

Algoritma Stemming Nazief & Adriani Dengan Metode Cosine Similarity Untuk Chatbot Telegram Terintegrasi Dengan E-layanan

Aan Choesni Herlingga¹, IGL. Putra Eka Prismana², Dedy Rahman Prehanto³, Dodik Arwin Dermawan⁴

^{1,2,3,4} Jurusan Teknik Informatika/Teknik Informatika, Universitas Negeri Surabaya

¹aanchoesni@gmail.com

²lanangprisma@unesa.ac.id

³dedyrahman@unesa.ac.id

⁴dodikdermawan@unesa.ac.id

Abstrak—Sistem Informasi yang berkembang memiliki sejumlah pertanyaan yang sering diajukan kepada *customer service* dengan tingkat kesamaan pertanyaan yang tinggi. Untuk membantu hal tersebut dibuat FAQ terkait sistem informasi. FAQ memiliki banyak informasi sehingga pengguna bingung dan memerlukan waktu untuk mencari informasi. Pengguna lebih untuk mengajukan pertanyaan ke *customer service*. Chatbot merupakan salah cara untuk membantu pengguna dan *customer service* dalam masalah ini. *Customer service* dapat menjawab pertanyaan secara otomatis dan pengguna dapat mengajukan pertanyaan seolah-olah bertanya kepada *customer service* secara langsung. Pada penelitian ini peneliti akan menerapkan algoritma Nazief & Adriani yang digunakan untuk melakukan stemming karena algoritma ini merupakan salah satu algoritma yang efektif dalam stemming bahasa Indonesia. Dan metode cosine similarity untuk mencari tingkat kemiripan dari pertanyaan dengan FAQ yang ada dalam database FAQ. Dengan mengimplementasikan algoritma dan metode tersebut dalam chatbot dengan bantuan layanan *Messenger Telegram* menghasilkan jawaban yang relative sesuai dengan yang diharapkan pengguna. Cara ini merupakan cara yang efektif untuk menjawab pertanyaan secara otomatis. Dan jika pertanyaan yang diajukan tidak menemukan jawaban atau jawaban tidak sesuai maka pertanyaan akan di sampaikan dalam e-Layanan PPTI UNESA.

Kata Kunci—algoritma nazief & andriani, cosine similarity, chatbot, bot.

I. PENDAHULUAN

Sistem informasi memiliki banyak fitur dan banyak menu yang bisa digunakan pengguna sistem informasi. Hal tersebut dapat memicu banyak pertanyaan seputar sistem informasi. Pertanyaan yang ditanyakan pengguna seringkali terdapat kemiripan antar pengguna sehingga *customer service* menjawab pertanyaan secara berulang-ulang. Dokumentasi dan video dapat digunakan untuk memberikan petunjuk penggunaan sistem informasi, tetapi pengguna lebih memilih bertanya langsung ke *customer service* daripada membaca atau melihat dokumentasi yang sudah disediakan. Penyediaan FAQ merupakan salah satu cara untuk membantu pengguna apabila menemui kesulitan. FAQ memiliki banyak informasi karena banyaknya pertanyaan. Ini dapat membuat pengguna bingung dan memerlukan waktu dalam mencari informasi. untuk mengatasi hal tersebut , dibutuhkan media

untuk membantu *customer service* dalam menjawab pertanyaan secara otomatis yaitu menggunakan Chatbot. Chatbot telah banyak digunakan dalam kehidupan sehari misalnya polling, game, dll. Untuk memanfaatkan chatbot yang dapat menanggapi pertanyaan secara dinamis diperlukan pemrosesan bahasa alami. Oleh sebab itu diperlukan algoritma yang mampu untuk memproses kata dan mencocokkan kalimat dalam database.

Penelitian yang relevan tentang algoritma *stemming*, menurut penelitian Wahyudi (2017) dengan judul “implementasi dan analisis algoritma stemming nazief & Adriani dan porter pada dokumen berbahasa indonesia” Algoritma *stemming* Nazief & Adriani memiliki hasil yang lebih bagus dibandingkan dengan algoritma *stemming* porter. Namun waktu proses, algoritma porter lebih baik dibandingkan dengan algoritma Nazief & Adriani. Menurut Suhada (2017) pada penelitian dengan judul “Implementasi Algoritma Rabin Karp Dan Stemming Najief Andriani Untuk Deteksi Plagiarisme Dokumen” Algoritma *stemming* Nazief & Adriani sangat cocok di terapkan dalam proses pencarian kata dasar dalam Bahasa Indonesia berdasarkan kamus kata dasar.

Penelitian yang relevan tentang metode pencocokan, menurut penelitian Gunawan (2018) dengan judul “The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents” metode *cosine similarity* memiliki hasil yang bagus untuk mencocokkan 2 dokumen. Menurut Lahitani (2016) dengan judul “Cosine Similarity to Determine Similarity Measure: Study Case in Online Essay Assessment” metode *cosine similarity* dapat memberikan nilai kemiripan dalam pencocokan penilaian esai secara daring.

Dengan masalah dan penelitian yang relevan diatas peneliti ingin menerapkan algoritma *stemming* nazief & adriani dan metode *cosine similarity* untuk dapat digunakan pada chatbot telegram yang terintegrasi dengan elayanan. Tujuan dari penelitian ini adalah: (1) Membuat chatbot yang bisa menjawab pertanyaan secara otomatis, (2) Menghitung tingkat akurasi kesesuaian respon chatbot terhadap pertanyaan yang masuk dengan algoritma steaming Nazief & Adriani dan metode kemiripan menggunakan *cosine similarity* .

II. TINJAUAN PUSTAKA

A. Algoritma Nazief & Adriani

Algoritma Nazief & Adriani pertama kali dikembangkan oleh Bobby Nazief & Mirna Adriani. Algoritma ini didasarkan pada aturan morfologi Bahasa Indonesia yang luas, yang dikumpulkan menjadi satu kelompok dan di enkapsulasi menjadi imbuhan yang diizinkan dan imbuhan yang tidak diizinkan. Terdapat kamus kata dasar yang digunakan untuk mendukung perekaman kata dan pencocokan kata setelah dilakukan stemming kata.

Proses stemming pada algoritma Nazief & Adriani sebagai berikut:

- 1) Sebelum proses selanjutnya kata akan dicocokkan pada database kata dasar. Apabila ada yang cocok maka algoritma berhenti.
- 2) Hapus akhiran (“-lah”, “-kah”, “-tah” atau “-pun”) jika itu adalah partikel (“-lah”, “-kah”, “-tah” atau “-pun”) langkah ini diulangi lagi untuk menghilangkan inflectional possessive pronoun suffixes (“-ku”, “-mu” atau “-nya”). Periksa kata dalam kamus kata dasar, jika ditemukan, algoritma berhenti.
- 3) Hapus Derivational Suffix (“-i” atau “-an”). Periksa kata dalam kamus kata dasar. Jika kata tersebut ditemukan dalam kamus kata dasar, maka algoritma berhenti. Jika tidak ditemukan maka proses lanjut ke langkah 3a:
 - a. Jika ada akhiran “-an” dan huruf terakhir adalah “-k” maka “-k” akan dihapus. Periksa kata dalam kamus kata dasar. Jika kata tersebut ditemukan dalam kamus maka proses berhenti. Jika tidak, maka lanjut ke langkah 3b.
 - b. Akhiran yang dihapus (“-i”, “-an”, atau “-kan”) tidak jadi dilakukan menghapus imbuhan dan lanjut ke langkah 4.
- 4) Hapus Derivational Prefix (“be-”, “di-”, “ke-”, “me-”, “pe-”, “se-” dan “te-”). Periksa kata dalam kamus kata dasar. Jika kata tersebut ditemukan dalam database kata dasar, maka proses berhenti. Pada tahapan ini berhenti jika memenuhi kondisi berikut:
 - a. Ditemukan kombinasi awalan dan akhiran yang tidak diizinkan
 - b. Awalan yang ditemukan sama dengan awalan yang dihapus sebelumnya.
 - c. Menghilangkan tiga awalan
- 5) Jika proses sudah dilakukan tetapi kata dasar tidak ditemukan pada database kamus kata dasar, maka kata asli yang belum dilakukan stemming akan dikembalikan.

Dalam penelitiannya Asian, dkk (Asian et al., 2005) melakukan beberapa pengembangan algoritma Nazief & Adriani sebagai berikut:

- 1) Menggunakan kamus kata yang lebih lengkap

- 2) Menambahkan aturan-aturan untuk kata-kata majemuk perulangan.
- 3) Menambahkan aturan awalan dan akhira, serta aturan lainnya, yaitu:
 - a. Menambahkan perikel (inflection suffix) “-pun”.
 - b. Penambahan aturan pemenggalan awalan.
 - c. Perubahan aturan pemnggalan untuk tipe awalan “me”.
- 4) Perubahan urutan proses stemming, yaitu:
 - a. Kata dengan awalan “be-” dan akhiran “-lah”, hilangkan awal terlebih dahulu kemudian akhiran.
 - b. Kata dengan awalan “be-” dan akhiran “-an”, hilangkan awalan terlebih dahulu kemudian akhiran.
 - c. Kata dengan awalan “me-” dan akhiran “-i”, hilangkan awalan terlebih dahulu kemudian akhiran.
 - d. Kata dengan awalan “di-” dan akhiran “-i”, hilangkan awalan terlebih dahulu kemudian akhiran.
 - e. Kata dengan awalan “pe-” dan akhiran “-i”, hilangkan awalan terlebih dahulu kemudian akhiran.
 - f. Kata dengan awalan “ter-” dan akhiran “-i”, hilangkan awalan terlebih dahulu kemudian akhiran.

B. Stemming

Steamming adalah proses terkandung dalam sistem IR (Information Retrieval) yang mengubah kata-kata yang terkandung dalam dokumen menjadi kata dasar menggunakan aturan tertentu. Sebagai contoh, kata membawa, dibawa, membawakan jika dilakukan stemming maka kata dasarnya yaitu “sama”. Proses stemming pada kata bahasa Indonesia berbeda dengan stemming pada kata dalam bahasa inggris. Pada bahasa inggris proses yang diperlukan adalah menghilangkan sufiks saja, sedangkan pada bahasa Indonesia selain menghilangkan sufiks juga menghilangkan prefix dan juga konfiks.

C. Cosine Similarity

Ariantini (2016:1) cosine similarity digunakan dalam ruang positif. Sehingga cosine similarity membandingkan level kemiripan dokumen menggunakan konsep derajat cosine dimana nilai memiliki batas antara nilai 0 dan 1. Jika hasil cosine similarity nilainya 0 maka dikatakan tidak mirip. Jika hasil cosine similarity 1 maka dokumen tersebut dianggap mirip

$$\text{Similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$
$$\cos(\theta) = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \cdot \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Keterangan :

A = Pesan pengguna
B = FAQ pada database
 A_i = Bobot kata i dalam blok A_i
 B_i = Bobot kata i dalam blok B_i
 i = Jumlah kata dalam kalimat
 n = Jumlah vector

(Arintini 2016:2)

D. Chatbot

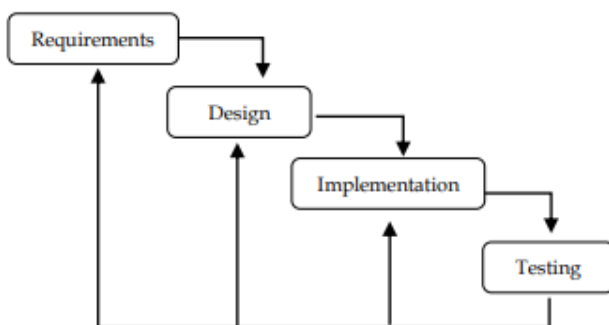
Chatbot adalah layanan yang mana seseorang dapat berinteraksi melalui antarmuka obrolan (chatting). Seseorang dapat berinteraksi menggunakan suara atau teks dan bisa juga mendapatkan respon yang sama yaitu menggunakan suara atau teks. Pada umumnya chatbot akan merespon dalam gaya percakapan, dan chatbot dapat melakukan tindakan sebagai respon dari percakapan yang ditanyakan atau yang diperintahkan. Aplikasi messaging yang sering digunakan untuk membuat Chatbot antara lain Facebook Messenger, Telegram, Whatsapp, Slack or SMS. Terdapat 2 jenis Chatbot.

1. Chatbot yang dapat beroperasi berdasarkan sekumpulan aturan yang sudah dibuat sebelumnya. Chatbot dapat merespon perintah yang spesifik. Jika ada seseorang melakukan perintah yang tidak ada dalam aturan yang disediakan, maka chatbot tidak tau maksud dari perintah orang tersebut.
2. Chatbot yang menggunakan machine learning dan artificial intelligence untuk menyediakan respon yang baik.

(Cherryleaf 2017)

III. METODOLOGI PENELITIAN

Berikut ini merupakan alur jalannya penelitian untuk melakukan penelitian “algoritma stemming nazief & adriani dengan metode *cosine similarity* untuk chatbot telegram terintegrasi dengan layanan” secara umum digambarkan pada gambar 1, sebagai berikut:



Gambar 1. Alur Penelitian

Gambar 1 membahas tentang alur penelitian algoritma stemming nazief & adriani dengan metode cosine similarity untuk chatbot telegram terintegrasi dengan layanan. Tahapan yang dilakukan pada penelitian ini yaitu analisis kebutuhan dengan melakukan studi pustaka untuk mencari literatur yang berkaitan untuk dapat dijadikan referensi pendukung dalam

penelitian ini. Literatur yang digunakan berhubungan dengan algoritma nazief & adriani dan metode cosine similarity. Kemudian membuat desain aplikasi chatbot menggunakan bantuan layanan telegram. Pada tahap implementasi, penelitian ini chatbot ditulis menggunakan bahasa PHP. Pada tahap pengujian sistem dilakukan untuk mengetahui kecocokan respon jawaban dari pertanyaan pengguna oleh bot.

A. Pengumpulan data

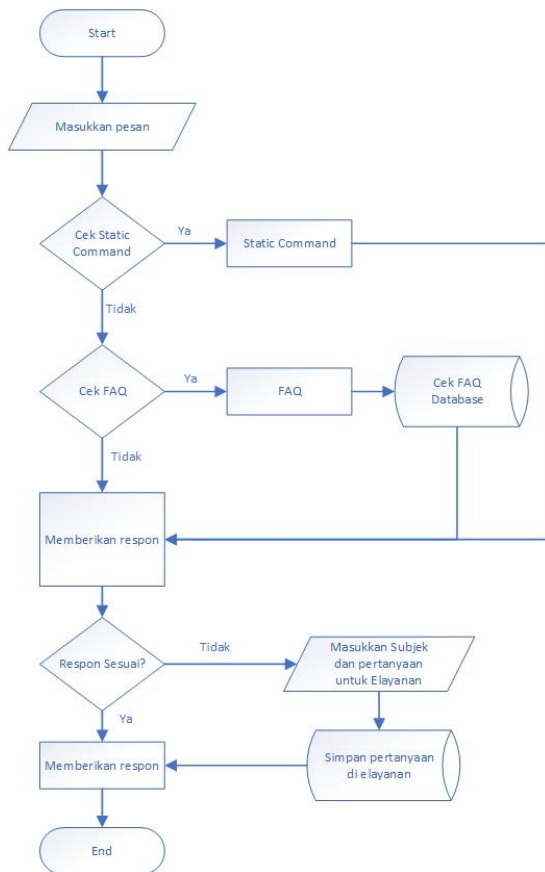
Pada tahap ini penulis melakukan studi literatur terkait teori stemming menggunakan algoritma stemming Nazief & Adriani dan metode kemiripan *Cosine Similarity*. Penelitian ini ditekankan untuk mengimplementasikan dan menganalisis teori stemming algoritma Nazief & Adriani dan metode kemiripan *Cosine Similarity* pada suatu chatbot.

Dalam penelitian ini, penulis menggunakan kamus kata dasar Indonesia yang diunduh dari internet dengan jumlah 28.530 kata. Kata dasar tersebut disimpan dalam database yang akan digunakan dalam proses *stemming*. Penulis juga menggunakan stopword list Tala yang dimuat pada penelitian Asian 2007.

Daftar pertanyaan atau FAQ yang akan menjadi respon dari bot didapatkan dari hasil pengelompokan pertanyaan yang sudah diajukan pada layanan Universitas Negeri Surabaya terkait dengan pertanyaan terkait SIAKADU (Sistem Informasi Akademik Terpadu) pada tanggal 1 Januari 2020 hingga 4 April 2020. Terdapat 15 pertanyaan yang sering diajukan pada layanan terkait Sistem Informasi Akademik Terpadu Universitas Negeri Surabaya.

B. Desain sistem

Dalam sistem ini menggunakan bantuan layanan aplikasi Telegram sebagai *user interface* untuk menampilkan respon dari proses algoritma nazief & adriani dengan metode *cosine similarity*. Alur sistem dijelaskan pada gambar 2. *flow chart* sebagai berikut :



Gambar 2 Alur Desain sistem

Pada Gambar 2. menampilkan tentang alur desain sistem algoritma stemming nazief & adriani dengan metode *cosine similarity* untuk *chatbot* telegram terintegrasi dengan layanan secara umum. Untuk memulai proses pengguna mengirimkan pesan melalui chatbot menggunakan layanan aplikasi Telegram. Selanjutnya bot akan memproses kalimat yang masuk dan merespon atau memberi balasan dari pertanyaan yang diajukan pengguna.

C. Rancang aplikasi

Penelitian ini mengimplementasikan algoritma nazief & adriani dengan metode *cosine similarity* menggunakan bahasa pemrograman PHP dan database PostgreSQL untuk membuat sebuah *website* dan *bot*. Website digunakan untuk memasukkan FAQ sedangkan *bot* akan memberikan respon atau jawaban dari pertanyaan yang masuk dengan bantuan layanan aplikasi Telegram.

D. Ujicoba produk

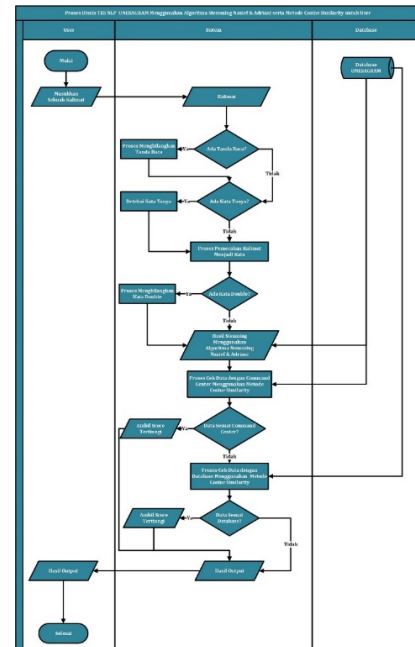
Ujicoba pada penelitian ini dilakukan oleh 15 pengguna layanan aplikasi telegram mahasiswa Universitas Negeri Surabaya. 15 pengguna tersebut mengajukan pertanyaan pada bot yang berkaitan dengan Sistem Informasi Akademik Terpadu (SIKADU).

IV. HASIL DAN PEMBAHASAN

Bab ini membahas tentang hasil penelitian dan pengujian terhadap implementasi yang dilakukan serta analisis algoritma yang digunakan. Pengujian dilakukan untuk menilai tingkat kesesuaian jawaban yang dihasilkan oleh *bot* dari pertanyaan yang diajukan yang di proses menggunakan algoritma stemming nazief & adriani dengan metode kemiripan *cosine similarity*.

A. Flowchart

Dibawah ini merupakan *flowchart* pemrosesan *bot*



Gambar 3. Flowchart pemrosesan bot

Pemrosesan kalimat yang masuk akan melalui beberapa tahap, yaitu:

1. Case Folding dan Tokenizing

Pada tahap *Case folding* ini kalimat akan diubah menjadi beberapa proses. Proses yang pertama kalimat akan diubah menjadi huruf kecil semua (lowercase). selanjutnya hanya huruf a sampai z yang diterima. Karakter selain huruf akan dihilangkan. Sedangkan *tokenizing* merupakan proses pemisahan kalimat kedalam bentuk kata yang dijadikan sebagai data, pada tahap ini penulis memisahkan kalimat menggunakan array.

2. Filtering

Pada tahap ini hasil dari *tokenizing* akan menghilangkan kata-kata double atau perulangan kata. Selanjutnya mengambil kata-kata yang dianggap penting dengan mengeliminasi kata-kata yang tidak deskriptif dari kamus stopwords list serta mengidentifikasi kata tanya.

3. Stemming

Dari hasil *filtering* proses selanjutnya itu mencari kata dasar. Pada proses ini algoritma nazief & adriani diimplementasikan untuk mencari kata dasar.

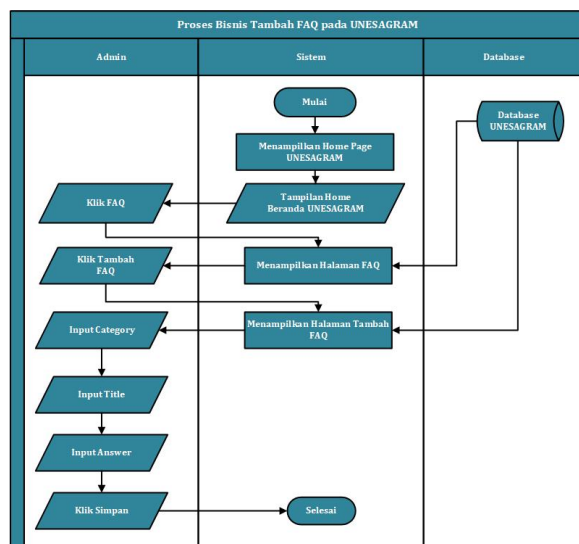
4. Analizing

Pada tahap ini metode cosine similarity diimplementasikan untuk memberi penilaian terhadap kalimat yang masuk dan dilakukan pencocokan dengan static command bot dan/atau dengan database FAQ yang sudah disediakan.

Static command akan menjadi prioritas utama pencocokan dari pada proses pencocokan. Apabila tidak ditemukan kecocokan pada static command proses akan lanjut melakukan pencocokan dengan database FAQ.

Hasil skor yang paling tinggi akan dijadikan sebagai jawaban atau respon oleh bot. Apabila skor sama dengan 0 artinya tidak ada yang cocok antara pertanyaan dengan static command dan/atau database FAQ yang sudah disediakan sehingga bot akan merespon dengan default jawaban. Selanjutnya bot menawarkan untuk mengirimkan ke elayanan Universitas Negeri Surabaya. Apabila pengguna menjawab iya maka menggunakan dapat mengajukan pertanyaan di elayanan. Apabila pengguna menjawab tidak maka percakapan berhenti.

Dibawah ini adalah flowchart tambah data FAQ,



Gambar 4. Proses pembuatan FAQ

Pada proses memasukkan FAQ, pertanyaan yang akan disimpan akan melalui proses, sebagai berikut:

1. Case Folding dan Tokenizing

Pada tahap *Case folding* ini prosesnya sama dengan proses *Case Folding* dan *Tokenizing* yang dilakukan pada pemrosesan bot diatas, yaitu kalimat akan diubah menjadi beberapa proses. Proses yang pertama kalimat akan diubah menjadi huruf kecil semua (lowercase). selanjutnya hanya huruf a sampai z yang diterima. Karakter selain huruf akan dihilangkan. Sedangkan *tokenizing* merupakan

proses pemisahan kalimat kedalam bentuk kata yang dijadikan sebagai data, pada tahap ini penulis memisahkan kalimat menggunakan array.

2. Filtering

Pada tahap ini hasil dari *tokenizing* akan menghilangkan kata-kata double atau perulangan kata. Selanjutnya mengambil kata-kata yang dianggap penting dengan mengeliminasi kata-kata yang tidak deskriptif dari kamus stopwords list serta mengidentifikasi kata tanya.

3. Stemming

Dari hasil *filtering* proses selanjutnya itu mencari kata dasar. Pada proses ini algoritma nazief & adriani diimplementasikan untuk mencari kata dasar.

4. Store Database

Menyimpan hasil dari berupa tags hasil dari case folding, filtering dan stemming dalam database yang nantinya akan digunakan dalam pencarian atau pencocokan dari pertanyaan yang masuk melalui aplikasi telegram.

B. Implementasi Database

Tabel 1. Tabel FAQ

faq		
name	type	length
id	uuid	0
articles	varchar	255
tags	text	0
categories	varchar	255
title	varchar	255
description	text	0
answer	text	0
questionmark	varchar	100
userid_created	varchar	255
userid_updated	varchar	255
created_at	timestamp	6
updated_at	timestamp	6

Tabel 2. Tabel Katadasar

Kata Dasar		
name	type	length
id_ktdasar	int4	32
katadasar	varchar	50
tipt_katadasar	varchar	20

C. Implementasi Sistem

Pada penelitian ini menggunakan FAQ atau daftar pertanyaan atau FAQ yang akan menjadi respon dari bot didapatkan dari hasil pengelompokan pertanyaan yang sudah diajukan pada elayanan Universitas Negeri Surabaya terkait dengan pertanyaan terkait SIAKADU (Sistem Informasi Akademik Terpadu) pada tanggal 1

Januari 2020 hingga 4 April 2020. Terdapat 15 pertanyaan yang sering diajukan pada elayanan terkait Sistem Informasi Akademik Terpadu Universitas Negeri Surabaya.

Berikut pertanyaan yang sering muncul pada elayanan :

Tabel 3. Tabel pertanyaan

No	Pertanyaan
1	Penghapusan matakuliah
2	Perubahan/ perbaikan nilai
3	Input/ entri/ upload nilai
4	NIK tidak sesuai/ salah
5	Permohonan/ pengajuan matakuliah krs
6	Akun/ email tidak bisa login/ akses/ log in/ masuk siakadu
7	Pengajuan/ pengaktifan akun siakad
8	Perbaiki transkrip nilai
9	Penambahan matakuliah
10	Perbaiki data forlap dikti
11	Permohonan pindah jalur
12	Nama tidak sesuai/ salah
13	Membuka kunci nilai
14	Pengaktifan status mahasiwa
15	Tanggal keluar tidak sama dengan ijazah/ ijazah

Berikut ini adalah tampilan master FAQ:

1. Tampilan halaman daftar FAQ
Untuk menampilkan daftar FAQ yang sudah dimasukkan oleh admin

Category	Title	Tag	Action
SIADU	Perbaikan nilai mata kuliah Jawaban : Silahkan membuat surat yang ditujukan ke Wakil Rektor 1	baik, nilai, mata, kuliah,	2 1
SIADU	Penghapusan nilai mata kuliah Jawaban : Silahkan membuat surat yang ditujukan ke Wakil Rektor 1	hapus, nilai, mata, kuliah,	2 1
SIADU	Ganti jataul berrtok siakadu Jawaban : Silahkan menghubungi admin prodi	ganti, jataul, berrtok, siakadu,	2 1
SIADU	Lupa Password siakadu Jawaban : Login menggunakan akun SSO	lupa, password, siakadu,	2 1

Gambar 5. Daftar FAQ

2. Tampilan halaman tambah FAQ
Form yang digunakan untuk menambahkan skenario FAQ dengan memasukkan kategori pertanyaan, pertanyaan dan jawaban.

Gambar 6. Tambah FAQ

3. Tampilan halaman ubah FAQ
Form untuk mengubah scenario FAQ.

Gambar 7. Ubah FAQ

Berikut ini adalah tampilan halaman untuk simulasi pertanyaan dan hasil jawaban atau respon dari *bot*. Pada menu ini digunakan untuk memudahkan admin dalam mencoba skenario FAQ dengan memasukkan sebuah kalimat. Dalam menu ini akan terlihat hasil step-by-step kalimat di proses hingga mendapatkan respon atau jawaban.

Gambar 8. Simulasi respon bot 1

Hasil skor pada *command center* merupakan hasil penilaian dari kalimat yang masuk terhadap *static command bot*.

CHECK COMMAND CENTER	
Word	Score
hi	0
covid	0
presensi	0
berita	0
server	0
daftar	0
aktivasi	0
broadcast	0
dafona	0
elayanan	0
/start	0

Gambar 9. Simulasi respon bot 2

Hasil Skor pada *database FAQ* merupakan hasil penilaian dari kalimat yang masuk terhadap *database FAQ* yang sudah dibuat oleh admin.

CHECK DATABASE	
Word	Score
Penghapusan nilai mata kuliah	0.5
Bagaimana caramu menjawab pertanyaan?	0.5
Perbaikan nilai mata kuliah	0.25
Call Center PPTI	0
Kamu lagi apa?	0
wkwk	0
Kamu makan minum?	0
Robot atau manusia?	0
jadwal kapan pembayaran UKT?	0
Lupa Password siakadu	0

Gambar 10. Simulasi respon bot 3

Tahap akhir dari proses yaitu hasil respon atau jawaban yang dihasilkan bot hasil dari penilaian mulai dari static command bot hingga penilaian pada *database FAQ*.

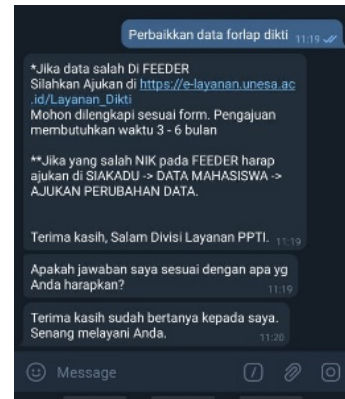
Hasil Output **Silahkan membuat surat yang ditujukan ke Wakil Rektor 1**

Gambar 11. Simulasi respon bot 4

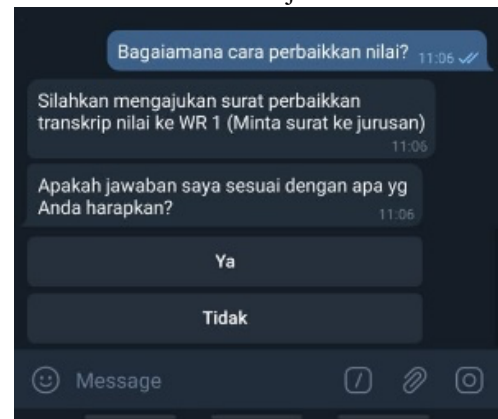
D. Pengujian Sistem

Pengujian sistem dilakukan oleh 15 pengguna aplikasi telegram mahasiswa Universitas Negeri Surabaya. 15 pengguna tersebut mengajukan pertanyaan pada bot yang berkaitan dengan Sistem Informasi Akademik Terpadu (SIKADU) yang didasarkan pada pertanyaan yang sering muncul pada elayanan Universitas Negeri Surabaya. Setiap pengguna bebas menggunakan kosa katanya masing-masing tidak terpaksa pada kalimat yang ada di *FAQ*.

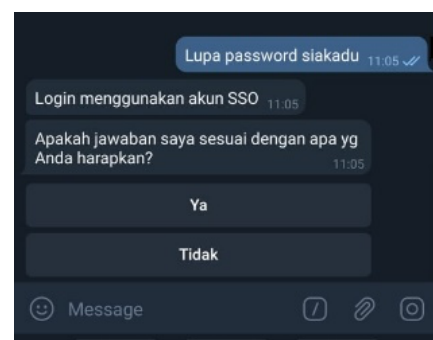
Dibawah ini merupakan contoh pertanyaan yang diajukan oleh pengguna pada bot melalui aplikasi telegram.



Gambar 12. Ujicoba 1

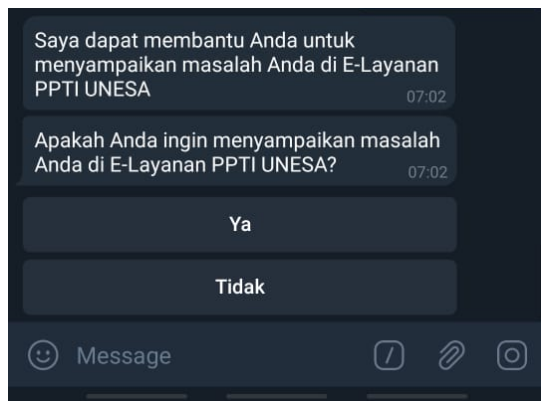


Gambar 13. Ujicoba 2



Gambar 14. Ujicoba 3

Apabila jawaban tidak sesuai maka akan ada pertanyaan untuk menyampaikan masalah ke E-layanan PPTI UNESA.



Gambar 15. Elayanan

V. KESIMPULAN & SARAN

A. Kesimpulan

Pengujian dilakukan dengan mengajukan pertanyaan secara acak. Setelah pengguna mengajukan pertanyaan dan dijawab oleh bot, bot akan kembali bertanya tentang respon yang dilakukan oleh bot. Dari ujicoba terdapat 236 pertanyaan yang masuk. Terdapat 217 respon yang menyatakan sesuai dengan harapan dan terdapat 19 respon yang menyatakan tidak sesuai. Dari hasil tersebut dapat dibuat persentase 91,95% menjawab sesuai dan 8,05 yang menjawab tidak sesuai.

Dari hasil tersebut dapat disimpulkan bahwa algoritma Nazief & Adriani dengan metode Cosine Similarity dapat digunakan untuk membantu menjawab pertanyaan secara otomatis dengan jawaban yang relatif sesuai dengan yang diharapkan pengguna. Dengan demikian cara ini merupakan cara yang efektif untuk membantu customer service dalam menjawab pertanyaan secara otomatis dan pengguna dalam mengajukan pertanyaan melalui chatbot.

B. Saran

Untuk penelitian selanjutnya dapat ditambahkan machine learning agar bot dapat mempelajari pertanyaan yang diajukan melalui chatbot dan dapat membaca bahasa gaul.

REFERENSI

- [1] Elcholiqi, Abidah dan Musdholifah, Aina. 2020. "Chatbot in Bahasa Indonesia Using NLP to Provide Banking Information". IJCCS (Indonesian Journal of Computing and Cybernetics Systems). Vol. 14 (1): hal 81-90.
- [2] Wahyudi, Dwi DKK. 2018. "Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia". Jurnal Ilmiah SINUS. hal 49-56.
- [3] Suhada, Satia dan Bahri Saeful. 2017. "Implementasi Algoritma Rabin Karp Dan Stemming Najief Andriani Untuk Deteksi Plagiarisme Dokumen". SWABUMI Vol.5: hal. 84-89.
- [4] D Gunawan, dkk. 2017. "The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents", 2nd International Conference on Computing and Applied Informatics. hal. 1-6.
- [5] Salim, Mohammad A dan Anistyasari, Yeni. 2017. "Pengembangan Aplikasi Penilaian Ujian Essay Berbasis Online

- Menggunakan Algoritma Nazief Dan Adriani Dengan Metode Cosine Similarity". Jurnal IT-EDU. Vol. 2 (1): hal. 126-135.
- [6] Lahitani, Alfira R, dkk. 2016. "Cosine similarity to determine similarity measure: Study case in online essay assessment". International Conference on Cyber and IT Service Management.
- [7] Benedictus, Ruspandi R, dkk. 2017. "Rancang Bangun Chatbot Helpdesk untuk Sistem Informasi Terpadu Universitas Sam Ratulangi", E-Journal Teknik Informatika. Vol. 11 (1)
- [8] Handoyo, Eko, dkk. 2018. "Proc. of 2018 5th Int. Conf. on Information Tech., Computer, and Electrical Engineering (ICITACEE)". hal. 325-330.
- [9] Mardiana, Tari, dkk. 2016. "Stemming Influence on Similarity Detection of Abstract Written in Indonesia", Jurnal TELKOMNIKA. Vol. 14 (1): hal. 219-227.
- [10] Prihatini PM, dkk. 2017. "Stemming Algorithm for Indonesian Digital News Text Processing", International Journal of Engineering and Emerging Technology. Vol. 2 (2): hal. 1-7.
- [11] Sari, Bunga dan Sibaroni, Yuliant. 2019. "Deteksi Kemiripan Dokumen Bahasa Indonesia Menggunakan Algoritma SmithWaterman dan Algoritma Nazief & Andriani". Ind. Journal on Computing. Vol. 4 (3): hal. 88-98.
- [12] F. B. Djafar, A. Lahinta and H. Lillyan, "Penerapan Algoritma Smith-Waterman Dalam Sistem Pendeteksi Kesamaan Dokumen".
- [13] M. Y. Rimaldo and S. "Sistem Pendeteksi Plagiat Dokumen Teks Menggunakan Algoritma SmithWaterman Serta Algoritma Arifin dan Setiono," 2012.
- [14] Ariantini, Dewa A, dkk. 2016. "Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity". Jurnal Inovtek Polbeng-Seri Informatika. E-Journal Teknik Informatika. Vol. 9 (1): hal. 1-8.
- [15] Deolika, Agatha, dkk. 2019. "Analisis Pembobotan Kata Pada Klasifikasi Text Mining". Jurnal Teknologi Informasi. Vol. 3 (2): hal. 179-184.
- [16] Diki Susandi, Pemanfaatan Vector Space Model pada Penerapan Algoritma Nazief Adriani, KNN dan Fungsi Similarity Cosine untuk Pembobotan IDF dan WIDF pada Prototipe Sistem Klasifikasi Teks Bahasa Indonesia. Jurnal ProTekInfo Vol. 3 No. 1 ISSN: 2406-7741, 2016.
- [17] Mandias, F. A., Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Di Universitas Klabat Dengan Metode Klasifikasi, Konferensi Nasional Sistem & Informatika 2015, STIKOM Bali, 9-10 Oktober 2015, 2015.
- [18] M. Syukri Mustafa, Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, Citec Journal, Vol. 4, No. 2 2460- 4259, 2017.
- [19] R. Shah, S. Lahoti, and Lavanya, "An Intelligence Chat-bot using Natural Language Processing," International Journal of Engineering Research, vol no. 6, Issue no. 5, pp: 281-286, 2017.
- [20] Ariantini, Dewi Ayu Rai. 2016. Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. Vol.9 ISBN:2301-8364: pp 1-8
- [21] Shutharland Smith. Plagiarism, the internal and student learning: Improving Academic integrity. New York: Routledge. 2014.
- [22] Manopo, R. I. 2016. "Perancangan Aplikasi Help Desk di UPT-TIK Unsrat", Tekno. vol. 8. no.1, 2016.
- [23] Gibson, Darril. 2015. Effective Help Desk Specialist Skills. USA: Pearson Education, Inc.
- [24] Dewi, A. & Setiaji, B. 2014. "Pemanfaatan Sentence Similarity Measurement untuk Proses Pencarian Pola pada Chatbot Berbasis Pattern-Matching," Seminar Nasional Teknologi Informasi dan Multimedia 2014, ISSN : 2302-3805, 8 Februari 2014.
- [25] Baiti, Zifora, Nugroho, Fresy. 2013. "Aplikasi Chatbot IM3 untuk Informasi Jurusan Teknik Informatika Berbasis Sistem Pakar menggunakan Forward Chaining", MATICS vol.5 no.3, September 2013.
- [26] Britz, Denny. 2016. "Deep Learning for Chatbots, Part 1 – Introduction" [Online]. Available: <http://www.wildml.com/2016/04/deep-learning-forchatbots-part-1-introduction>. Diakses tanggal 12 Juni 2017