	0.068741776

			0.001, total_example per pattern = 1.			
--	--	--	--	--	--	--

Note: Hasil pengujian jika model dilatih dengan menjalankan epoch secara langsung untuk semua dataset intent dan pattern, tanpa perlu mengakses pattern tiap intent satu per satu. Namun tetap diberikan dalam bentuk list of string (sentence).

Tabel 5.5 Lanjutan Hasil Pengujian ada Tabel 5.4 dengan Modifikasi Program

Pengujian	Nama model	Pre-Trained Word2Vec	Parameter	Hasil Similarit y 1	Hasil Similarit y 2	Hasil Similarit y 3
4	model_finetuned_2 .2	w2v_try_v2.0	30 epoch, start alpha = 0.01, end alpha = 0.01.	0.46590 608	0.05564 805	0.03760 023
5	model_finetuned_2 .3	w2v_try_v2.0	50 epoch, start alpha = 0.01, end alpha = 0.01.	0.46590 608	0.05564 805	0.03760 023
6	model_finetuned_2 .4	w2v_try_v2.0	50 epoch	0.46590 608	0.05564 805	0.03760 023

Note: Pengujian dilakukan dengan mengganti dataset yang digunakan menjadi list of word per pattern sentences sebagai input atau training data untuk model Word2Vec.

Tabel 5.6 Lanjutan Hasil Pengujian pada Tabel 5.5 dengan Modifikasi Dataset

Pen guji an	Nama model	Pre-Trained Word2Vec	Epoch	Hasil Similarity 1	Hasil Similarity 2	Hasil Similarity 3
7	model_finetuned_2.5	w2v_try_v2.0	50	0.46590608	0.13324256	0.48833928
8	model_finetuned_2.6	w2v_try_v2.0	30	0.46590608	0.11594536	0.47770867
9	model_finetuned_2.7	w2v_try_v2.0	10	0.46590608	0.11415142	0.476405

Note: Melakukan pengujian ulang dikarenakan terdapat perubahan pada dataset.

Tabel 5.7 Hasil Pengujian Menggunakan Dataset yang Telah Dimodifikasi

Pengu jian	Nama model	Pre-Trained Word2Vec	Epoch	Hasil Similarity 1	Hasil Similarity 2	Hasil Similarity 3
10	model_finetuned_2.8	w2v_try_v2.0	50	0.46590608	0.17647395	0.3968388
11	model_finetuned_2.9	w2v_try_v2.0	30	0.46590608	0.15491052	0.40215805
12	model_finetuned_2.10	w2v_try_v2.1	10	0.46590608	0.10538931	0.43066478
12	model_finetuned_2.11	w2v_try_v2.1	50	0.46590608	0.13036771	0.4845084
13	model_finetuned_2.12	w2v_try_v2.1	30	0.46590608	0.17190465	0.4363562

Pengujian terhadap hasil similarity 1 tidak mengalami perubahan dikarenakan kedua kata tersebut tidak terdapat pada dataset chat. Oleh karena itu tidak terjadi perubahan pada similarity vector kedua kata tersebut.

5.2.2. Fine-Tuning Percobaan 2

Pada percobaan kedua dilakukan dengan menggabungkan pattern tiap intent menjadi 1 paragraf, dan melakukan tokenisasi ke paragraf tersebut sebelum di-concat dengan tiap intent.

Kondisi awal pada percobaan 2 melanjutkan kondisi dataset terakhir yang digunakan pada percobaan 1. Data yang digunakan berupa list of word per *tokenized paragraph* setelah melalui proses *concatenation intent* sebagai *input* atau *training data* untuk model Word2Vec.

Pengujian dilakukan dengan mengganti jenis pre-trained word2vec yang digunakan, serta durasi epoch model. Perbandingan hasil akurasi dilakukan diantara penghitungan hasil similarity dari dua kata. Terdapat tiga pasang kata yang digunakan sebagai metrik pengukuran untuk tes kedekatan similarity, karena kata-kata tersebut memiliki nilai semantik yang serupa. Hasil similarity 1 didapatkan dengan membandingkan antara kata 'belajar' dan 'membaca'. Sedangkan hasil similarity 2 didapatkan dengan membandingkan antara kata 'submit' dan 'kumpul'. Hasil similarity 3 didapatkan dengan membandingkan antara kata 'dosen' dan 'dosbing'. Penghitungan similaritynya sendiri dilakukan menggunakan *cosine similarity* dan dijalankan dengan function `wv.similarity` yang disediakan oleh gensim.

Tabel 5.8 Hasil Pengujian Fine-Tuning Menggunakan Percobaan 2

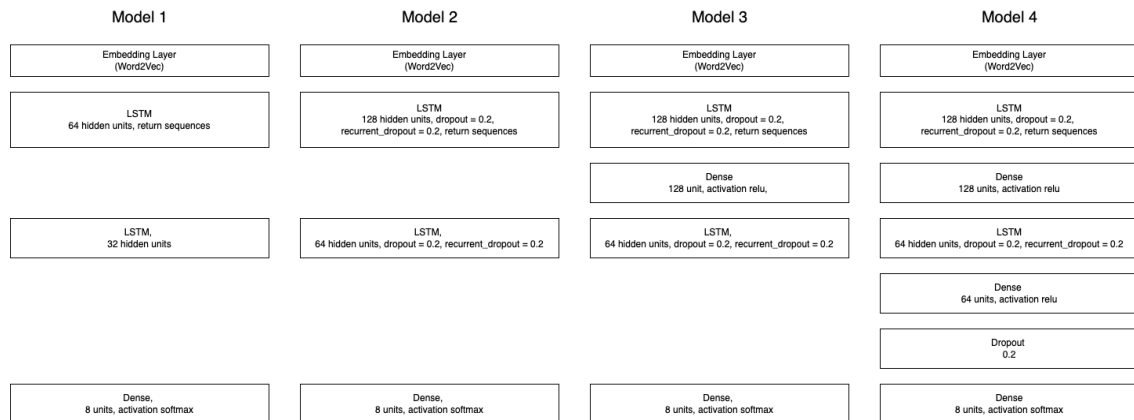
Pengujian	Nama model	Pre-Trained Word2Vec	Epoch	Hasil Similarity 1	Hasil Similarity 2	Hasil Similarity 3
1	model_finetuned_3.0	w2v_try_v2.0	10	0.46590608	0.05564805 5	0.12322900 4
2	model_finetuned_3.1	w2v_try_v2.0	30	0.46590608	0.12619388	0.12322900 4
3	model_finetuned_3.2	w2v_try_v2.0	50	0.46590608	0.12619388	0.30409625

4	model_finetuned_3.3	w2v_try_v2.1	50	0.46590608	0.15456474	0.23810437
5	model_finetuned_3.4	w2v_try_v2.1	30	0.46590608	0.16976874	0.21907449

5.3. Pengujian Model LSTM

Pengujian ini dilakukan untuk mengambil model LSTM dengan akurasi tertinggi menggunakan dataset yang telah disediakan. Percobaan ini dilakukan dengan mengubah struktur dari model, serta parameter yang digunakan untuk mendapatkan model dengan tingkat akurasi tertinggi terhadap dataset test.

Terdapat 4 model yang digunakan sebagai uji coba dengan struktur sebagai berikut:



Gambar 5.1 Perbandingan Struktur Model LSTM

Tabel 5.9 Hasil Pengujian Akurasi pada Model 1

Epoch	model_finetuned_2.5	model_finetuned_2.8	model_finetuned_2.9	model_finetuned_2.11	model_finetuned_2.12
10	0.9396	0.9258	0.9203	0.9066	0.9148
15	0.9341	0.9286	0.9258	0.9093	0.9203

Epoch	model_finetuned_2.5	model_finetuned_2.8	model_finetuned_2.9	model_finetuned_2.11	model_finetuned_2.12
20	0.9341	0.9396	0.9341	0.9148	0.9286
10 + refit 10	0.9451	0.9313	0.9231	0.9121	0.9121
30	0.9423	0.9286	0.9341	0.9341	0.9286

Telah dilakukan percobaan dengan mengganti rasio splitting data pada training model, dari 80:20 menjadi 70:30. Percobaan ini menghasilkan akurasi berjumlah **0.9103** dalam training selama 20 epoch dan menggunakan model_finetuned_2.9. Namun dikarenakan akurasi yang dihasilkan lebih kecil daripada akurasi model yang menggunakan pembagian data dengan rasio semula, maka percobaan dilanjutkan dengan rasio 80:20.

Setelah melakukan pengujian pada model 1, terjadi perubahan pada dataset yang digunakan, sehingga pengujian selanjutnya dijalankan tanpa menggunakan model_finetuned_2.5. Dari tabel 5.9 juga dapat dilihat bahwa kebanyakan model mencapai akurasi tertinggi pada pengujian selama 20 epoch dan 30 epoch, oleh karena itu pengujian selanjutnya akan lebih berfokus pada 20 epoch dan 30 epoch.

Tabel 5.10 Hasil Pengujian Akurasi pada Model 2

Epoch	model_finetuned_2.8	model_finetuned_2.9	model_finetuned_2.11	model_finetuned_2.12
20	0.9478	0.9396	0.9258	0.9423
30	0.9423	0.9286	0.9396	0.9560

Pada pengujian untuk model 3 dan model 4, dilakukan pengujian tambahan dengan melatih model selama 50 epoch untuk menunjukkan apakah model telah mengalami overfitting.

Tabel 5.11 Hasil Pengujian Akurasi pada Model 3

Epoch	model_finetuned_2.8	model_finetuned_2.9	model_finetuned_2.11	model_finetuned_2.12
20	0.9368	0.9368	0.9451	0.9451

30	0.9451	0.9478	0.9505	0.9533
50	0.9423	0.9396	0.9423	0.9313

Tabel 5.12 Hasil Pengujian Akurasi pada Model 4

Epoch	model_finetuned_ 2.8	model_finetuned_ 2.9	model_finetuned_ 2.11	model_finetuned_ 2.12
20	0.9451	0.9313	0.9396	0.9396
30	0.9313	0.9478	0.9396	0.9505
50	0.9341	0.9368	0.9231	0.9451

Hasil pengujian yang ada menunjukkan bahwa 30 epoch merupakan durasi training yang optimal untuk mencapai akurasi tertinggi di sebagian besar model embedding. Selain itu, dengan membandingkan hasil dari pelatihan tiap model (tabel 5.13, 5.14, 5.15, dan 5.16), dapat dilihat bahwa hasil tertinggi didapatkan pada Model 2 (tabel 5.14) yang dijalankan selama 30 epoch menggunakan model_finetuned_2.12 sebagai embedding model. Oleh karena itu model inilah yang nantinya akan digunakan pada implementasi hybrid chatbot.

5.4. Pengujian Survey Pengguna

Pengujian ini dilakukan untuk mengetahui kepuasan pengguna terhadap ketiga model chatbot untuk menentukan jenis chatbot sesuai preferensi mahasiswa jurusan teknik Informatika Universitas Kristen Petra. Pengujian ini dilakukan dengan melakukan survey terhadap 10 mahasiswa jurusan teknik informatika angkatan 2020 dan 2021 yang telah melalui maupun sedang melakukan pengerjaan proposal skripsi sebagai responden. Responden akan melakukan uji coba terhadap ketiga model chatbot dan mengisi survey sesuai dengan pengalaman selama mencoba model chatbot. Berikut pertanyaan survey yang diajukan untuk tiap chatbot:

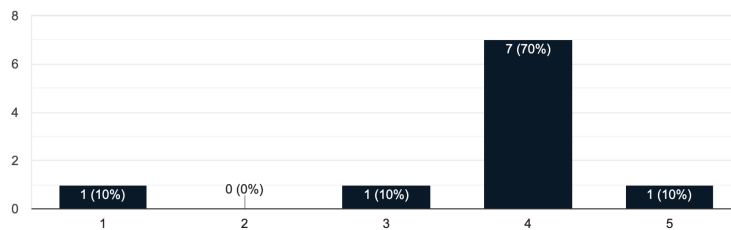
- Apakah jawaban dari chatbot ini membantu selama proses pengerjaan proposal skripsi hingga submit, baik secara teknis maupun alur?
- Apakah chatbot ini efisien dan mudah untuk digunakan?
- Apakah jawaban dari chatbot ini mudah untuk dipahami?

Setelah menjawab survey, beberapa dari responden akan diberi pertanyaan terkait kepuasan selama penggunaan chatbot serta alasan dari jawaban yang telah mereka berikan pada survey. Sebagian besar dari responden cenderung memiliki respon positif terhadap penggunaan Scripted Chatbot untuk mengakomodasi keperluan selama penggunaan chatbot dikarenakan penggunaannya yang cenderung simpel, cepat dan mengakomodasi jawaban sesuai dengan kebutuhan. Berikut ini hasil dari survey yang telah dilakukan:

5.4.1. Hybrid Chatbot

Apakah jawaban dari chatbot ini membantu selama proses pengerjaan proposal skripsi hingga submit, baik secara teknis maupun alur?

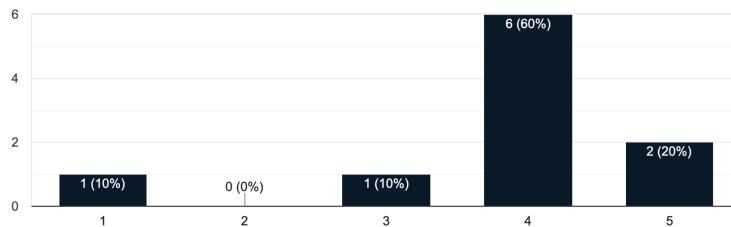
10 responses



Gambar 5.2 Hasil Pertanyaan 1 pada Hybrid Chatbot.

Apakah chatbot ini efisien dan mudah untuk digunakan?

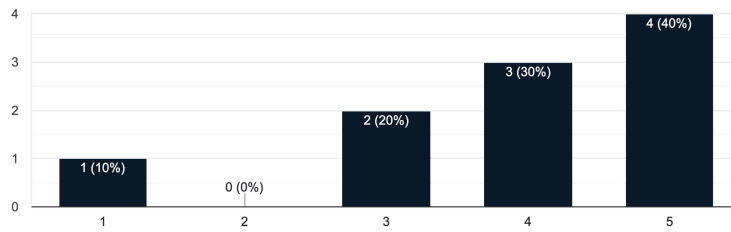
10 responses



Gambar 5.3 Hasil Pertanyaan 2 pada Hybrid Chatbot

Apakah jawaban dari chatbot ini mudah untuk dipahami?

10 responses

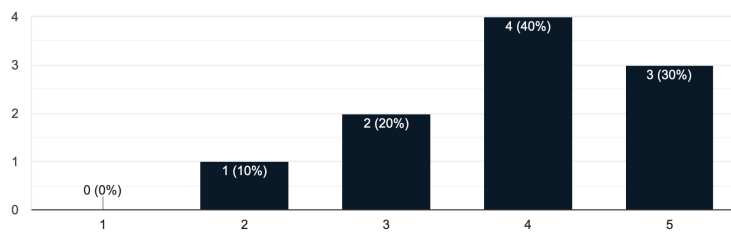


Gambar 5.4 Hasil Pertanyaan 3 pada Hybrid Chatbot

5.4.2. Rule-Based Chatbot

Apakah jawaban dari chatbot ini membantu selama proses pengerjaan proposal skripsi hingga submit, baik secara teknis maupun alur?

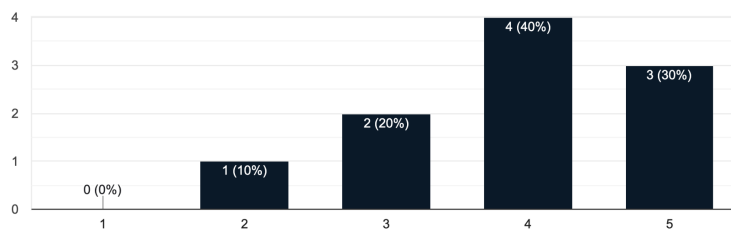
10 responses



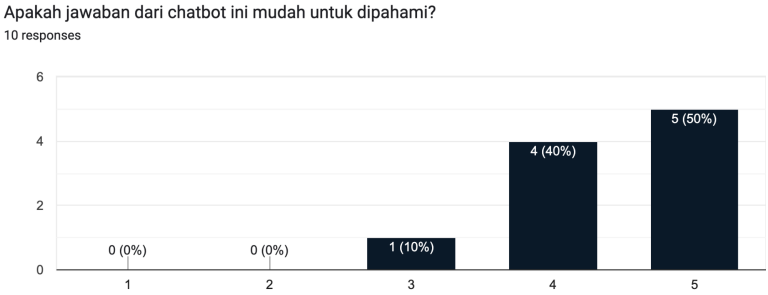
Gambar 5.5 Hasil Pertanyaan 1 pada Rule-Based Chatbot

Apakah chatbot ini efisien dan mudah untuk digunakan?

10 responses

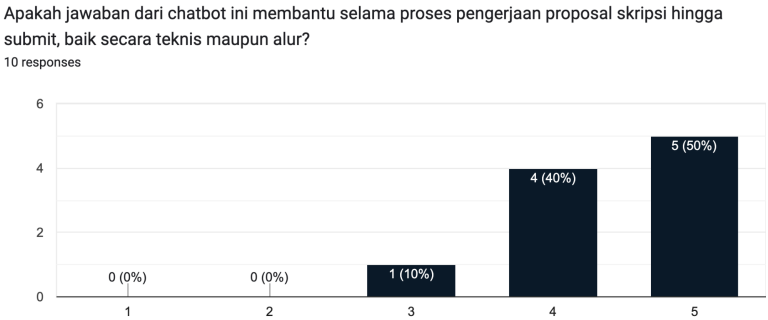


Gambar 5.6 Hasil Pertanyaan 2 pada Rule-Based Chatbot

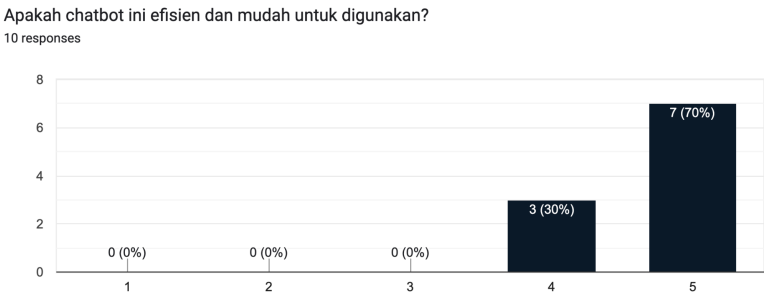


Gambar 5.7 Hasil Pertanyaan 3 pada Rule-Based Chatbot

5.4.3. Scripted Chatbot

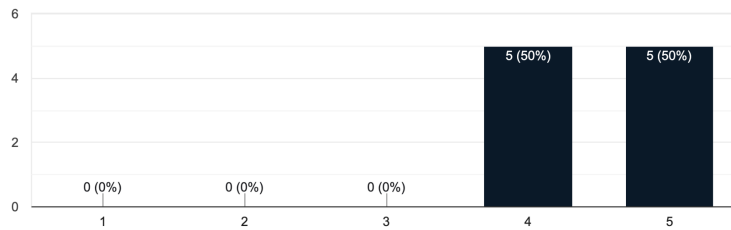


Gambar 5.8 Hasil Pertanyaan 1 pada Scripted Chatbot



Gambar 5.9 Hasil Pertanyaan 2 pada Scripted Chatbot

Apakah jawaban dari chatbot ini mudah untuk dipahami?
10 responses



Gambar 5.10 Hasil Pertanyaan 3 pada Scripted Chatbot

Tabel 5.13 Rata-Rata Penilaian Survei Jenis Chatbot

Jenis Chatbot	Avg. Question 1	Avg. Question 2	Avg. Question 3	Total Rating
Hybrid Chatbot	3.7	4	3.9	3.87
Rule-Based Chatbot	3.9	3.9	4.4	4.07
Scripted Chatbot	4.5	4.7	4.5	4.57

Berdasarkan Tabel 5.13, dapat dilihat bahwa scripted chatbot memiliki keseluruhan rata-rata tertinggi di antara jenis lainnya. Salah satu penyebab yang diketahui dikarenakan kebanyakan merasa lebih nyaman untuk memilih pertanyaan yang telah disediakan daripada mengetik pertanyaan pada chatbot. Alasan lainnya yang dapat mempengaruhi hasil ini dikarenakan keterbatasan intent pada chatbot sehingga kebanyakan pertanyaan yang diajukan oleh pengguna tidak mendapatkan jawaban yang memuaskan sehingga kebanyakan pengguna lebih memilih untuk menggunakan scripted dikarenakan jawaban yang mudah didapatkan.

Berikut data pertanyaan gabungan yang ditanyakan oleh 10 responden. Data berikut dinilai dengan memberi TRUE pada jawaban yang menjawab pertanyaan user dan FALSE pada jawaban yang tidak menjawab maupun diluar dari pertanyaan user. Pertanyaan yang bernilai UNKNOWN menandakan pertanyaan user tersebut tidak dapat dijawab oleh chatbot.

Tabel 5.14 Data Prediksi Intent pada Hybrid Chatbot

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
kumpul kapan	PERSIAPAN	PERSIAPAN_Alur	TRUE
mau topik	JUDUL	JUDUL_TawaranTopik	TRUE
rekom dosbing	DOSBING	DOSBING_Deskripsi	FALSE
rekomendasi dosbing	DOSBING		FALSE
minta dosbing	DOSBING	DOSBING_EtikaPermohonan	TRUE
apakah boleh maju sidang ber 2?	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	FALSE
kapan tau dosen penguji?	POST-SIDANG		FALSE
kumpul dimana?	POST-SIDANG	POST-SIDANG_PengumpulanRevisi	TRUE
topik sim	PERSIAPAN	PERSIAPAN_ProsedurPendaftaranProposal	TRUE
mau makan apa?	DOSBING		UNKNO WN
makan apa hari ini	PROPOSAL		UNKNO WN
ga mau skripsi	DOSBING	DOSBING_EtikaPermohonan	UNKNO WN
ingin lulus	POST-SIDANG		UNKNO WN
gimana minta dosbing?	DOSBING	DOSBING_EtikaPermohonan	TRUE
kumpul dimana?	POST-SIDANG	POST-SIDANG_PengumpulanRevisi	TRUE
kumpul kapan?	PERSIAPAN	PERSIAPAN_Alur	TRUE
gimana tulis latar belakang?	SIDANG		UNKNO WN
ruang lingkup harus sebanyak apa?	PROPOSAL	PROPOSAL_RuangLingkup	TRUE
latar belakang berapa halaman?	PROPOSAL		FALSE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
beda tinjauan studi dan tinjauan pustaka	PROPOSAL	PROPOSAL_TinjauanStudi	TRUE
proposal kapan kumpul	PERSIAPAN	PERSIAPAN_Alur	TRUE
apakah boleh maju sidang ber 2?	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	FALSE
apakah boleh sidang ber2?	PERSIAPAN	PERSIAPAN_KetentuanSidang	TRUE
bagaimana jika tidak bisa hadir onsite?	PERSIAPAN	PERSIAPAN_ProsedurPembatalanProposal	FALSE
sidang online	SIDANG	SIDANG_ProsedurSidangOnline	TRUE
sidang onlen	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	TRUE
apakah bisa jika tidak datang onsite untuk sidang?	PERSIAPAN	PERSIAPAN_KeperluanSidang	FALSE
apakah bisa sidang online	SIDANG	SIDANG_KetentuanSidangOnline	TRUE
bagaimana jika tidak bisa datang sidang?	PERSIAPAN	PERSIAPAN_KeperluanSidang	FALSE
apakah bisa jika ingin mendaftarkan sidang onsite?	PERSIAPAN	PERSIAPAN_ProsedurPendaftaranDosbing	FALSE
submit dosbing kapan?	PERSIAPAN	PERSIAPAN_Alur	TRUE
review kbid	PERSIAPAN	PERSIAPAN_ReviewKabid	TRUE
apakah topik bisa beda dari konsentrasi	JUDUL	JUDUL_BedaTopikKonsentrasi	TRUE
bagaimana format latar belakang yang benar?	POST-SIDANG		FALSE
bagaimana cara mengumpulkan proposal skripsi	PERSIAPAN	PERSIAPAN_ProsedurPendaftaranDosbing	FALSE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
kumpul proposal skripsi lewat mana	PERSIAPAN	PERSIAPAN_Alur	FALSE
siapa kepala bidang kita	PROPOSAL		UNKNOWN
rekomendasi dosbing untuk topik tema AI ke siapa	JUDUL		FALSE
apakah bisa memilih dosbing sebelum memilih topik	JUDUL	JUDUL_Konsultasi	FALSE
submit judul topik kemana	JUDUL	JUDUL_RevisiJudul	FALSE
revisi judul kesiapa	DOSBING		FALSE
latar belakang maksimal berapa halaman	PROPOSAL		UNKNOWN
latar belakang maksimal berapa halaman	PROPOSAL		UNKNOWN
proposal kumpul ke siapa dan kapan terakhir kumpul	PERSIAPAN		FALSE
kapan terakhir kumpul	PERSIAPAN	PERSIAPAN_Alur	TRUE
apa bedanya periode 1 dan periode 2	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	FALSE
halo	SIDANG		UNKNOWN
kerja proposal harus mulai dari mana ya	PERSIAPAN		FALSE
mau kerja proposal skripsi mulai dari mana ya	PERSIAPAN		FALSE
apa yang perlu dipersiapkan untuk proposal skripsi	PERSIAPAN	PERSIAPAN_KeperluanSidang	FALSE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
gimana cara pilih dosen pembimbing	DOSBING	DOSBING_EtikaPermohonan	FALSE
apa yang bisa membuang proposal tertolak	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	FALSE
apa syarat supaya porposal terterima	POST-SIDANG		FALSE
kapan sidang dimulai	PERSIAPAN		FALSE
kapan pengumpulan	PERSIAPAN	PERSIAPAN_Alur	TRUE
kapan sidang proposal	PERSIAPAN	PERSIAPAN_Alur	TRUE
kapan deadline propoal ya ?	PERSIAPAN	PERSIAPAN_Alur	TRUE
anjay jadi	DOSBING		UNKNO WN
kalau dosbing yang enak kira" apa ya ??	DOSBING		UNKNO WN
kapan deadline kumpul proposal?	PERSIAPAN	PERSIAPAN_Alur	TRUE
jurnal yang dipake minimal berapa tahun terakhir?	PROPOSAL		FALSE
referensi yang digunakan minimal tahun berapa?	PROPOSAL		FALSE
apakah ruang lingkup harus sampai perkiraan output?	PROPOSAL	PROPOSAL_RuangLingkup	TRUE
tinjauan pustaka itu apa	PROPOSAL	PROPOSAL_LembarUsulanSkripsi	UNKNO WN
tinjauan studi harus ada di latar belakang kah	PROPOSAL	PROPOSAL_TinjauanStudi	TRUE
tinjauan studi minimal berapa	PROPOSAL	PROPOSAL_TinjauanStudi	TRUE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
latar belakang minimal berapa paragraf?	PROPOSAL		UNKNOWN
apakah topik proposal harus sesuai dengan konsentrasi yang diambil?	JUDUL	JUDUL_BedaTopikKonsentrasi	TRUE
kapan bisa mengetahui dosen penguji proposal?	POST-SIDANG	POST-SIDANG_AlurRevisi	FALSE
sampai manakah ruang lingkup harus dituliskan	PROPOSAL	PROPOSAL_RuangLingkup	TRUE
apakah ada maksimal halaman dalam penulisan proposal	PROPOSAL		FALSE
bagaimana cara menentukan pembimbing skripsi	DOSBING	DOSBING_EtikaPembatalan	FALSE
saya mau mencari pembimbing skripsi	DOSBING	DOSBING_EtikaPermohonan	TRUE
bagaimana saya mencari pembimbing skripsi	DOSBING	DOSBING_EtikaPembatalan	TRUE
cara penulisan tinjauan studi	PROPOSAL	PROPOSAL_TinjauanStudi	TRUE
bagaimana jadwal harus dituliskan	PERSIAPAN		FALSE
adakah minimal rumusan masalah	PROPOSAL		FALSE
adakah minimal rumusan masalah yang harus dituliskan	PROPOSAL		FALSE
bagaimana cara membuat latar belakang untuk proposal	POST-SIDANG		UNKNOWN
bagaimana menurut anda tentang topik dengan judul "Pengembangan Sistem	JUDUL		FALSE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
Informasi Manajemen Tugas Akhir (SIMTA)"			
topik apa yang bisa direkomendasikan dengan jenis proyek akhir website	PROPOSAL		FALSE
dosen yang cocok untuk artificial intelligence	DOSBING	DOSBING_Rekomendasi	TRUE
berikan ide untuk topik proposal	JUDUL		TRUE
apa itu proposal	PERSIAPAN	PERSIAPAN_KeperluanSidang	FALSE
apa itu skripsi proposal bukan sidang	PERSIAPAN		FALSE
apa itu sidang	PERSIAPAN		FALSE
apa itu bimbingan	PERSIAPAN		FALSE
Apakah topik "Pengembangan Website Sistem Informasi Manajemen Tugas Akhir" termasuk topik untuk jurusan Informatika?	JUDUL		FALSE
cara buat proposal	PROPOSAL	PROPOSAL_KerangkaProposal	TRUE
cara menentukan topik	JUDUL		FALSE
saya benci skripsi	SKRIPSI		UNKNO WN
saya benci proposal	PERSIAPAN		UNKNO WN
cara melakukan bimbingan	PERSIAPAN	PERSIAPAN_ProsedurPendaftaranDosbing	FALSE
berapa harus bimbingan	PERSIAPAN	PERSIAPAN_JumlahBimbingan	TRUE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
apa itu bimbingan	PERSIAPAN		FALSE
cara batalkan dosen bimbingan	PERSIAPAN	PERSIAPAN_ProsedurPembatalanProposal	FALSE
cara memilih dosen pembimbing	DOSBING		FALSE
apa itu dosen pembimbing	DOSBING	DOSBING_Deskripsi	TRUE
untuk setiap dosbing berapa bimbingan	PERSIAPAN	PERSIAPAN_JumlahBimbingan	FALSE
apa itu bimbingan	PERSIAPAN		FALSE
judul topik	JUDUL	JUDUL_Konsultasi	TRUE
judul topik yang bagus	JUDUL		FALSE
informasi periode	JUDUL		FALSE
periode	PERSIAPAN	PERSIAPAN_ProsedurSidangUlang	FALSE
periode deadline pengumpulan	PERSIAPAN	PERSIAPAN_Alur	TRUE
deadline kapan?	PERSIAPAN	PERSIAPAN_Alur	TRUE
deadline pengumpulan proposal	PERSIAPAN	PERSIAPAN_Alur	TRUE
kalau tidak bisa hadir waktu sidang bagaimana	PERSIAPAN		FALSE
kalau yang di presentasikan tidak sesuai dengan file yang di submit bagaimana?	PROPOSAL		UNKNOWN
bagaimana format penulisan proposal	PROPOSAL	PROPOSAL_KerangkaProposal	TRUE
berapa durasi pengerjaan proposal skripsi?	PROPOSAL		FALSE
apabila proposal pada sidang periode terakhir ditolak, masih bisa mengajukan lagi?	PERSIAPAN	PERSIAPAN_ProsedurPendaftaranProposal Ulang	TRUE

User Question	Predicted top level intent	Predicted sub_level_intent	Truth Value
apabila proposal saya diterima tetapi ternyata saya harus cuti bagaimana?	PERSIAPAN		UNKNOWN
siapa dosen pembimbing yang masih available untuk membimbing?	DOSBING	DOSBING_EtikaPembatalan	FALSE
kalau topik saya mengenai data mining, siapa dosen pembimbing yang sesuai?	DOSBING		FALSE
apa ruang lingkup proposal?	PERSIAPAN	PROPOSAL_RuangLingkup	TRUE
ruang lingkup di proposal?	PROPOSAL	PROPOSAL_RuangLingkup	TRUE
ruang lingkup proposal?	PERSIAPAN	PROPOSAL_RuangLingkup	TRUE
isi ruang lingkup apa aja?	PROPOSAL	PROPOSAL_RuangLingkup	TRUE
		Benar	46
		Salah	55
		Tidak Terjawab	18
		Akurasi	45%

Dari data pertanyaan yang tidak terjawab, dapat dianalisa lebih lanjut apakah pertanyaan tersebut termasuk dalam kategori intent yang baru, atau data tersebut tidak termasuk dalam cakupan chatbot.

Tabel 5.15 Data Analisa Pertanyaan Tidak Terjawab pada Hybrid Chatbot

No	User Question	Predicted Intent	Suggested Intent	Note
1	mau makan apa?	DOSBING	-	
2	makan apa hari ini	PROPOSAL	-	

No	User Question	Predicted Intent	Suggested Intent	Note
3	ga mau skripsi	DOSBING_EtikaPer mohonan	-	
4	ingin lulus	POST-SIDANG	-	
5	latar belakang maksimal berapa halaman	PROPOSAL	PROPOSAL_LatarBe lakang	Intent Baru
6	latar belakang maksimal berapa halaman	PROPOSAL	PROPOSAL_LatarBe lakang	Intent Baru
7	siapa kepala bidang kita	PROPOSAL	LAIN-LAIN_Kepala bidang	Intent Baru
8	tinjauan pustaka itu apa	PROPOSAL_Lembar UsulanSkripsi	PROPOSAL_Tinjaua n Pustaka	Intent Baru
9	kalau yang di presentasikan tidak sesuai dengan file yang di submit bagaimana?	PROPOSAL	PERSIAPAN_PPT	Dataset baru
10	apabila proposal saya diterima tetapi ternyata saya harus cuti bagaimana?	PERSIAPAN	-	Kasus khusus yang tidak bisa dijawab menggunakan chatbot
11	gimana tulis latar belakang?	SIDANG	PROPOSAL_LatarB elakang	Intent Baru
12	saya benci skripsi	SKRIPSI	-	
13	saya benci proposal	PERSIAPAN	-	
14	bagaimana cara membuat latar belakang untuk proposal	POST-SIDANG	PROPOSAL_LatarBe lakang	Intent Baru
15	latar belakang minimal berapa paragraf?	PROPOSAL	PROPOSAL_LatarBe lakang	Intent Baru
16	anjay jadi	DOSBING	-	
17	kalau dosbing yang enak kira" apa ya ??	DOSBING	-	

No	User Question	Predicted Intent	Suggested Intent	Note
18	halo	SIDANG	SAPAAN	Intent Baru

Dari Tabel 5.15 dapat dilihat bahwa 9 dari 18 pertanyaan merupakan pertanyaan yang tidak dapat terjawab dikarenakan kurangnya data intent pada chatbot.

Tabel 5.16 Data Prediksi Intent pada Rule-Based Chatbot

User Question	Predicted Intent	Truth Value
apakah format penulisan proposal?	PROPOSAL_KerangkaProposal	TRUE
tugas akhir	PROPOSAL_Pengujian	FALSE
Pak greg	LAIN-LAIN_EmailDosen	TRUE
Bu, saya bingung belum menemukan judul untuk proposalku	JUDUL_Konsultasi	TRUE
Bu, saya bingung alur dari seminar proposal	PROPOSAL_LampiranGDD	FALSE
bu jelaskan prosedur dari seminar proposal	SIDANG_ProsedurSidangOnsite	TRUE
Bu kenal kelvin sidharta sie	LAIN-LAIN_InformasiDosen	UNKNOWN
Calvert	PERSIAPAN_ProsedurPendaftaranProposal	UNKNOWN
wisuda	PERSIAPAN_KetentuanSidang	UNKNOWN
bimbingan apa itu	DOSBING_Deskripsi	UNKNOWN
judul topik yang bagus	DOSBING_Rekomendasi	FALSE
apakah judul "pengembangan website simta" merupakan yang bagus?		FALSE
deadline kapan proposal?	PERSIAPAN_Alur	TRUE
apa itu proposal?	PROPOSAL_KerangkaProposal	TRUE
cara memilih dosbing yang baik		FALSE
dosbing	DOSBING_Deskripsi	TRUE
sidang proposal harus bagaimana	PERSIAPAN_KeperluanSidang	TRUE
apa itu proposal?	PROPOSAL_KerangkaProposal	TRUE
apakah proposal itu sulit?		UNKNOWN
kapan sidang	PERSIAPAN_Alur	TRUE
siapa dosen yang masih tersedia untuk bimbingan		UNKNOWN
template proposal seperti apa		FALSE

User Question	Predicted Intent	Truth Value
kapan pengumpulan	POST-SIDANG_PengumpulanRevisi	TRUE
kapan periode 1	PERSIAPAN_Alur	TRUE
kapan sidang?	PERSIAPAN_Alur	TRUE
siapa dosen pembimbing yang terbaik untuk jurusan informatika?		UNKNOWN
jurnal yang dipakai apakah harus internasional?	PROPOSAL_Referensi	TRUE
bagaimana cara mencari dosen pembimbing skripsi	DOSBING_EtikaPembatalan	FALSE
saya mau mencari dosbing skripsi		FALSE
bagaimana menemukan dosen pembimbing skripsi	DOSBING_EtikaPembatalan	FALSE
apakah judul proposal harus sama dengan konsentrasi	JUDUL_Konsultasi	FALSE
apakah topik proposal harus sama dengan konsentrasi yang diambil	JUDUL_TopikMitra	FALSE
kapan penjadwalan sidang proposal dikeluarkan	SIDANG_Jadwal	TRUE
kapan bisa mengetahui dosen penguji	DOSBING_EtikaPermohonan	FALSE
sampai manakah tinjauan pustaka harus ditulis		UNKNOWN
bagaimana penulisan tinjauan pustaka	PROPOSAL_RuangLingkup	UNKNOWN
tinjauan studi harus berapa penelitian	PROPOSAL_TinjauanStudi	TRUE
tinjauan pustaka isinya apa		UNKNOWN
bagaimana alur proposal		FALSE
proposal kapan dikumpulkan	POST-SIDANG_PengumpulanRevisi	TRUE
proposal kumpul ke siapa		UNKNOWN
revisi judul ke siapa		UNKNOWN
kapan terakhir kumpul	POST-SIDANG_PengumpulanRevisi	TRUE
format latar belakang		UNKNOWN
bagaimana format latar belakang yang benar		UNKNOWN
apa bedanya periode 1 dan periode 2	DOSBING_Deskripsi	UNKNOWN
apakah dosen pembimbing harus 2	DOSBING_EtikaPembatalan	FALSE
siapa dosen pembimbing yang cocok untuk topik AI	DOSBING_Rekomendasi	TRUE
kumpul proposal skripsi lewat mana	POST-SIDANG_PengumpulanRevisi	TRUE

User Question	Predicted Intent	Truth Value
jika tidak ada revisi proposal bagaimana	POST-SIDANG_TandaTanganRevisi	FALSE
kapan pengumpulan?	POST-SIDANG_AlurRevisi	TRUE
kumpul kapan?	POST-SIDANG_AlurRevisi	TRUE
dimana kumpul?	POST-SIDANG_PengumpulanRevisi	TRUE
submit judul	JUDUL_RevisiJudul	FALSE
tanya dosbing	DOSBING_Deskripsi	TRUE
gimana tulis proposal?	PROPOSAL_KerangkaProposal	TRUE
proposal submit mana?	PROPOSAL_LampiranGDD	FALSE
topik skripsi dimasukkin sim dimana?	PERSIAPAN_ProsedurPendaftaranPropo sal	TRUE
topik sim	PERSIAPAN_ProsedurPendaftaranPropo sal	TRUE
minta judul	LAIN-LAIN_CodeClassroom	FALSE
mau judul skripsi	JUDUL_Konsultasi	FALSE
ada judul skripsi	JUDUL_Konsultasi	FALSE
tawaran jduul		FALSE
gimana prosposal	DOSBING_EtikaPermohonan	FALSE
gimana proposal	DOSBING_EtikaPermohonan	FALSE
mau maju sidang ber 2		FALSE
maju ber2		FALSE
apakah boleh maju sidang ber 2?	PERSIAPAN_ProsedurSidangUlang	FALSE
apakah boleh memilih dosen penguji?	JUDUL_Konsultasi	FALSE
dosen penguji tau kapan?		FALSE
kapan tau dosen penguji?		FALSE
rekomendasi dosbing		FALSE
rekom dosbing	DOSBING_Deskripsi	FALSE
proposal kapan kumpul?	POST-SIDANG_AlurRevisi	FALSE
proposal sebelum sidang kapan kumpul?	PERSIAPAN_Alur	TRUE
revisi kapan kumpul	POST-SIDANG_AlurRevisi	TRUE
apakah boleh maju sidang ber2?	SIDANG_KetentuanSidangOnline	FALSE
bagaimana jika tidak bisa datang sidang?	PERSIAPAN_KeperluanSidang	FALSE
bagaimana jika berhalangan hadir onsite?		FALSE
tidak bisa hadir sidang onsite	PERSIAPAN_ProsedurPembatalanPropo sal	FALSE

User Question	Predicted Intent	Truth Value
apakah bisa isdang onlen?	JUDUL_TawaranTopik	FALSE
apakah bisa sidang online?	SIDANG_KetentuanSidangOnline	TRUE
apakah boleh tidak hadir sidang?	PERSIAPAN_ProsedurPembatalanProposal	FALSE
	Salah	37
	Benar	30
	Tidak Terjawab	16
	Akurasi	44.7%

Tabel 5.17 Data Analisa Pertanyaan Tidak Terjawab pada Rule-Based Chatbot

No	User Question	Predicted Intent	Suggested Intent	Note
1	Bu kenal kelvin sidharta sie	LAIN-LAIN_InformasiDosen	-	
2	Calvert	PERSIAPAN_ProsedurPendaftaranProposal	-	
3	wisuda	PERSIAPAN_KetentuanSidang	-	
4	bimbingan apa itu	DOSBING_Deskripsi	-	
5	apakah proposal itu sulit?		-	
6	siapa dosen yang masih tersedia untuk bimbingan		DOSBING_KuotaDosing	
7	siapa dosen pembimbing yang terbaik untuk jurusan informatika?		-	
8	sampai manakah tinjauan pustaka harus ditulis		PROPOSAL_TinjauanPustaka	
9	bagaimana penulisan tinjauan pustaka	PROPOSAL_RuangLingkup	PROPOSAL_TinjauanPustaka	
10	tinjauan pustaka isinya apa		PROPOSAL_TinjauanPustaka	

No	User Question	Predicted Intent	Suggested Intent	Note
11	proposal kumpul ke siapa		PROPOSAL_ProsedurPendaftaranProposal	Dataset baru
12	revisi judul ke siapa		POST-SIDANG_PengumpulanRevisi	Dataset baru
13	apa bedanya periode 1 dan periode 2	DOSBING_Desripsi	SIDANG_Periode	
14	format latar belakang		PROPOSAL_LatarBelakang	
15	bagaimana format latar belakang yang benar		PROPOSAL_LatarBelakang	

Dari Tabel 5.17, dapat dilihat pada rule-based chatbot terdapat 9 pertanyaan dari 16 pertanyaan yang tidak dapat terjawab dikarenakan oleh kurangnya data intent pada chatbot.

5.5. Pengujian Running Time

Dikarenakan terdapat perbedaan yang signifikan terhadap lama waktu chatbot memproses input user dan memberikan respon balik, maka dilakukan pengujian untuk menemukan penyebab terjadinya delay yang cukup lama pada respons chatbot.

5.5.1. Pengujian *Running Time* Rule Based

Pengujian ini dilakukan untuk menemukan langkah yang memakan waktu terbanyak dari proses pengolahan input user, hingga pengembalian respons kepada user. Pengujian ini dilakukan dengan menghitung selisih waktu pada awal dan akhir di tiap proses. Proses yang diukur berupa: *running time* untuk melakukan embedding pada pattern di tiap intent (Waktu Pattern Embedding), *running time* untuk melakukan rata-rata perhitungan embedding pattern (Waktu Rule Embedding), dan *running time* untuk melakukan penghitungan cosine similarity antara user input dan mean nilai embedding pattern (Waktu Cosine Similarity). Pada waktu pengujian, juga dilakukan pengukuran *running time* untuk proses embedding dari user input. Pada pertanyaan yang berisi 2 kata, waktu embedding yang diperlukan adalah 0:00:00.029000.

Sedangkan pada pertanyaan yang berisi 46 kata, waktu embedding yang diperlukan adalah 0:00:00.411469.

Tabel 5.18 Data Running Time Rule Based Chatbot

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Counting
JUDUL_Konsultasi	0:00:03.274201	0:00:00.000824	0:00:00.001560
JUDUL_BedaTopikKonsentrasi	0:00:05.660565	0:00:00.000250	0:00:00.000148
JUDUL_KetentuanPenulisan	0:00:02.279902	0:00:00.000098	0:00:00.000075
JUDUL_TopikMitra	0:00:03.661949	0:00:00.000079	0:00:00.000072
JUDUL_TawaranTopik	0:00:01.180692	0:00:00.000106	0:00:00.000070
JUDUL_RevisiJudul	0:00:02.023734	0:00:00.000096	0:00:00.000072
DOSBING_Deskripsi	0:00:05.958180	0:00:00.000139	0:00:00.000061
DOSBING_Rekomendasi	0:00:03.353670	0:00:00.000073	0:00:00.000044
DOSBING_EtikaPermohonan	0:00:06.216481	0:00:00.000084	0:00:00.000055
DOSBING_EtikaPembatalan	0:00:01.519688	0:00:00.000096	0:00:00.000112
DOSBING_Administrasi	0:00:04.236775	0:00:00.000095	0:00:00.000048
PERSIAPAN_Alur	0:00:04.258559	0:00:00.000096	0:00:00.000076
PERSIAPAN_JumlahBimbingan	0:00:00.938485	0:00:00.000155	0:00:00.000068
PERSIAPAN_KeperluanSidang	0:00:01.799113	0:00:00.000101	0:00:00.000061
PERSIAPAN_ProsedurPendaftar anDosbing	0:00:06.639135	0:00:00.000080	0:00:00.000085

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Counting
PERSIAPAN_ProsedurPendaftaranProposal	0:00:04.078815	0:00:00.000129	0:00:00.000087
PERSIAPAN_ProsedurPendaftaranProposalUlang	0:00:05.304209	0:00:00.000094	0:00:00.000058
PERSIAPAN_ProsedurPembatalanProposal	0:00:00.285275	0:00:00.000098	0:00:00.000055
PERSIAPAN_KetentuanSidang	0:00:02.309585	0:00:00.000085	0:00:00.000053
PERSIAPAN_ProsedurSidangUlang	0:00:01.884825	0:00:00.000179	0:00:00.000105
PERSIAPAN_AppointmentGoogleCalendar	0:00:10.630820	0:00:00.000099	0:00:00.000060
PERSIAPAN_ReviewKabid	0:00:13.797294	0:00:00.000369	0:00:00.000576
PROPOSAL_KerangkaProposal	0:00:00.985174	0:00:00.000166	0:00:00.000070
PROPOSAL_LembarUsulanSkripsi	0:00:00.865882	0:00:00.000180	0:00:00.000096
PROPOSAL_ProsedurProposalMitra	0:00:04.751308	0:00:00.000114	0:00:00.000062
PROPOSAL_TinjauanStudi	0:00:01.590491	0:00:00.000087	0:00:00.000052
PROPOSAL_RuangLingkup	0:00:00.509408	0:00:00.000083	0:00:00.000060
PROPOSAL_Pengujian	0:00:02.667453	0:00:00.000092	0:00:00.000047
PROPOSAL_MuatanKonsentrasi	0:00:01.267677	0:00:00.000069	0:00:00.000052
PROPOSAL_Referensi	0:00:06.284999	0:00:00.000122	0:00:00.000112

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Counting
PROPOSAL_LampiranGDD	0:00:00.541398	0:00:00.000107	0:00:00.000056
PROPOSAL_LampiranGambar	0:00:03.574901	0:00:00.000117	0:00:00.000065
PROPOSAL_LampiranTranskrip	0:00:00.548846	0:00:00.000183	0:00:00.000056
PROPOSAL_LampiranSurveyNar sum	0:00:03.789482	0:00:00.000109	0:00:00.000053
PROPOSAL_KetentuanPenggun aanData	0:00:02.320217	0:00:00.000139	0:00:00.000065
SIDANG_Dresscode	0:00:03.592656	0:00:00.000100	0:00:00.000061
SIDANG_Jadwal	0:00:00.451455	0:00:00.000059	0:00:00.000045
SIDANG_PPT	0:00:02.691695	0:00:00.000146	0:00:00.000205
SIDANG_ProsedurSidangOnsite	0:00:03.592527	0:00:00.000084	0:00:00.000065
SIDANG_ProsedurSidangOnline	0:00:07.949381	0:00:00.000113	0:00:00.000076
SIDANG_KetentuanSidangOnlin e	0:00:05.015317	0:00:00.000079	0:00:00.000049
POST-SIDANG_Hasil	0:00:00.382011	0:00:00.000069	0:00:00.000083
POST-SIDANG_AlurRevisi	0:00:03.640256	0:00:00.000064	0:00:00.000048
POST-SIDANG_PengumpulanRe visi	0:00:00.455750	0:00:00.000066	0:00:00.000060
POST-SIDANG_TandaTanganRev isi	0:00:08.994102	0:00:00.000144	0:00:00.000214
SKRIPSI_DescJenisTugasAkhir	0:00:01.727557	0:00:00.000092	0:00:00.000060

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Counting
SKRIPSI_ChooseJenisTugasAkhir	0:00:02.371319	0:00:00.000081	0:00:00.000050
LAIN-LAIN_CodeClassroom	0:00:09.738214	0:00:00.000236	0:00:00.000081
LAIN-LAIN_EmailDosen	0:00:00.412466	0:00:00.000103	0:00:00.000071
LAIN-LAIN_InformasiDosen	0:00:04.258105	0:00:00.000206	0:00:00.000093
LAIN-LAIN_JadwalBimbingan	0:00:05.245010	0:00:00.000133	0:00:00.000063
LAIN-LAIN_HistorySkripsi	0:00:01.461818	0:00:00.000133	0:00:00.000114
Total waktu tiap proses	0:03:02.968827	0:00:00.006901	0:00:00.005885
Total Response Time	0:03:03.019953		

Dapat dilihat pada Tabel 5.18, bahwa proses terlama terdapat pada embedding tiap pattern. Hal ini dikarenakan proses embedding dijalankan ke 35 pattern pertanyaan di tiap intent. Oleh karena itu tidak heran bahwa proses ini memakan waktu paling banyak dibandingkan proses-proses lainnya. Dapat dilihat juga bahwa jumlah kata dari pertanyaan user tidak memberikan pengaruh yang signifikan terhadap lama *running time* penentuan respons chatbot.

5.5.2. Pengujian *Running Time* Hybrid Model

Pengujian pada *running time* hybrid model dilakukan untuk mengetahui perbandingan lama proses pada hybrid model dengan rule-based model. Pengujian ini dilakukan dengan menghitung selisih waktu pada awal dan akhir di tiap proses. Dikarenakan hybrid chatbot memiliki 2 tahapan pemrosesan input user, maka penyajian hasil pengujian akan dibagi menjadi 2 bagian.

Bagian pertama berupa penghitungan *running time* keseluruhan yang mencakup: pemrosesan input user (preproses, tokenisasi, padding) sebelum dilanjutkan ke model, lama model melakukan prediksi intent berdasarkan input user, dan lama pemrosesan input lebih lanjut menggunakan *rule-based* untuk pemilihan response chatbot.

Tabel 5.19 Data *Running Time* Hybrid Model Keseluruhan

Waktu Preprocess User Input	Waktu Prediction Model	Waktu Pengerjaan Rule-Based
0:00:00.031285	0:00:00.227508	0:00:50.135837
0:00:50.135837		

Bagian kedua berupa penghitungan running time di dalam proses pemilihan respon berbasis rule-based. Dapat dilihat pada Tabel 5.19 proses ini memakan waktu paling banyak dibandingkan proses lainnya. Diperlukan adanya pengujian ini untuk melihat proses mana yang memakan waktu paling banyak. Proses yang diukur berupa: running time untuk melakukan embedding pada pattern di tiap intent (Waktu Pattern Embedding), *running time* untuk melakukan rata-rata perhitungan embedding pattern (Waktu Rule Embedding), dan *running time* untuk melakukan penghitungan cosine similarity antara user input dan mean nilai embedding pattern (Waktu Cosine Similarity).

Tabel 5.20 Data *Running Time Rule-Based Process*

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Similarity
PERSIAPAN_Alur	0:00:04.090079	0:00:00.000113	0:00:00.000082
PERSIAPAN_JumlahBimbingan	0:00:00.937447	0:00:00.000102	0:00:00.000070
PERSIAPAN_KeperluanSidang	0:00:01.778743	0:00:00.000094	0:00:00.000052
PERSIAPAN_ProsedurPendaftaran anDosbing	0:00:06.774072	0:00:00.000103	0:00:00.000075
PERSIAPAN_ProsedurPendaftaran anProposal	0:00:04.068736	0:00:00.000139	0:00:00.000090
PERSIAPAN_ProsedurPendaftaran anProposalUlang	0:00:05.300510	0:00:00.000106	0:00:00.000067

Intent	Waktu Pattern Embedding	Waktu Rule Embedding	Waktu Cosine Similarity
PERSIAPAN_ProsedurPembatalanProposal	0:00:00.278076	0:00:00.000091	0:00:00.000056
PERSIAPAN_KetentuanSidang	0:00:02.230566	0:00:00.000122	0:00:00.000104
PERSIAPAN_ProsedurSidangUlang	0:00:01.756567	0:00:00.000090	0:00:00.000067
PERSIAPAN_AppointmentGoogleCalendar	0:00:11.233774	0:00:00.000230	0:00:00.000129
PERSIAPAN_ReviewKabid	0:00:11.677072	0:00:00.000230	0:00:00.000184
Total waktu tiap proses	0:00:50.125642	0:00:00.001420	0:00:00.000976
Total Rule-Based Process Time	0:00:50.135837		

Dikarenakan proses rule-based ini merupakan proses lanjutan dari tahapan 1 (penentuan top level intent menggunakan LSTM model), maka jumlah intent yang perlu dihitung menyesuaikan dengan jumlah sub-intent dari top-level-intent yang didapatkan pada tahapan 1. Dalam proses ini dapat dilihat bahwa proses yang memakan waktu paling banyak masih terletak pada proses embedding dari tiap pattern pertanyaan yang dimiliki oleh tiap sub_intent.