

CHATBOT UNTUK KONSULTASI AKADEMIK MENGUNAKAN NATURAL LANGUAGE PROCESSING (NLP) DI IBI-K 1957

Syamsu Hidayat¹, Astried Silvanie^{2*}, Rino Subekti³

Departemen Teknik Informatika¹²³
Fakultas Ilmu Komputer¹²³
Institut Bisnis dan Informatika Kosgoro 1957¹²³

Email: Syamsuhi3009@gmail.com¹, astried.silvanie.bk@gmail.com²,
rino.subekti@gmail.com³

Received: February 17, 2022. **Revised:** March 12, 2022. **Accepted:** March 26, 2022. **Issue Period:** Vol.6 No.2 (2022), Pp.396-410

Abstrak: Kemajuan Teknologi komunikasi seiring sejalan dengan teknologi informasi, *Chatbot* merupakan inovasi perpaduan antara teknologi komunikasi dan teknologi informasi, merupakan aplikasi yang dapat berkomunikasi dengan manusia layaknya asisten virtual yang dapat merespon dan menjawab setiap pertanyaan yang diajukan. Beberapa pelaku utama industri komunikasi telah menerapkan teknologi *chatbot* ini salah satunya *Apple*, Inc dengan produk SIRI. IBI-K 1957 sebagai salah satu Perguruan Tinggi Swasta di Jakarta akan terbantu dengan aplikasi *chatbot* ini khususnya dalam hal konsultasi akademik yang bertujuan untuk meringankan beban dosen pembimbing dengan menyediakan bimbingan virtual tanpa ada batasan ruang dan waktu. Mengacu pada penelitian-penelitian sebelumnya, pada penelitian ini kami memanfaatkan teknologi dasar *chatbot* yakni *Machine Learning* dengan *Natural Language Processing*(NLP) yang dapat memproses setiap hasil teks kemudian menganalisa dan melakukan pelatihan mesin menggunakan algoritma Jaringan Syaraf Tiruan (JST) dengan model *Supervised Machine Learning*, kemudian membuat NLP Engine dengan kode *Python* yang terinstal di *Server*. Disisi pengguna dibuat aplikasi berbasis *Android* dengan bahasa *Kotlin* . Uji coba aplikasi dengan menggunakan kotak hitam, hasil baik yang didapat dari beberapa uji coba yang dilakukan bahwa dengan pemanfaatan teknologi *Chatbot* dapat berfungsi sebagai media konsultasi akademik virtual dengan menggunakan NLP dan pemodelan *Machine Learning* dapat memberikan kontribusi dalam bidang kecerdasan buatan khususnya NLP.

Kata Kunci : *Data Mining, Text Mining, Machine Learning, Chatbot, Python, Natural Language Processing, Penelitian Kuantitatif.*

Abstract: *Advances in communication technology are in line with information technology, Chatbot is an innovative combination of communication technology and information technology, an application that can communicate with humans like a virtual assistant who can respond and answer all questions asked. Several major players in the communications industry have implemented this chatbot technology, including Apple, Inc. with SIRI products. IBI-K 1957 as one of the private universities in Jakarta will be helped by this chatbot application, especially in terms of academic consultation which aims to ease the burden on supervisors by providing virtual guidance without any limitations of space and time. Referring to previous studies, in this study we use the core technology of chatbots, namely machine learning with natural language processing (NLP) which can process each text result, then*



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

analyze and perform machine training using the Artificial Neural Network (ANN) Algorithm with Machine Learning Model. the Supervised. , then build NLP Engine with Python code installed on the server. On the user side, an Android-based application is created with the Kotlin language. Application trials using black boxes, good results are obtained from several trials conducted where using chatbot technology can function as a means of virtual academic counseling using NLP and machine learning modeling can contribute to the field of artificial intelligence, especially NLP

Keywords : *Data Mining, Text Mining, Machine Learning, Chatbot, Python, Natural Language Processing, quantitative research.*

I. Pendahuluan

Perkembangan media komunikasi seiring sejalan dengan perkembangan Teknologi Informasi dan ini terjadi setiap waktu. Hal ini sangat berdampak positif bagi masyarakat yang menggunakan manfaatnya. Selain itu pasti ada hal negatif yang dapat terjadi dari manfaat kemajuan komunikasi, bilamana digunakan secara salah cenderung melanggar hukum. Pada era kemajuan teknologi seperti sekarang ini , media komunikasi telah mengalami banyak perubahan dan telah bertransformasi kedalam berbagai bentuk yang baru. Media komunikasi hadir dengan bentuk yang beraneka ragam seperti pesan singkat, *text chat* maupun komunikasi dengan fitur panggilan *audio video*.

Chatbot (atau *Chatter Bot*) merupakan sebuah perangkat lunak yang dapat berkomunikasi kepada manusia seperti layaknya asisten virtual yang dapat berfungsi untuk dapat menjawab beberapa pertanyaan yang diajukan oleh pengguna dan dapat merespon setiap pertanyaan yang diberikan [4]. Saat ini penggunaan *Chatbot* sangat populer dan sudah diaplikasikan seperti SIRI oleh *Apple, Inc*, dan Alexa oleh *Amazon, Inc*. *Chatbot* ini telah diimplementasikan diberbagai bidang, salah satunya adalah bidang Pendidikan guna mendukung siswa belajar maupun konsultasi akademik. Teknologi *Chatbot* ini dapat dianggap sebagai sebuah inovasi penting untuk *e-learning*, selain itu dapat menjadi sebuah solusi paling inovatif untuk mengisi kekosongan antara teknologi dan Pendidikan [4].

Perguruan Tinggi berupaya dalam meningkatkan model dari Pendidikan dengan menyesuaikan dengan kebutuhan siswa, salah satu caranya adalah dengan pengembangan model pendidikan dengan pemanfaatan teknologi informasi dan komunikasi (TIK). Teknologi informasi dan komunikasi telah mengubah paradigma masyarakat, bahkan dapat mengubah cara kita berkomunikasi, dan ini merupakan kontribusi pada pembangunan yang berkelanjutan dalam sektor Pendidikan.

Salah satu dari kelebihan yang dimiliki oleh *Chatbot* adalah [8] :

- Ruang lingkup yang fokus. *Chatbot* dapat menjawab pertanyaan-pertanyaan di area spesifik dan jawabannya sudah diketahui.
- Chatbot* dapat memahami pertanyaan yang membutuhkan beberapa langkah dengan memahami sintaks bahasa menggunakan *Natural Language Processing (NLP)*.
- Chatbot* cocok untuk menjawab pertanyaan yang umum ditanyakan dan sering dengan respon yang hampir mirip.

Menurut Thomson dkk Kesuksesan akademik dipengaruhi banyak faktor, diantaranya manajemen waktu (*goal setting*), nasehat atau bimbingan akademik dilakukan oleh dosen pembimbing yang membimbing konsultasi puluhan siswa untuk masalah akademik [13].

Adapun kendala yang kerap kali dihadapi saat proses konsultasi adalah karena ketidakcocokan waktu antara kedua belah pihak yakni mahasiswa dan dosen pembimbingnya [5]. Sebagai salah satu Perguruan Tinggi Swasta yang berlokasi di Jakarta Selatan, Institut Bisnis dan Informatika Kosgoro 1957 (IBI-K 57) memiliki beberapa fakultas diantaranya fakultas ekonomi, fakultas ilmu social dan politik serta fakultas ilmu komputer, dengan jumlah mahasiswa aktif sebanyak 2.492 mahasiswa. Dengan potensi berupa jumlah mahasiswa aktif yang sebanyak ini peneliti mempunyai ide untuk membuat aplikasi komunikasi antara mahasiswa fakultas ilmu komputer dengan pihak kampus dalam melakukan konsultasi akademik dengan tujuan membantu meringankan beban dosen pembimbing serta menyediakan juga bimbingan tanpa ada Batasan ruang dan waktu yang dilakukan secara virtual.



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

II. Metode dan Materi

Dalam penelitian ini masalah yang akan dijawab adalah :

- a) Bagaimana mengembangkan aplikasi *chatbot* untuk konsultasi akademik mahasiswa yang dapat diakses kapan saja dan dimana saja?.
- b) Bagaimana pengembangan aplikasi *chatbot* untuk dapat memahami data tidak terstruktur yaitu teks atau bahasa.
- c) Bagaimana pengembangan aplikasi *chatbot* untuk dapat memahami data tidak terstruktur yaitu teks atau bahasa.

Adapun kegunaan dari penelitian ini adalah sebagai berikut :

- a) Aplikasi *chatbot* ini dapat digunakan untuk konsultasi akademik secara *virtual* untuk mahasiswa IBI-K 57.
- b) Aplikasi *chatbot* ini dapat dikembangkan sebagai bagian dari Sistem Informasi Akademik berbasis kecerdasan buatan di IBI-K 57.

Dalam penelitian ini peneliti menggunakan beberapa penelitian sebelumnya dan juga buku pengetahuan sebagai referensi, berikut adalah beberapa penelitian sebelumnya Penelitian yang dilakukan oleh Sharma dengan judul Analisa Mengenai Kecerdasan Dalam Sistem *Virtual Chatbot* [11]. Penelitian yang dilakukan oleh (Colace et al., 2018) dengan penelitian berjudul Pembuatan *Chatbot* Untuk Pembelajaran Daring [4]. Penelitian yang dilakukan oleh ((Cucus, 2019), dengan penelitian berjudul Pembuatan *Model Chatbot* Untuk Konsultasi Akademik di Perguruan Tinggi [5]. Penelitian yang dilakukan oleh (Eka Yuniar & Heri Purnomo, 2019) , dengan penelitian berjudul Pembuatan Asisten *Virtual* Berbentuk *Chatbot* Untuk Pusat Informasi BALITTAS [6]. Penelitian yang dilakukan oleh (Adamopoulou & Moussiades, 2020), dengan penelitian berjudul Rangkuman Teknologi *Chatbot* [1]. *Chatbots* Sebuah program komputer yang dapat berinteraksi dengan pengguna yang menggunakan bahasa alami, chatbots menggunakan bahasa alami dalam berkomunikasi dengan pengguna Vichare [14].

Terdapat tiga komponen pada *chatbot* [8] yaitu:

1. Antar muka yang mirip dengan fasilitas pengirim pesan.
2. Kecerdasan buatan
3. Integrasi dengan Sistem Informasi

Proses penemuan pengetahuan sebagai urutan dari langkah-langkah berikut (*Data Mining : Concepts and Techniques*, 2006) [5]:

1. *Data Cleaning*
2. *Data Integration*
3. *Data Selection*
4. *Data Transformation*
5. *Data Mining*
6. *Pattern Evaluation*
7. *Knowledge Presentation*



Teknologi dasar dari chatbot adalah *Machine Learning*, *Natural Language Program*, dan *Artificial Intelegence* [12], NLP bertindak sebagai pilar dasar untuk pengenalan bahasa yang telah digunakan oleh *Siri Apple* dan *Google*. Hal ini memungkinkan teknologi dapat mengenali bahasa alami manusia. *Natural Language Processing* (NLP) adalah sub-bidang dari ilmu komputer dengan Teknik komputasi untuk mempelajari, memahami dan memproduksi bahasa manusia [2]. Menurut S. Divya, Dkk dalam (Soufyane Ayanouz, Boudhir Anouar Abdelhakim, 2020). *Natural Language Understand* bertanggung jawab dalam menangani dan mengubah data tak berbentuk menjadi bentuk yang tepat sehingga mudah dipahami [12]. Adapun lima langkah penting agar pesan dapat dimengerti oleh *chatbot* [12]:

1. *Lexical Analysis*
2. *Syntactic Analyzer*
3. *Semantic Analysis*
4. *Discoource Integration*
5. *Pragmatic Analysis*

Natural Language Generation melibatkan antara realisasi teks dan perencanaan teks untuk tanggapan yang dimengerti. Adapun tantangan utama yang dihadapi oleh NLP adalah untuk memahami komplikasi bahasa alami manusia. Ide dasar dari jaringan saraf berulang dua arah adalah untuk menyajikan setiap urutan pelatihan maju dan mundur ke dua jaringan berulang yang terpisah, yang keduanya terhubung ke lapisan *output* yang sama. Ini berarti bahwa untuk setiap titik dalam urutan tertentu, BRNN memiliki informasi berurutan yang lengkap tentang semua titik sebelum dan sesudahnya. Juga, karena net bebas untuk menggunakan konteks ini sebanyak atau sesedikit yang diperlukan, tidak perlu menemukan jendela waktu (tergantung tugas) atau ukuran penundaan target [7]. Niat didefinisikan sebagai tindakan yang bertujuan. Klasifikasi niat (berfokus pada tindakan dimasa depan) adalah sebuah klasifikasi teks [10].

Penelitian ini menggunakan metode kuantitatif, merupakan metode dengan melakukan kuantifikasi dan analisis *variable* untuk mendapatkan hasil dengan menggunakan analisis data numerik dengan statistic tertentu untuk dapat menjawab pertanyaan dipenelitian ini [3], adapun pertanyaan yang ingin dijawab adalah :

- 1) Bagaimana kita melatih mesin agar memahami data tidak terstruktur yaitu teks.
- 2) Bagaimana membuat aplikasi *chatbot* untuk konsultasi akademik masalah umum studi kasih kampus IBI-K 57.

Penelitian ini dilakukan dalam waktu 5 bulan sejak bulan April 2021 sampai Agustus 2021. Berlokasi di Institut Bisnis dan Informatika Kosgoro 1957 yang beralamat di Jl. M. Kahfi II No. 56 Lenteng Agung Jakarta Selatan. Penelitian di bidang pembelajaran mesin mayoritas adalah kuantitatif karena melibatkan pemodelan data, pendekatan dan formulasi *statistic* untuk mengolah data [9], adapun langkah-langkah dalam penelitian ini :

1. Pengumpulan data dan pemodelan data
2. Pelatihan mesin dengan algoritma Jaringan Syaraf Tiruan (JST)
3. Pengujian model JST dengan menggunakan metrik akurasi.
4. Pembangunan NLP *Engine* dengan *Python* dan terinstall di *server*
5. Membuat aplikasi *Android* dengan bahasa kotlin dan IDE *Android Studio*
6. Uji coba *black box* aplikasi *Android*.

Data pada penelitian ini bersumber dari data primer aktivitas akademik di IBI-K 57 didalamnya termasuk kegiatan belajar mengajar, ujian dan skripsi, semua data ini didapatkan dari observasi dokumen akademik. Dari kegiatan observasi dokumen, kita dapat mendefinisikan sebuah dokumen teks pendek yang



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

relevan, mendefinisikan sebagai pesan yang berisi kata kunci terkait peristiwa apapun, sehubungan dengan acara kata kunci disediakan secara manual berdasarkan dokumen dari IBI-K 57. Klasifikasi maksud menjawab pertanyaan : Apa yang penanya (mahasiswa/i) coba katakan/tuliskan?.

Dalam klasifikasi maksud, kalimat $d \in X$ dari sebuah dialog diberikan, di mana X adalah ruang ucapan dan K adalah jumlah d .

$$D = \{D_i | 1 \leq i \leq K\}$$

Satu set tetap kelas maksud (*Intent Class*), sebanyak j didefinisikan sebagai berikut:

$$C = \{C_i | 1 \leq i \leq J\}$$

Satu set pelatihan D berlabel ucapan dialog $\{d_i, c_i\}_{i=1}^N$, di mana $(d, c) \in X \times C$. Dalam penelitian ini satu kalimat d akan diberi label tunggal atau dengan kata lain klasifikasi label tunggal, di mana setiap d sesuai dengan satu dan hanya satu elemen C . Contohnya sebagai berikut ini:

$(d, c) = (\text{'Kapankah UTS dilakukan'}, \text{'Jadwal UTS'})$.

Sumber data training berupa pola maksud disimpan dalam file JSON (*JavaScript Object Notation*) dengan nama *intents2id.json*.

Format:

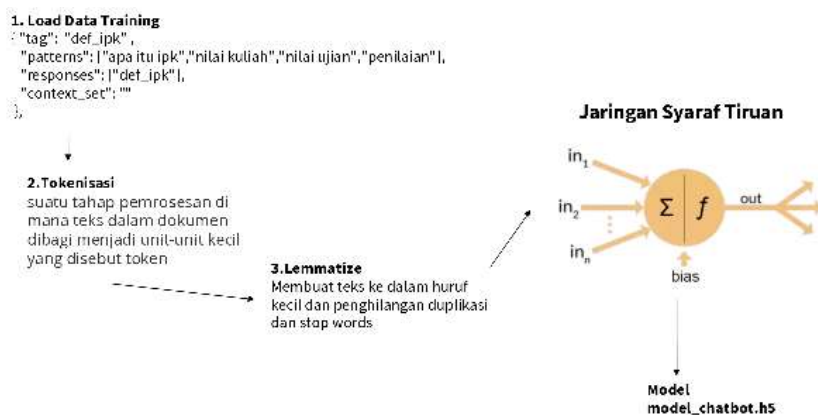
tags : maksud

polatanya: macam-macam pertanyaan berkaitan dengan maksud

respon : maksud

Cara bertanya dibagi berdasarkan maksud arti, cara dan waktu. Misalkan “pengertian registrasi”, “bagaimana cara melakukan registrasi”, “kapan registrasi dilakukan”. Ditambahkan juga kata-kata yang tidak EYD (ejaan yang disempurnakan) seperti “gimana”, “hey”, “hi”, “sih”.

Model JST terdiri dari lapisan masukan, lapisan tersembunyi dan lapisan keluaran. Jumlah neuron masukan berjumlah banyak kelas yang dihasilkan dari tags dalam *intents2id.json*. Jumlah neuron luaran terdiri dari satu saja karena akan memilih satu respon yang tepat sesuai dengan maksud dari teks.



Gambar 1. Proses Pembelajaran mesin dengan *intents*

Fungsi aktivasi yang digunakan adalah *Rectified Linear Unit* (ReLU), fungsi ini mengembalikan nilai 0 jika input negatif dan nilai 1 jika input positif.



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

$$\begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$

Gambar 2. Rumus Fungsi Aktivasi *Rectified Linear Unit* (Relu)

Proses pengembangan bekerja melalui iterasi berarti bahwa pengembangan aplikasi dibagi menjadi bagian-bagian yang lebih kecil. Dalam setiap iterasi fitur didefinisikan, dirancang, dikembangkan dan diuji. Siklus iterasi diulang sampai perangkat lunak yang berfungsi penuh siap dikirim ke produksi. Prosesnya tidak langsung dimulai dengan set lengkap persyaratan dan desain. Sebaliknya, tim mencoba untuk mempersiapkan apa yang dibutuhkan untuk keberhasilan pengiriman iterasi berikutnya [15].

III. PEMBAHASAN DAN HASIL

Pada penelitian ini digunakan *python* dan *jupyter notebook*, berikut adalah langkah yang dilakukan dalam pelatihan.

```
import json
import pickle
import random
import numpy as np
import nltk

from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()

words=[]
classes = []
documents = []
acuhkan_kata_ini = ['?', '!', ',', ':', ';', ':', ':']
data_file = open('intents2id.json').read()
intents = json.loads(data_file)
```

Gambar 3. Pelatihan Bagian 1

Langkah pertama adalah mengimpor semua pustaka yang dibutuhkan seperti tertera pada (gambar 4.1). kemudian membaca data dari file, kemudian memecahkan setiap pola tanya menjadi unit-unit kecil yaitu *token*. Memasangkan setiap pertanyaan tersebut dengan artinya, lalu simpan sebagai dokumen . hasilnya akan seperti berikut ini :

```
(['waktu', 'bayar'], 'waktu_registrasi'),
(['waktu', 'registrasi'], 'waktu_registrasi'),
(['kapan', 'bayar'], 'waktu_registrasi'),
(['kapan', 'registrasi'], 'waktu_registrasi')
```



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).


```
for intent in intents[intents]:
    for pattern in intent['polanya']:
        #tokenize
        output = stemmer.stem(pattern)

        w = output.split()
        words.extend(w)

        #tambah corpus
        documents.append(w, intent['tag'])

        #tambah ke class
        if intent['tag'] not in classes:
            classes.append(intent['tag'])
```

Gambar 4. Pelatihan Kode 2

Dilanjutkan dengan proses Lemmatisasi dengan tujuan agar dapat dianalisis menjadi satu item, diidentifikasi dengan kata lemma atau bentuk kamus kata (gambar 5)

```
# lemmatize, lowe case, remove duplicate
words = [lemmatizer.lemmatize(w.lower()) for w in words if w not in acuhkan_kata_ini]
words = sorted(list(set(words)))

# sorting
classes = sorted(list(set(classes)))

# documents = kombinasi patterns dan intents
print (len(documents), 'documents')

# classes = intents
print (len(classes), 'classes', classes)

# words = kosakata
print (len(words), 'unique lemmatized words', words)

pickle.dump(words,open('words.pkl','wb'))
pickle.dump(classes,open('classes.pkl','wb'))
```

Gambar 5. Pelatihan Kode 3

Kemudian melakukan proses pengubahan teks menjadi vektor dengan Panjang tetap dengan menghitung setiap kata yang muncul . adapaun model yang dihasilkan dari proses ini disebut *Bag of Words* (BOW) misal terdapat 1 dokumen seperti dibawah ini :

dokumen 1 : “bayar kuliah”

dokumen 2 : “syarat bayar kuliah”

Bag of words : [“bayar”, “kuliah”, “syarat”]

Vektorisasi dokumen 1 adalah [1,1,0] dan vektorisasi dokumen 2 adalah [1,1,1]. Kemudian kita masukkan data teks ke pelatihan dimana pola tanya menjadi atribut predikator dan tags menjadi kelas target.



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).


```

from keras.models import Sequential
from keras.layers import Dense, Activation, Dropout
#from keras.optimizers import SGD
from tensorflow.keras.optimizers import Adam

# JST 3 layers. Layer1 128 neurons, layer2 64 neurons and layer3 output neurons
# output adalah banyak intents yang diprediksi
model = Sequential()
model.add(Dense(128, input_shape=(len(train_x[0]),), activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(64, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(len(train_y[0]), activation='softmax'))

# Set Optimizer
opt = Adam(lr=0.001, decay=1e-6)

# Compile model
model.compile(
    loss='categorical_crossentropy',
    optimizer=opt,
    metrics=['accuracy']
)

#fitting & saving model
hist = model.fit(np.array(train_x), np.array(train_y), epochs=200, batch_size=5, verbose=1)
model.save('charbot_model.h5', hist)

```

Gambar 7. Pelatihan Kode 5

Untuk mengukur kinerja jaringan saraf dengan menggunakan akurasi seperti berikut :

$$Accuracy = (TP+TN) / (TP+TN+FP+FN)$$

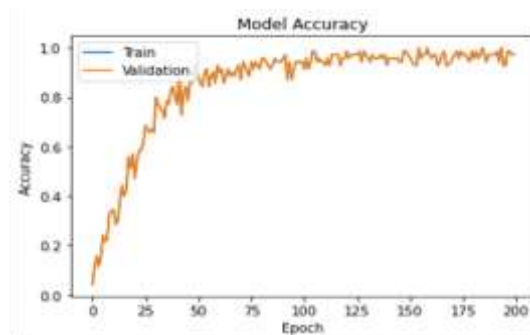
Diketahui:

TP = *True Positive*, adalah jumlah sampel yang diklasifikasikan secara benar dan memang benar.

TN = *True Negative*, adalah jumlah sampel yang diklasifikasikan secara negatif dan memang benar.

FN = *False Neagtive*, adalah jumlah sampel yang diklasifikasikan sebagai negatif tetapi sebenarnya positif.

FP = *False Positive*, adalah jumlah sampel yang diklasifikasikan sebagai positif tetapi sebenarnya negatif



Gambar 8. Metrik Akurasi



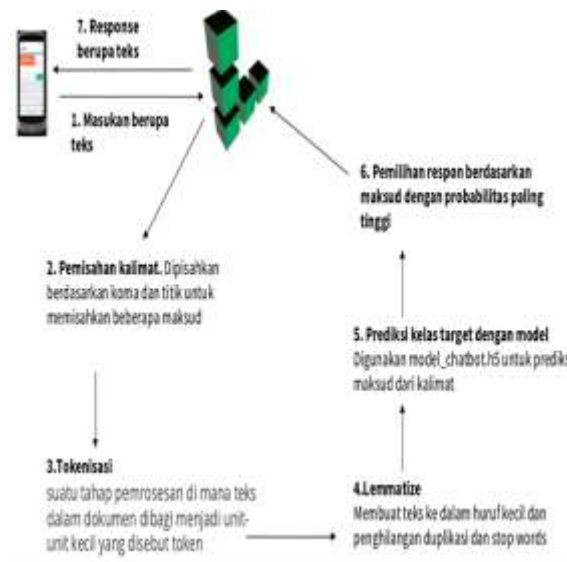
DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Kebutuhan fungsional dari sistem *chatter bot* berbasis NLP yang harus dipenuhi:

- 1) *Login* Aplikasi dengan input nama dan jurusan
- 2) Memasukkan pesan ke *Robot* via *textbox* pesan.
- 3) *Robot* merespon via kotak pesan.
- 4) Koneksi ke *Cloud Server*.
- 5) *Cloud Server* menjalankan kode *python* dengan model *machine learning*.

Setelah mendapatkan model JST lengkap beserta bobotnya ,lalu kita akan mengimplementasikan model ini untuk membangun aplikasi.



Gambar 9. Implementasi dan Integrasi Model JST dengan Aplikasi

Adapun fungsi-fungsi *python* dalam menjalankan JST adalah sebagai berikut.

```
import nltk
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()

# import StemmerFactory class
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
# create stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()

from keras.models import load_model

import json
import random
import json
import pickle
import numpy as np
import re
```



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

```

intents = json.loads(open('intents2id.json').read())
words = pickle.load(open('words.pkl', 'rb'))
classes = pickle.load(open('classes.pkl', 'rb'))

def clean_up_sentence(sentence):
    output = stemmer.stem(sentence)
    sentence_words = output.split()
    sentence_words = [lemmatizer.lemmatize(word.lower()) for word in sentence_words]
    return sentence_words

def bow(sentence, words, show_details=True):
    sentence_words = clean_up_sentence(sentence)
    bag = [0] * len(words)
    for s in sentence_words:
        for i, w in enumerate(words):
            if w == s:
                bag[i] = 1
    return np.array(bag)

def predict_class(sentence, model):
    p = bow(sentence, words, show_details=False)
    res = model.predict(np.array([p]))[0]
    ERROR_THRESHOLD = 0.70
    return_list = []
    results = [(i, j) for i, j in enumerate(res) if j > ERROR_THRESHOLD]

    if len(results) == 0:
        return_list = []
    else:
        results.sort(key=lambda x: x[1], reverse=True)
        for r in results:
            return_list.append({"intent": classes[r[0]], "probability": str(r[1])})
    return return_list

def getResponse(ints, intents_json):
    tag = ints[0]['intent']
    list_of_intents = intents_json['intents']
    for i in list_of_intents:
        if (i['tag'] == tag):
            result = random.choice(i['responses'])
            break
    return result

def chatbot_response(text):
    lower_case = text.lower()
    ints = predict_class(lower_case, model)
    if len(ints) == 0:
        else:
            res = getResponse(ints, intents)
            return res

def iterate_in_text(text):
    res = ""
    res_list = []
    question_list = []
    output_list = []
    array_words = re.split(r'[\s]', text)

    for x in array_words:
        res = chatbot_response(x)
        output_list.append(x, res)
        if(res == "404"):
            file1 = open("files/kalimatbaru.txt", "a") # append mode
            file1.write(text + "\n")
            file1.close()

    return output_list

```

Gambar 10. Flask digunakan untuk menjalankan kode python berupa NLP engine (Gambar 4.8).



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

```

GNU nano 4.8
from flask import Flask, request, jsonify
from chatbot_func import iterate_in_text
import json
from pathlib import Path

app = Flask(__name__)

@app.route('/api/chat/', methods=['GET', 'POST'])
def classify():
    text = request.args.get("text")
    question_intent = iterate_in_text(text)
    replies = []
    intent = ""
    reply = ""

    for pair_row in question_intent:
        datatxt = "Maaf bisa diulangi"
        if(pair_row[1]=="404"):
            datatxt = "Maaf, saya kurang mengerti pertanyaan '"+pair_row[0]+'"'
        else:
            datatxt = Path('files/' + pair_row[1] + '.txt').read_text()

        replies.append([pair_row[1],pair_row[0],datatxt])

        intent = intent + pair_row[1] + ";"
        reply = reply + datatxt + ";"

    x = [{"intent": intent, "reply":reply}]

    return jsonify(x)

if __name__ == '__main__':
    app.run(host="0.0.0.0",debug=True)

```

Gambar 11. Eksekusi Flask untuk menjalankan fungsi-fungsi python

Berikut ini adalah uji coba dalam bentuk *black-box* ,dengan setiap kasus diujikan lebih dari satu kalimat secara berkali-kali.

Tabel 1 Uji Coba Kotak Hitam / Black Box

No	Kasus	Hasil yang diharapkan	Hasil yang keluar
1	Kalimat yang ad akelas target di intens2id.json	Pengembangan kelas target dan respon	Tercapai
2	Kalimat yang tidak ad akelas target di intenst2id.json	Pengembangan kelas target 404, respon “Pertanyaan kurang dimengerti” dan kalimat tersebut disimpan di server untuk latihan selanjutnya	Tercapai
3	Kalimat yang mempunyai lebih dari satu kelas target di intents2id.json	Pengembangan beberapa kelas target dan responnya	Tercapai

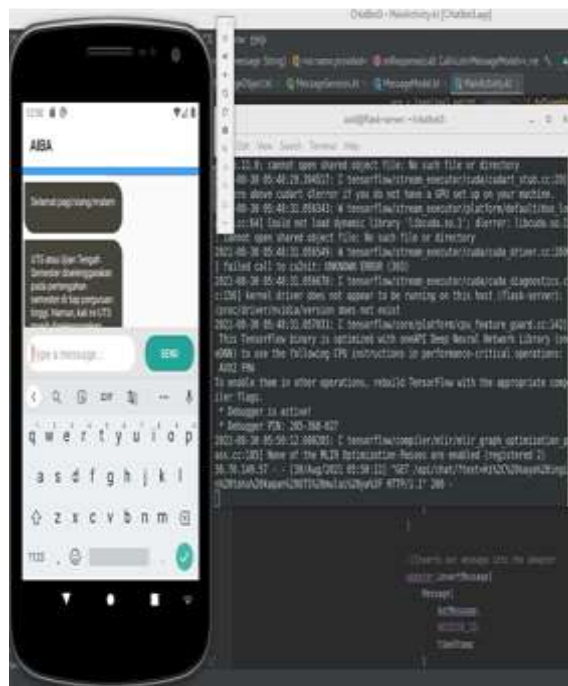


DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



Gambar 12. Kalimat Pembuka di Chatbot



Gambar 13.. Jawaban dari chatbot, disini terlihat terdeteksi dua maksud dalam kalimat tanya sehingga mengembalikan response dua kali



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Adapun hambatan yang ditemui dalam penelitian ini adalah sebagai berikut :

- (1) Spesifikasi komputer dan *cloud* dipilih paling minimum sesuai budget penelitian, hal ini akan membuat latihan machine learning memakan waktu lama (1 jam atau lebih) jika data cukup banyak.
- (2) Pemilihan algoritma permodelan ML dicoba 3 model, dipilih model yang paling sederhana (karena alasan kompleksitas implementasi) dan lebih kecil (karena alasan keterbatasan kapasitas *server*)
- (3) Pemilihan antara (a) integrasi *python + android* atau (b) *cloud server + python + android*. Poin a memerlukan *TensorFlow Lite* untuk menjalankan kode ML *python* di dalam kode *kontlin android*. Poin b memerlukan *cloud server* sebagai pihak yang menjalankan kode *python*.
- (4) Pemilihan beberapa kandidat VPS (1. *Google Cloud*, 2. *AWSLight*, 3. *Linode*). Dipilih *Linode* dengan alasan dapat voucher \$100 dan juga lebih mudah untuk koneksi via *remote*.

IV. KESIMPULAN DAN SARAN

- 1) Pembuatan *Asisten Virtual/Chatter Robot* sebagai media konsultasi akademik virtual untuk menjawab pertanyaan-pertanyaan umum mengenai akademik.
- 2) Pembuatan Aplikasi mobile based terkoneksi dengan VPS (*Virtual Private Server*) sebagai media yang dapat diakses tanpa batasan ruang dan waktu.
- 3) *Chatter Robot* dibangun dengan model Machine Learning dan Natural Language Processing (NLP) yang dapat memahami maksud kalimat. Hasil ini memberikan kontribusi dalam bidang kecerdasan buatan khususnya NLP.

Saran untuk penelitian yang akan dilakukan adalah sebagai berikut :

- 1) NLP untuk memahami maksud yang digunakan dalam penelitian ini dapat di pasangkan dengan Sistem Pakar bisa menggunakan *Use Case Based Reasoning* atau *Soft Computing*.
- 2) Pelatihan Model dapat ditambahkan kamus besar bahasa indonesia untuk menambah kosa kata. Memang akan membutuhkan storage yang lebih besar.

REFERENSI

- [1] Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. In *IFIP Advances in Information and Communication Technology: Vol. 584 IFIP*. Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31
- [2] Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and analyzing Chatbot responses using natural language processing. *International Journal of Advanced Computer Science and Applications*, 10(9), 60–68. <https://doi.org/10.14569/ijacsa.2019.0100910>
- [3] Apuke, O. D. (2017). Quantitative Research Methods : A Synopsis Approach. *Kuwait Chapter of Arabian Journal of Business and Management Review*, 6(11), 40–47. <https://doi.org/10.12816/0040336>
- [4] Colace, F., De Santo, M., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, 7(5), 528–533. <https://doi.org/10.18178/ijmerr.7.5.528-533>
- [5] Cucus, A. (2019). *Fakultas ilmu komputer.Data Mining : Concepts and Techniques*. (2006). *Data Mining : Concepts and Techniques* (Second). University of Illinois at Urbana Campaign.
- [6] Eka Yuniar, & Heri Purnomo. (2019). Implementasi Chatbot “Alitta” Asisten Virtual Dari Balittas Sebagai Pusat Informasi Di Balittas. *Antivirus : Jurnal Ilmiah Teknik Informatika*, 13(1), 24–35. <https://doi.org/10.35457/antivirus.v13i1.714>
- [7] Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., & Wierstra, D. (2015). DRAW: A recurrent neural network for image generation. *32nd International Conference on Machine Learning, ICML 2015*, 2, 1462–1471.
- [8] Interactive, A. (2016). *Chatbots in Customer Service*.



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

- [9] Kamiri, J., & Mariga, G. (2021). Research Methods in Machine Learning: A Content Analysis. *International Journal of Computer and Information Technology*(2279-0764), 10(2), 78–91. <https://doi.org/10.24203/ijcit.v10i2.79>
- [10] Purohit, H., Dong, G., Shalin, V., Thirunarayan, K., & Sheth, A. (2015). Intent classification of short-text on social media. *Proceedings - 2015 IEEE International Conference on Smart City, SmartCity 2015, Held Jointly with 8th IEEE International Conference on Social Computing and Networking, SocialCom 2015, 5th IEEE International Conference on Sustainable Computing and Communications, SustainCom 2015, 2015 International Conference on Big Data Intelligence and Computing, DataCom 2015, 5th International Symposium on Cloud and Service Computing, SC2 2015*, 222–228. <https://doi.org/10.1109/SmartCity.2015.75>
- [11] Sharma, V., Goyal, M., & Malik, D. (2017). An Intelligent Behaviour Shown by Chatbot System. *International Journal of New Technology and Research*, 4, 52–54.
- [12] Soufyane Ayanouz, Boudhir Anouar Abdelhakim, M. B. (2020). A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance. *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*.
- [13] Suprihatin, T. (2017). *Prestasi akademik ditinjau dari model bimbingan akademik*. 321–330.
- [14] Vichare, A., Gyani, A., Shrikhande, Y., & Rathod, N. (2015). A chatbot system demonstrating Intelligent Behaviour using NLP. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 4(10), 4–6.
- [15] Victor, R. (2003). *L2a- Iterative Incremental Development - A Brief History [Larman,Basili IEEE.Computer 2003] (10pp)*. June, 47–56.



DOI: 10.52362/jisamar.v6i2.780

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).