

QUESTION ANSWERING SYSTEM BERBASIS CHATBOT PADA PLATFORM LINE UNTUK MATA PELAJARAN SEJARAH INDONESIA SMA DENGAN MENGUNAKAN METODE SENTENCE SIMILARITY MEASUREMENT

CHATBOT-BASED QUESTION ANSWERING SYSTEM ON THE LINE PLATFORM FOR HIGH SCHOOL INDONESIAN SUBJECTS USING THE SENTENCE SIMILARITY MEASUREMENT METHOD

Radita Najmi Kamilia¹, Oktariani Nurul Pratiwi², Riska Yanu Fa'rifah³.

^{1,2,3} Universitas Telkom, Bandung

¹ditakamilia@student.telkomuniversity.ac.id, ²onurulp@telkomuniversity.ac.id,

³riskayanu@telkomuniversity.ac.id

Abstrak

Pembangunan aplikasi yang berjudul "Question Answering System Berbasis Chatbot Pada Platform LINE Untuk Mata Pelajaran Sejarah Indonesia SMA Dengan Menggunakan Metode Sentence Similarity Measurement" untuk membangun sebuah aplikasi chatbot pada platform LINE yang diharapkan dapat membantu pelajar SMA dalam mencari jawaban mengenai mata pelajaran sejarah Indonesia yang dianggap kurang efektif dan efisien ketika mencari pada buku pelajaran dan search engine. Pertanyaan yang diinputkan oleh pengguna akan diproses dengan text preprocessing untuk mendapatkan keyword pada database. Untuk mendapatkan respon atau jawaban yang tepat dan benar sesuai dengan pertanyaan yang diinputkan dari pengguna, maka chatbot dibangun menggunakan metode Sentence Similarity Measurement dan Bigram. Kelebihan dari menggunakan metode Similarity Sentence Measurement yaitu metode ini dapat bekerja dengan baik dari pengumpulan, lalu menyatukan informasi, kemudian menemukan kesimpulan yang dapat diambil dari informasi tersebut. Pengujian evaluasi pada penelitian ini melakukan perhitungan pola algoritma dimana memperoleh nilai 0,86. Kemudian melakukan tanya jawab pada chatbot oleh responden sehingga memperoleh hasil akurasi sebesar 80%. Pada penelitian ini dapat disimpulkan bahwa chatbot dapat melakukan tanya jawab dan membantu siswa dalam memberikan jawaban secara tepat dan benar mengenai mata pelajaran sejarah Indonesia dengan menggunakan metode Sentence Similarity Measurement dan Bigram.

Kata Kunci: Question Answering System, Chatbot, Metode Sentence Similarity Measurement, LINE, mata pelajaran sejarah Indonesia

Abstract

Development of an application entitled "Chatbot-Based Question Answering System on the LINE Platform for High School Indonesian History Subjects Using the Sentence Similarity Measurement Method" to build a chatbot application on the LINE platform which is expected to assist high school students in finding answers about Indonesian history subjects considered to be less effective and efficient when searching on textbooks and search engines. Questions entered by the user will be processed with text preprocessing to get keywords in the database. To get the right and correct response or answer according to the questions entered from the user, the chatbot was built using the Sentence Similarity Measurement and Bigram methods. The advantage of using the Similarity Sentence Measurement method is that this method can work well from collecting, then uniting information, then finding conclusions that can be drawn from that information. The evaluation test in this study calculated the algorithm pattern which obtained a value of 0.86. Then do a question and answer on the chatbot by the respondent so as to obtain an accuracy result of 80%. In this study, it can be concluded that chatbots can conduct questions and answers and assist students in providing correct and correct answers regarding Indonesian history subjects using the Sentence Similarity Measurement and Bigram methods.

Keyword: Question Answering System, Chatbot, Sentence Similarity Measurement Method, LINE, Indonesian history subjects

1. Pendahuluan

Kemajuan teknologi informasi memberikan dampak positif di berbagai bidang, salah satunya bidang pendidikan. Pendidikan adalah menjadi salah satu pondasi pembangunan nasional. Semakin tinggi kualitas pendidikan suatu negara, maka semakin tinggi pula kualitas bangsanya [1]. Dalam pendidikan terdapat beberapa mata pelajaran salah satunya yaitu mata pelajaran sejarah Indonesia. Mata pelajaran sejarah Indonesia memiliki posisi yang amat penting dalam kurikulum 2013 untuk kehidupan berbangsa dan bernegara [2]. Menurut hasil kerja dari History Working Group Hunt (2007:7) menjelaskan bahwa ada sembilan tujuan dari pembelajaran mata pelajaran sejarah Indonesia yang ada di sekolah, beberapanya yaitu (1) membangkitkan minat dan pengetahuan masa lalu, (2) memberikan dan menanamkan identitas kebangsaan pada siswa, (3) memberikan pemahaman mengenai akar dan warisan budaya negara, dan (4) mempersiapkan siswa menuju kehidupan dewasa [3].

Namun, dalam pelaksanaannya dalam pembelajaran masih banyak siswa yang mendapatkan nilai rendah [4]. Hariyono (2017) juga mengatakan ada beberapa faktor dampak yang akan membuat siswa dapat nilai rendah pada mata pelajaran sejarah, salah satunya yaitu siswa dalam mengikuti pelajaran sejarah umumnya mengalami banyak kesulitan dalam memahami pelajaran sejarah. Siswa hanya memanfaatkan buku catatan dan modul yang diberikan guru saat belajar di rumah sehingga untuk siswa yang tidak aktif bertanya dan mencatat akan merasa kesulitan saat belajar di rumah dikarenakan tidak memahami materi yang ada [5]. Padahal sejatinya mata pelajaran sejarah Indonesia mendapat amanah untuk membentuk karakter peserta didik lewat nilai-nilai yang terkandung di dalamnya. Untuk memudahkan siswa dan guru dalam mencari informasi terkait mata pelajaran sejarah Indonesia maka dibutuhkan sistem *Question Answering System* atau QAS berbentuk *Chatbot*. Berdasarkan permasalahan di atas, maka dibutuhkan suatu media pembelajaran interaktif yang diharapkan menjadi salah satu alat bantu yang dapat membantu siswa dalam kegiatan belajar yaitu *Chatbot Question Answering System* (QAS). *Chatbot Question Answering System* atau *Chatbot QAS* ini berbasis *platform* yang berjalan pada salah satu media sosial *messenger* yang banyak digunakan oleh masyarakat di Indonesia yaitu LINE. Metode yang digunakan yaitu *Similarity Sentence Measurement* atau SSM. SSM atau *Sentence Similarity Measurement* merupakan perhitungan sebuah tingkat dari kemiripan atau kecocokan antara dua kalimat atau dua string. Hasil dari yang didapatkan dari perhitungan *similarity* adalah berupa bilangan *real* 0 sampai 1 [6].

2. Landasan Teori

2.1 Question and Answering System

Question Answering System merupakan sebuah sistem yang dapat memahami tujuan dan arti dari sebuah pertanyaan yang diberikan oleh Pengguna dengan menggunakan bahasa yang alami dan merespon pertanyaan tersebut berdasarkan informasi yang ada [7]. Sistem QAS (*Question Answering System*) mengijinkan *user* untuk menginputkan pertanyaan dalam bahasa *natural*, yaitu bahasa yang digunakan dalam percakapan sehari-hari, dan memperoleh jawaban dengan cepat serta ringkas [8].

2.2 Chatbot

Chatbot merupakan salah satu kecerdasan yang dibuat dan dirancang sebagai program untuk berinteraksi secara langsung dengan manusia. Yang menjadi pembeda *Chatbot* dengan *Natural Language Processing System* atau NLP ialah kesederhanaan Algoritma yang digunakan dalam membangunnya [9]. Penyajian materi pembelajaran yang menarik dapat dimanfaatkan oleh *Chatbot* pada bidang pendidikan sebagai suatu pengembangan media untuk pembelajara siswa sehingga dapat membantu siswa dalam pembelajaran.

2.3 LINE

LINE messenger merupakan sebuah aplikasi yang dapat melakukan pesan kirim gratis secara instan yang dapat digunakan pada *platform* seperti tablet, telepon cerdas, hingga komputer. LINE messenger hanya dapat berfungsi ketika terhubung dengan jaringan internet sehingga pengguna LINE dapat melakukan aktivitas seperti mengirim pesan teks, mengirim gambar, video, pesan suara, dan lain lain [10].

2.4 Text Preprocessing

Text preprocessing merupakan suatu proses mengubah sebuah bentuk data struktural yang sebelumnya belum terstruktur menjadi data yang terstruktur [11]. Preprocessing juga dapat dikatakan proses pengurangan kata yang tidak penting dan tidak memiliki arti dari dokumen atau teks, sehingga membuat data menjadi lebih terstruktur dan siap diolah [12]. Pada penelitian ini menggunakan beberapa *preprocessing*, yaitu *case folding*, *tokenizing*, *filtering* (*stopword removal*), dan *stemming*.

- a. *Case Folding* merupakan proses untuk mengubah semua huruf menjadi huruf kecil, menghilangkan semua tanda baca seperti koma (,), titik (.), tanda tanya(?) dan lain sebagainya [11].

Kalimat: Apa yang dimaksud dengan Reformasi?

Case Folding : apa yang dimaksud dengan reformasi

- b. *Tokenizing* merupakan proses memisahkan kalimat menjadi perkata [11].
Kalimat: Apa yang dimaksud dengan Reformasi
Tokenizing: 'Apa', 'yang', 'dimaksud', 'Reformasi'
- c. *Filtering (Stopword Removal)* merupakan proses mengurangi jumlah suatu kata dengan menghilangkan kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words* [11].
Kalimat: Apa yang dimaksud dengan Reformasi
Filtering (Stopword Removal): 'dimaksud', 'Reformasi'
- d. *Stemming* merupakan proses menghilangkan kata imbuhan yang terdapat pada kalimat atau suatu kata dan mengubah menjadi kata dasar [11].
Kalimat : Apa yang dimaksud dengan Reformasi
Stemming : 'maksud', 'Reformasi'

2.5 Sentence Similarity Measurement

SSM atau Sentence Similarity Measurement merupakan perhitungan sebuah tingkat dari kemiripan atau kecocokan antara dua kalimat atau dua string. (Dewi and Setiaji 2014). Hasil dari yang didapatkan dari perhitungan similarity adalah berupa bilangan real 0 sampai 1 [6].

Dengan menggunakan bigram maka dapat dibuat sebuah rumus sederhana yang dapat digunakan untuk menghitung sebuah nilai/skor. Berikut ini adalah rumus yang digunakan:

$$\frac{\text{count}(b1 \in b2) + \text{count}(b2 \in b1)}{\text{count}(b1) + \text{count}(b2)}$$

- a. $b1$ merupakan himpunan bigram *string* pertama, dan $b2$ merupakan himpunan bigram *string* kedua.
- b. $\text{count}(b1 \in b2)$ merupakan jumlah himpunan dari "setiap bigram yang ada di $b1$ merupakan anggota bigram di $b2$ ". Sama dengan sebaliknya untuk $\text{count}(b2 \in b1)$.
- c. $\text{count}(b1)$ merupakan jumlah himpunan bigram yang terdapat pada *string* pertama, dan $b2$ merupakan jumlah himpunan bigram yang terdapat pada *string* kedua.

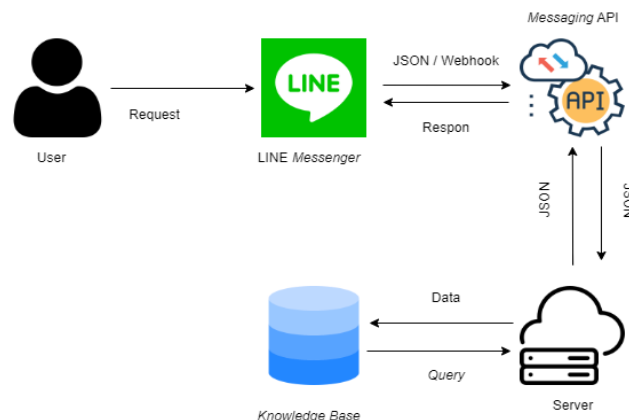
2.6 Mata Pelajaran Sejarah

Mata pelajaran sejarah Indonesia menduduki salah satu posisi yang penting dalam kurikulum 2013 untuk kehidupan masyarakat berbangsa dan bernegara [2]. 1 Karena memiliki nilai-nilai di dalamnya, mata pelajaran sejarah Indonesia mendapatkan amanah dalam membentuk karakter peserta didik. Sejarah termasuk golongan ilmu sosial karena sejarah mempelajari perilaku sosial di sekitar kita atau di masyarakat [2].

3. Hasil dan Pembahasan

3.1 Perancangan Sistem

3.1.1 Arsitektur Aplikasi



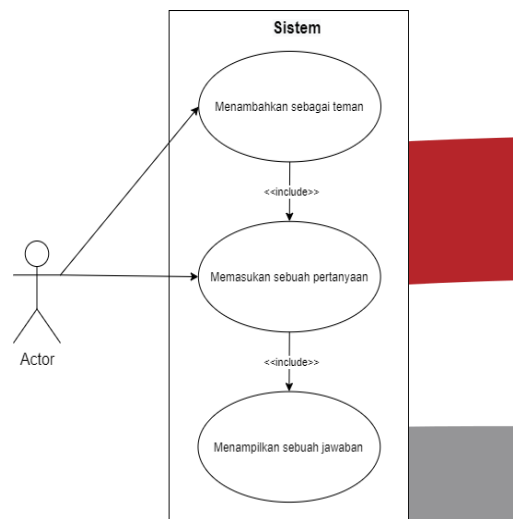
Gambar 1 Arsitektur Aplikasi

Beberapa komponen penting yang ada pada Gambar 3 adalah sebagai berikut :

1. LINE Messenger sebagai *Chat Channel Service* yang berfungsi sebagai *user* berinteraksi dengan *bot* untuk mencari informasi mengenai mata pelajaran sejarah.
2. DB Browser for SQLite sebagai *knowledge base* pada sistem *chatbot* yang berfungsi untuk menyimpan data berisikan jawaban mengenai mata pelajaran sejarah beserta kata kunci.
3. Server menggunakan *Cloud Platform Service Heroku* yang merupakan *Platform As a Service (PaaS)* yang berfungsi sebagai *web service* untuk mengelola informasi dan *knowledge base*.

3.1.2 Use Case

Use Case Diagram merupakan sebuah diagram yang dapat menjelaskan interaksi yang terjadi antar *user* dengan sistem.

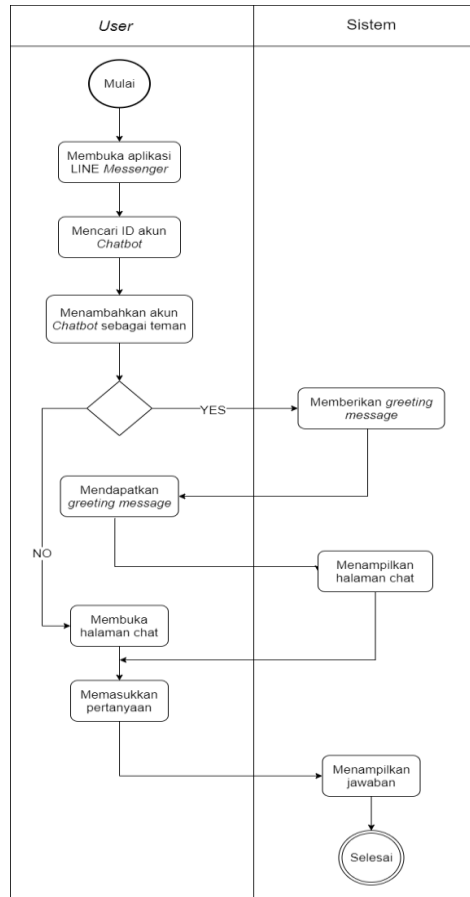


Gambar 2 Use Case Diagram

4

3.1.3 Activity Diagram

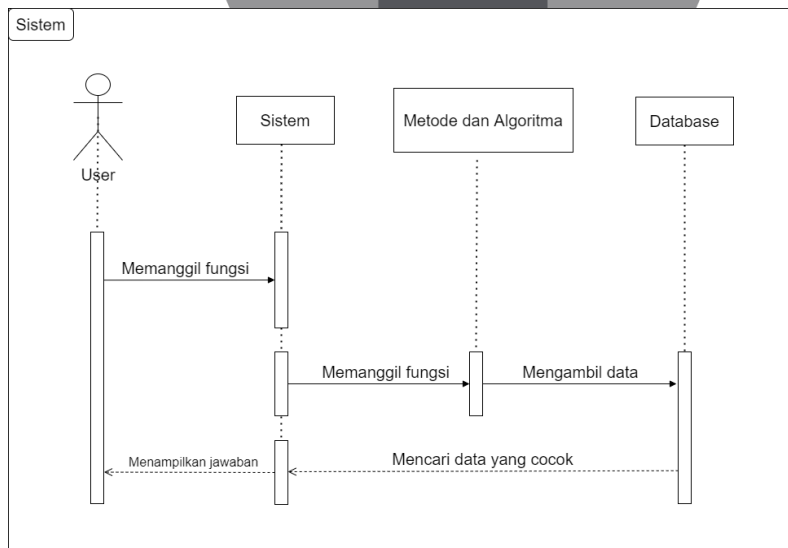
Activity Diagram merupakan diagram yang menggambarkan aktivitas dalam sebuah sistem yang sedang berjalan.



Gambar 3 Activity Diagram

3.1.4 Sequence Diagram

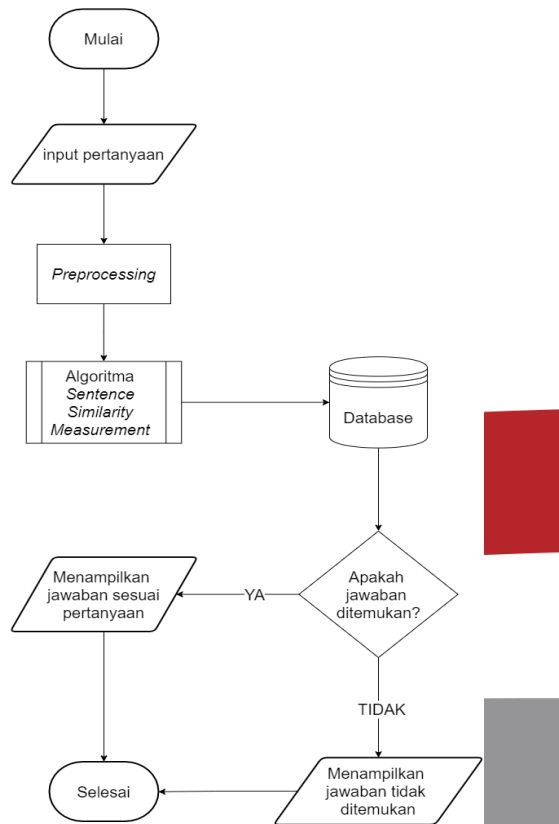
Sequence Diagram adalah suatu diagram yang menggambarkan interaksi antar objek pada sistem dalam sebuah rangkaian waktu.



Gambar 4 Sequence Diagram

3.1.5 Flowchart Sistem

Flowchart adalah bagan dengan simbol tertentu yang menggambarkan suatu proses dalam suatu sistem. Berikut merupakan flowchart pada chatbot. Yang dapat dilihat pada Gambar 7.



Gambar 5 Flowchart

3.2 Implementasi

Pada tahap ini yaitu menghitung pengulangan dan pencocokan pola *Sentence Similarity Measurement* dan bigram antara pertanyaan yang ada di database dan pertanyaan yang yang diberikan oleh pengguna. Terdapat empat tahap pada pencocokan pola algoritma:

- a. Tahap 1 : pengguna menginputkan pertanyaan pada aplikasi *Chatbot Question Answering System* dengan pertanyaan berisi “Apakah kemerdekaan negara Indonesia menjadi kado dari Jepang ?”
- b. Tahap 2 : pada tahap ini melakukan *Preprocessing*
 - *Case Folding*

Table 1 *Case Folding* Untuk Pencocokan Pola

Sebelum proses	Hasil proses
Apakah kemerdekaan negara Indonesia menjadi kado dari Jepang ?	apakah kemerdekaan negara indonesia menjadi kado dari jepang

Pada tabel *Case Folding* Untuk Pencocokan Pola ini melakukan proses *Preprocessing Case Folding* untuk menghilangkan tanda baca yaitu tanda tanya (?) dan merubah huruf kapital menjadi huruf kecil.

- *Tokenizing*

Table 2 *Tokenizing* Untuk Pencocokan Pola

Sebelum proses	Hasil proses
apakah kemerdekaan negara indonesia menjadi kado dari jepang	'apkah', 'kemerdekaan', 'indonesia', 'merupakan', kado, 'dari', 'jepang'

Pada tabel *Tokenizing* Untuk Pencocokan Pola melakukan proses *Preprocessing Tokenizing* untuk memisahkan kalimat menjadi sebuah kata perkata.

- *Filtering (Stopword Removal)*

Table 3 *Filtering (Stopword Removal)* Untuk Pencocokan Pola

Sebelum proses	Hasil proses
'apkah', 'kemerdekaan', 'indonesia', 'merupakan', kado, 'dari', 'jepang'	'kemerdekaan', 'indonesia', 'kado', 'jepang'

Pada tabel *Filtering (Stopword Removal)* Untuk Pencocokan Pola ini melakukan proses *Preprocessing Filtering (Stopword Removal)* untuk menghilangkan kata yang kurang penting yang terdapat pada daftar *stopword*. Daftar *stopword* dapat mencegah munculnya kalimat yang kurang penting sehingga dapat fokus mencari kata *keyword*. Daftar *stopword* terdapat pada file `ta_stopword_list.txt`.

- *Stemming*

Table 4 *Stemming* Untuk Pencocokan Pola

Sebelum proses	Hasil proses
'kemerdekaan', 'indonesia', 'kado', 'jepang'	'merdeka', 'indonesia', 'kado', 'jepang'

Pada tabel *Stemming* Untuk Pencocokan Pola ini melakukan proses *Preprocessing Stemming* untuk menghilangkan kata imbuhan yang terdapat pada sebuah kalimat.

- c. Tahap 3 : pertanyaan pengguna yang telah di *preprocessing* akan dilakukan proses mencari *keyword* pada database dengan cara menyeleksi dan mencocokkan kata yang terdapat pada *keyword*. Konteks pertanyaan terdapat *keyword* yang ada pada *database*

Table 5 Contoh data kalimat menjadi *keyword* untuk Database

Sebelum proses	Hasil proses <i>keyword</i>
Apakah kemerdekaan negara Indonesia menjadi kado dari Jepang ?	'merdeka', 'indonesia', 'kado', 'jepang'

Terdapat empat *keyword* untuk pertanyaan “Apakah kemerdekaan negara Indonesia menjadi kado dari Jepang ?” dengan *keyword* “merdeka”, “indonesia”, “kado”, dan “jepang”. Dengan begitu tabel di bawah ini adalah tabel beberapa pertanyaan diambil dari database yang memiliki

beberapa *keyword* yang sama dengan pertanyaan dari pengguna.

Table 6 Beberapa pertanyaan yang akan di uji

Pertanyaan	Keyrowd
Mengapa Jepang mengaku sebagai saudara tua Indonesia ?	['jepang', 'aku', 'saudara', 'tua', 'indonesia']
Bagaimana sambutan rakyat Indonesia terhadap kedatangan Jepang ?	['sambut', 'rakyat', 'indonesia', 'datang', 'jepang']
Apakah kemerdekaan Indonesia merupakan hadiah dari Jepang ?	['merdeka', 'indonesia', 'hadiah', 'jepang']

d. Tahap 4 : melakukan pencocokan pola SSM dan bigram. Apabila memiliki nilai yang besar dan cocok dengan keyword maka itu jawaban yang sesuai dengan pertanyaan yang diinputkan.

Metode SSM dan bigram menggunakan rumus :

$$\frac{count(b1 \in b2) + count(b2 \in b1)}{count(b1) + count(b2)}$$

- $b1$ merupakan himpunan bigram *string* pertama, dan $b2$ merupakan himpunan bigram *string* kedua.
- $count(b1 \in b2)$ merupakan jumlah himpunan dari “setiap bigram yang ada di $b1$ merupakan anggota bigram di $b2$ ”. Sama dengan sebaliknya untuk $count(b2 \in b1)$.
- $count(b1)$ merupakan jumlah himpunan bigram yang terdapat pada *string* pertama, dan $b2$ merupakan jumlah himpunan bigram yang terdapat pada *string* kedua.

3.3 Pengujian dan Evaluasi

Pengujian melakukan perhitungan pencocokan pola yang dilakukan oleh metode *Sentence Similarity Measurement* dan *Bigram* untuk menentukan jawaban yang sesuai dan dapat dikirimkan oleh sistem kepada Pengguna. Evaluasi yang dilakukan yaitu untuk menghitung hasil akurasi kinerja *Chatbot* dalam melakukan tanya jawab dengan Pengguna. Dalam melakukan evaluasi, Pengguna melakukan tanya pada *Chatbot* Sejarah BOT dengan pertanyaan-pertanyaan yang ditanyakan oleh Pengguna, kemudian *Chatbot* akan langsung menjawab sesuai dengan pertanyaan dan *keyword* yang ada di *database* kemudian dicocokkan oleh sistem.

a. Perhitungan Pertama

Perngguna : Bagaimana sambutan rakyat Indonesia terhadap kedatangan Jepang

Database : Apakah kemerdekaan Indonesia merupakan hadiah dari Jepang ?

Bigram1 pengguna : {B a, a g, g a, a i, i m, m a, a n, n a, a , s, s a, a m, m b, b u, u t, t a, a n, n , r, r a, a k, k y, y a, a t, t, I, i n, n d, d o, o n, n e, e s, s i, i a, a , t, t e, e r, r h, h a, a d, d a, a p, p , k k e, e d, d a, a t, t a, a n, n g, g a, a n, J, j e, e p, p a, a n, n g, g }

Bigram2 database : {a p, p a, a k, k a, a h, h , k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n , n, n e, e g, g a, a r, r a, a , i, i n, n d, d o, o n, n e, e s, s i, i a, a , m, m e, e n, n j, j a, a d, d i, i , h, h a, a d, d i, i a, a h, h , d, d a, a r, r i, i , j, j e, e p, p a, a n, n g, g }

Bigram1 ∈ Bigram2 : {p a, a k, k k, k e, e r, a n, n n, g a, r a, a, i, i n, n d, d o, o n, n e, e s, s i, i a, d a, a, j, j e, e p, p a, a n, n g, g }

Bigram2 ∈ Bigram1 : {g a, a n, a, r a, a k, i, i n, n d, d o, o n, n e, e s, s i, i a, a , e, r, a d, a p, h a, a d, k e, a n, n g, a n, n, j, j e, e p, p a, a n, n g, g }

Diperoleh hasil sebagai berikut :

Count Bigram 1 = 32

Count Bigram 2 = 31

Count Bigram1 \in Bigram2 = 63

Count Bigram2 \in Bigram1 = 62

$$\frac{\text{count}(b1 \in b2) + \text{count}(b2 \in b1)}{\text{count}(b1) + \text{count}(b2)}$$

$$\frac{32 + 31}{63 + 62} = \frac{63}{125} = 0,50$$

Count Bigram 1 = 59 didapat dari menghitung jumlah himpunan bigram yang terdapat pada *string* pertama. Sama dengan sebaliknya untuk Count Bigram 2.

Count Bigram1 \in Bigram2 = 51 didapat dari menghitung jumlah himpunan “setiap bigram yang ada di *b1* merupakan anggota bigram di *b2*”. Sama dengan sebaliknya untuk Count Bigram2 \in Bigram1.

b. Perhitungan Kedua

Perngguna : apakah kemerdekaan negara indonesia menjadi kado dari jepang ?

Database : Apakah kemerdekaan Indonesia merupakan hadiah dari Jepang ?

Bigram1 pengguna : {a p, p k, k a, a h, h , k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n , n, n e, e g, g a, a r, r a, a n, n , i, i n, n d, d o, o n, n e, e s, s i, i a, a , m, m e, e n, n j, j a, a d, d i, i , k, k a, a d, d o, o , d, d a, a r, r i, i , j, j e, e p, p a, a n, n g, g }

Bigram2 database : {a p, p a, a k, k a, a h, h , k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n , i, i n, n d, d o, o n, n e, e s, s i, i a, a , m, m e, e r, r u, u p, p a, a k, k a, a n, n , h, h a, a d, d i, i a, a h, h , d, d a, a r, r i, i , j, j e, e p, p a, a n, n g, g }

Bigram1 \in Bigram2 : a p, k a, a h, h , k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n , n, n e, a n, n , i, i n, n d, d o, o n, n e, e s, s i, i a, a , m, m e, a d, d i, i , k, k a, a d, d o, d, d a, a r, r i, i , j, j e, e p, p a, a n, n g, g }

Bigram2 \in Bigram1 : a p, p a, k a, a h , k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n , i, i n, n d, d o, o n, n e, e s, s i, i a, a , m, m e, e r, p a, k a, a n, , a d, d i, i a, h , d, d a, a r, r i, i , j, j e, e p, p a, a n, n g, g }

Diperoleh hasil sebagai berikut :

Count Bigram 1 = 59

Count Bigram 2 = 56

Count Bigram1 \in Bigram2 = 51

Count Bigram2 \in Bigram1 = 49

$51 + 49 / 59 + 56 = 0,87$

c. Perhitungan Ketiga

Perngguna : Mengapa Jepang mengaku sebagai saudara tua Indonesia

Database : Apakah kemerdekaan Indonesia merupakan hadiah dari Jepang ?

Bigram1 pengguna : {M e, e n, n g, g a, a p, p a, a , j, j e, e p, p a, a n, n g, g , m, m e, e n, n g, g a, a k, k u, u , s, s e, e b, b a, a g, g a, a i, i , s, s a, a u, u d, d a, a r, r a, a , t, t u, u a, a , i, i n, n d, d o, o n, n e, e s, s i, i a, a }

Bigram2 database : { a p, p a, a k, k a, a h, h k, k e, e m, m e, e r, r d, d e, e k, k a, a a, a n, n n, n e, e g, g a, a r, r a, a i, i n, n d, d o, o n, n e, e s, s i, i a, a m, m e, e n, n j, j a, a d, d i, i h, h a, a d, d i, i a, a h, h d, d a, a r, r i, i j, j e, e p, p a, a n, n g, g }

Bigram1 \in Bigram2 : { m e, e n, n g, g a, a p, p a, a j, j e, e p, p a, a n, n g, g m, m e, e n, n g, g a, a k, k a, a i, i d, d a, a r, r a, a i, i n, n d, d o, o n, n e, e s, s i, i a, a }

Bigram2 \in Bigram1 : { a p, p a, a k, k m, m e, e g, g a, a r, r a, a i, i n, n d, d o, o n, n e, e s, s i, i a, a m, m e, e n, n i, i a, a d, d a, a r, r i, i j, j e, e p, p a, a n, n g, g }

Diperoleh hasil sebagai berikut :

Count Bigram 1 = 52

Count Bigram 2 = 62

Count Bigram1 \in Bigram2 = 37

Count Bigram2 \in Bigram1 = 33

$$\frac{37 + 33}{52 + 62} = \frac{70}{114} = 0,61$$

Dari hasil perhitungan pencocokan algoritma pertama hingga ketiga memperoleh nilai 0.50, 0.87, dan 0,6. Hasil tertinggi dari perhitungan pencocokan algoritma adalah 0,87. Dengan memperoleh nilai 0.87 maka jawaban yang sesuai dapat dikirimkan kepada pengguna karena memiliki keyword yang sama dan hasil perhitungan yang tinggi.

Berdasarkan hasil evaluasi yang dilakukan oleh pengguna dapat diperoleh 30 pertanyaan dan jawaban dari 5 Responen yang berpartisipasi untuk melakukan evaluasi *Chatbot*. Terdapat 24 pertanyaan dijawab sesuai dengan pertanyaan yang ditanyakan dan 8 pertanyaan yang dijawab tidak sesuai dengan pertanyaan. Untuk mengukur tingkat akurasi sistem maka dilakukan perhitungan sesuai rumus.

$$Accuracy = \frac{\text{Jumlah pertanyaan yang benar}}{\text{Jumlah pertanyaan}} \times 100 \%$$

Diketahui : Jumlah pertanyaan yang benar = 24

Jumlah pertanyaan = 30

Nilai Akurasi :

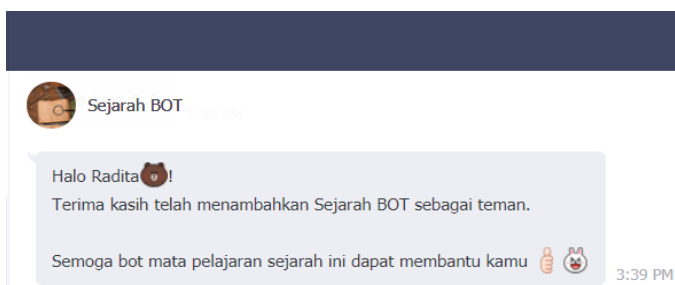
$$Accuracy = \frac{24}{30} \times 100 \%$$

$$Accuracy = 80\%$$

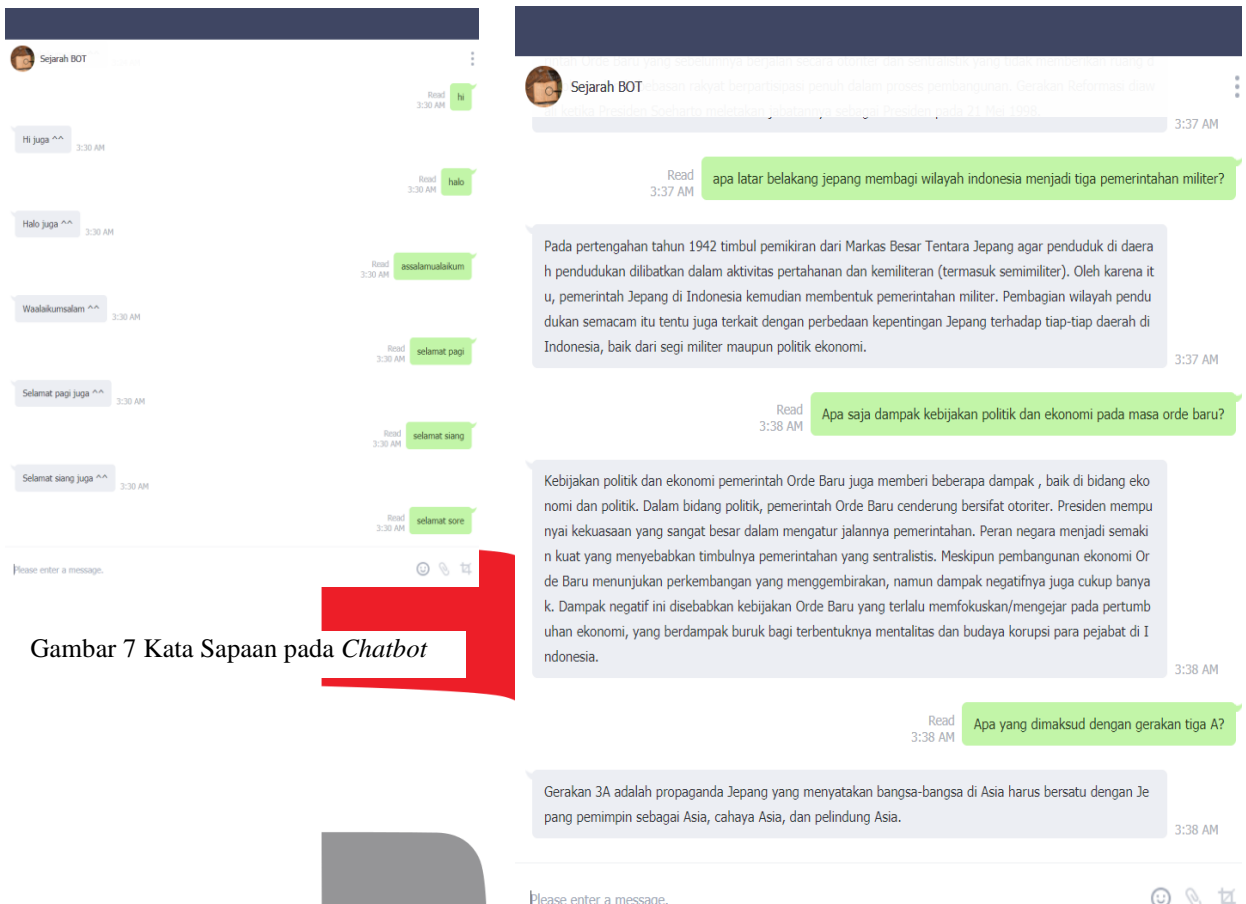
Hasil pengujian menunjukkan bahwa tingkat akurasi yang dimiliki chatbot ketika menjawab pertanyaan yang ditanyakan langsung oleh siswa adalah 80%.

3.4 Tampilan Chatbot

Berikut merupakan tampilan pada *chatbot* :



Gambar 6 Greeting Message



Gambar 7 Kata Sapaan pada Chatbot

Gambar 8 Tanya Jawab Pada Chatbot

4. Kesimpulan dan Saran

4.1 Kesimpulan

Pada penelitian ini dapat dikatakan bahwa:

1. Penulis dapat dikatakan berhasil dalam merancang dan membangun Chatbot untuk Question Answering System dengan menggunakan metode Sentence Similarity Measurement dan Bigram pada platform LINE Messenger. Perancangan Dengan dibangunnya Chatbot pada platform LINE Messenger Pengguna dapat berinteraksi dan melakukan tanya jawab mengenai mata pelajaran Sejarah Indonesia.
2. Dalam kinerjanya Chatbot Question Answering System menggunakan text preprocessing dan metode Sentence Similarity Measurement dan Bigram. Text Preprocessing berguna mengubah bentuk data menjadi terstruktur, kemudian proses dilakukan untuk mengubah pertanyaan menjadi sebuah keyword. Dengan melakukan proses preprocessing dan pencocokan pola algoritma Sentence Similarity Measurement dan Bigram maka sistem dapat merespon pertanyaan Pengguna dengan tepat dan benar.
3. Pada implementasi Chatbot Question Answering System ini menggunakan metode Sentence Similarity Measurement dan Bigram sehingga dihasilkan dari perhitungan pencocokan pola yaitu senilai 0,87, dimana itu merupakan nilai yang mendekati 1 sebagai nilai similarity. Kemudian untuk hasil dari akurasi yang didapatkan ketika Pengguna menggunakan Chatbot yaitu 80%. Dimana hasil tersebut membuktikan bahwa dengan menggunakan metode *Sentence Similarity Measurement* dan Bigram Chatbot dapat digunakan dengan baik dan merespon pertanyaan dengan jawaban yang sesuai.

4.2 Saran

Adapun saran yang dapat diberikan untuk pengembangan penelitian ini adalah

1. Terbatasnya *keyword* pada *database*, sehingga sistem tidak dapat merespon apabila pertanyaan tidak cocok dengan *keyword* yang terdapat pada *database*. Untuk penelitian selanjutnya diharapkan untuk memperbanyak data *keyword* dan jawaban pada *database*.
2. Untuk pengembangan selanjutnya diharapkan sistem tidak hanya dapat menjawab pertanyaan dalam bentuk teks saja, namun dapat dalam bentuk gambar, video, atau pesan suara.

Referensi

- [1] Fauzi, Arfan Rifqi. 2017. "IMPLEMENTASI MULTIPLE INTELLIGENCES RESEARCH (MIR) DALAM PENGELOMPOKAN KELAS DAN PEMBELAJARANNYA(STUDI KASUS DI MIM PK KARTASURA TAHUN AJARAN 2015/2016)." *eprints.ums.ac.id* (1): 1–11. <http://eprints.ums.ac.id/41842/>.
- [2] Agustinova, Danu Eko. 2018. "PENERAPAN KURIKULUM 2013 PADA MATA PELAJARAN SEJARAH PADA SEKOLAH MENENGAH ATAS." *ISTORIA: Jurnal Pendidikan dan Ilmu Sejarah*.
- [3] Ulhaq, Zia. 2017. "Pembelajaran Sejarah Berbasis Kurikulum 2013 Di SMA Kotamadya Jakarta Timur." *Jurnal Pendidikan Sejarah* 6(2): 49–60.
- [4] Hariyono, H. (2017). Peningkatan Hasil Belajar Sejarah melalui Pendekatan Kooperatif Teknik Jigsaw Siswa Kelas XII. *Jurnal Pendidikan : Riset Dan Konseptual*. https://doi.org/10.28926/riset_konseptual.v1i1.2
- [5] Anwarulloh, T. P., Agustia, R. D., & Bandung, J. D. (2017). VIRTUAL PEMBELAJARAN FISIKA DI RUMAH MENGGUNAKAN Teknik Informatika – Universitas Komputer Indonesia. *Elibrary.Unikom.Ac.Id*.
- [6] Dewi, Ayu, and Bayu Setiaji. 2014. "Pemanfaatan Sentence-Similarity Measurement Untuk Proses Pencarian Pola Pada Chatbot Berbasis Pattern-Matching." *Seminar Nasional Teknologi Informasi dan Multimedia 2014*: 39–44. <https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/263>.
- [7] Lenni, Marga, and R. Kristoforus Jawa Bendi. 2019. "Question Answering System Informasi Pariwisata Kota Palembang." *Jurnal Ilmiah Matrik* 21(2): 128–38.
- [8] Handojo, Andreas, Erwin Lie, and Rolly Intan. 2011. "APLIKASI QUESTION ANSWERING SYSTEM DENGAN METODE RULE-BASED QUESTION ANSWERING SYSTEM PADA ALKITAB." *Jurnal Informatika*.
- [9] Ananda Dwi, Firdha Imamah, Yusuf Mei Andre, and Ardiansyah. 2018. "Aplikasi Chatbot (Milki Bot) Yang Terintegrasi Dengan Web CMS Untuk Customer Service Pada UKM MINSU." *Jurnal Cendikia*.
- [10] Sondakh, Merry G.A, Johnny J Senduk, and Sintje A Rondonuwu. 2017. "Pengaruh Berita Line Today Terhadap Perilaku Mahasiswa Program Studi Ilmu Komunikasi Fakultas Ilmu Sosial Dan Ilmu Politik Universitas Sam Ratulangi Manado." *e-Journal "Acta Diurna"* 6(1): 1–16.
- [11] Rahman, Amelia, and Afrizal Doewes. 2017. "Online News Classification Using Multinomial Naive Bayes." *ITSMART: Jurnal Ilmiah Teknologi dan Informasi* 6(1): 32–38.
- [12] Pujadayanti, Irma, Moch Ali Fauzi, and Yuita Arum Sari. 2018. "Prediksi Rating Otomatis Pada Ulasan Produk Kecantikan Dengan Metode Naïve Bayes Dan N-Gram." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)* 2(11): 4421–27.