

Implementasi *Natural Language Processing* (NLP) Dalam Pengembangan Aplikasi Chatbot Pada SMK YPE Nusantara Slawi

Implementation of Natural Language Processing (NLP) in the Development of a Chatbot Application at SMK YPE Nusantara Slawi

Dyah Apriliani¹, Sharfina Febbi Handayani², Irfan Triadi Saputra³

^{1,2,3} Teknik Informatika, Politeknik Harapan Bersama

E-mail: ¹dyahapril@poltektegal.ac.id, ²sharfina.handayani@poltektegal.ac.id,

³irfans1710@gmail.com

Abstrak

Perkembangan teknologi informasi memberikan berbagai perubahan dalam berbagai bidang kehidupan, salah satunya adalah dalam dunia pendidikan. Perkembangan teknologi tersebut ditandai dengan munculnya teknologi berbasis *Artificial Intelligence* (AI) yang sudah mulai digunakan. Salah satunya adalah aplikasi *chatbot*. Aplikasi *chatbot* merupakan suatu aplikasi yang dibangun dengan mengadopsi pengetahuan manusia yang dapat memberikan respon secara cepat dan tepat. SMK YPE Nusantara Slawi sudah memiliki aplikasi *chatbot* akan tetapi aplikasi tersebut hanya digunakan untuk pembelajaran materi di salah satu prodi saja. Pihak sekolah mengalami kesulitan dalam promosi dan juga penyampaian informasi kepada masyarakat. Banyak masyarakat yang ketika membutuhkan informasi harus datang ke sekolah. Bahkan seringkali pertanyaan yang ditanyakan adalah sama antar satu dengan yang lainnya, sehingga terjadi ketidak efisienan. Oleh karenanya dibutuhkan suatu aplikasi interaktif yang dapat memberikan informasi tentang sekolah tersebut. Aplikasi *chatbot* yang dibuat berisi informasi secara menyeluruh seputar sekolah yang terdiri dari jurusan, guru maupun data lainnya yang berkaitan dengan sekolah. Penelitian dilakukan dalam beberapa tahapan mulai dari pengambilan data, analisis data, pembuatan aplikasi pengujian sampai ke implementasi sistem. Aplikasi dikembangkan dengan menggunakan *Natural Language Processing* (NLP) yang akan dikembangkan berbasis website.

Kata kunci: Aplikasi, Chatbot, Informasi, NLP, Sekolah, SMK

Abstract

The development of information technology brings various changes in many fields of life, one of which is education. This technological development is marked by the emergence of Artificial Intelligence (AI)-based technology that has begun to be used. One example is a chatbot application. A chatbot application is an application built by adopting human knowledge that can provide quick and accurate responses. SMK YPE Nusantara Slawi already has a chatbot application, but this application is only used for learning materials in one department. The school has difficulty in promoting and conveying information to the public. Many people need to come to school when they need information. Often, the questions asked are the same from one person to another, resulting in inefficiency. Therefore, an interactive application is needed to provide information about the school. The created chatbot application contains comprehensive information about the school consisting of departments, teachers, and other data related to the school. The research was conducted in several stages, from data collection, analysis, and application creation testing to system implementation. The application was developed using Natural Language Processing (NLP) which will be web-based.

Keywords: Application, Chatbot, Information, NLP, School, SMK

1. PENDAHULUAN

Seiring perkembangan zaman, teknologi informasi juga mengalami berbagai macam perkembangan. Perkembangan teknologi informasi ini memiliki peranan penting dalam berbagai bidang kehidupan, namun yang paling penting adalah pengguna dapat menerima keberadaan teknologi informasi tersebut [1]. Perkembangan teknologi yang berkembang saat ini ditandai dengan semakin banyaknya pengguna komputer maupun smartphone di semua kalangan dan semua bidang baik dalam bidang pendidikan, kesehatan maupun bidang lainnya [2]. Perkembangan teknologi yang sangat pesat ini juga membawa perubahan dalam kehidupan manusia, salah satunya adalah dengan perkembangan teknologi berbasis *Artificial Intelligence* (AI) [3]. Dengan memanfaatkan teknologi AI tersebut, komputer dapat melakukan tugas tertentu yang hampir sama dengan manusia, salah satunya adalah dengan menggunakan *chatbot* [4]. *Chatbot* merupakan suatu aplikasi yang dibangun dengan mengadopsi pengetahuan manusia [5]. Aplikasi *chatbot* ini digunakan untuk membantu pengguna mendapatkan informasi yang dibutuhkan dengan melakukan percakapan dengan menggunakan bahasa alami [6].

Aplikasi *chatbot* ini sudah banyak digunakan berbagai bidang, salah satunya adalah bidang pendidikan. Teknologi *chatbot* dalam dunia pendidikan digunakan untuk membantu pelayanan dalam menjawab pertanyaan-pertanyaan dari siswa terkait dengan informasi sekolah yang biasanya ditanyakan secara berulang. dengan adanya Aplikasi *Chatbot* ini diharapkan mampu memberikan peningkatan kualitas layanan terhadap siswa. Salah satu metode yang dapat digunakan dalam pembuatan *chatbot* ini adalah dengan menggunakan metode *Natural Language Processing* (NLP). NLP merupakan cabang dari AI yang membantu komputer membaca dan memahami bahasa alami manusia, dengan tujuan untuk meningkatkan komunikasi antara mesin dengan manusia [7]. Metode NLP digunakan dalam aplikasi *chatbot* dikarenakan metode ini lebih mudah, fleksibel dan mampu memberikan jawaban atau data yang lebih akurat [8]. Aplikasi dalam penelitian ini dibuat dengan menggunakan *transfer learning* pada pengolahan data pada *chatbot* berbasis teknik pengolahan bahasa alami (NLP) dan jaringan saraf tiruan (*neural networks*). *Transfer learning* yang digunakan dengan menggunakan *DenseNet*. *DenseNet* adalah salah satu jenis *pre-trained* model atau model yang sudah pernah dilatih sebelumnya dengan dataset yang lebih besar, seperti dataset CIFAR, *Street View House Numbers* (SVHN), dan *ImageNet* [9]. Aplikasi *chatbot* mampu memberikan kepuasan kepada pengguna ketika mendapatkan informasi dengan cepat. Aplikasi chatbot dapat memberikan respon berbagai pertanyaan yang berbeda dengan cepat.

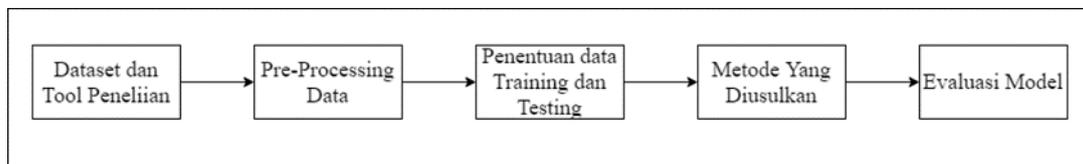
Pada penelitian [10], aplikasi *chatbot* juga dibuat untuk memberikan layanan informasi sekolah di SMK Dwi Sejahtera Pekanbaru. Aplikasi *chatbot* yang dibuat digunakan untuk memberikan layanan informasi dengan cara memberikan respon pertanyaan yang diberikan oleh *user* dan memberikan jawaban *default* yang sudah ditentukan. Adapun pertanyaan yang diberikan diantaranya adalah alamat, biaya, pendaftaran, syarat pendaftaran dan informasi lain tentang sekolah yang dibutuhkan. Tujuan dibuatnya sistem ini adalah untuk mempermudah calon siswa yang ingin mendaftar untuk mengetahui informasi seputar sekolah dengan menggunakan aplikasi tanpa harus datang ke sekolah. Selain itu penelitian sebelumnya tentang membandingkan bagaimana menciptakan rasa emosional, relasional, dan psikologis antara percakapan dengan manusia dan *chatbot*. Studi ini mempelajari dampaknya apabila manusia diubah menjadi teknologi, serta konsekuensi pengungkapan ke chatbot. Hasilnya diperoleh bahwa percakapan antara manusia maupun chatbot sama-sama efektif dalam menciptakan manfaat emosional, relasional, dan psikologis [11].

SMK YPE Nusantara Slawi merupakan salah satu sekolah di Kabupaten Tegal yang akan mengimplementasikan pengembangan aplikasi chatbot dari penelitian ini. Sampai dengan kajian ini dibuat, sekolah telah memiliki aplikasi chatbot yang digunakan dalam pembelajaran akuntansi, akan tetapi belum diimplementasikan untuk mengelola informasi sekolah secara menyeluruh, baik dari jurusan maupun informasi lainnya. Berdasarkan observasi yang telah dilakukan di SMK YPE Nusantara Slawi, proses promosi yang dilakukan masih kurang optimal, yaitu masih mengguankan cara manual dan belum memanfaatkan teknologi informasi. Masyarakat maupun

calon siswa yang akan mengetahui tentang sekolah harus datang ke sekolah ketika ingin mendapatkan informasi. Hal ini, menyebabkan ketidakefisien dalam proses penyebaran informasi seputar sekolah sebagai media promosi. Bahkan tidak jarang pertanyaan yang diajukan ke pihak sekolah pun cenderung sama dan berulang. Berdasarkan masalah tersebut, melalui penelitian ini penulis mengembangkan aplikasi chatbot dengan mengimplementasikan *Natural Language Processing* (NPL) dalam format tanya jawab berupa teks atau audio [12]. *Chatbot* diprogram untuk bekerja secara independen dan dapat menjawab pertanyaan yang dirumuskan berdasarkan kombinasi dari skrip yang telah ditentukan dalam Bahasa alami serta merespons layaknya manusia sungguhan [13]. Penggunaan *Chatbot* berbasis *Artificial Intelligence* (AI) memberikan dampak positif bagi penggunanya, *chatbot* memberikan kemudahan dalam melakukan otomatis menjawab pertanyaan sesuai kebutuhan pengguna [14]. Implementasi *Natural Language Processing* (NLP) memungkinkan chatbot untuk memahami pesan dan merespons dengan tepat. Kemampuan *Natural Language Processing* (NLP) dapat mengidentifikasi kesalahan ejaan dan tata bahasa dan memungkinkan chatbot untuk menafsirkan pesan meskipun terdapat kesalahan [15]. Tujuan dari pembuatan aplikasi *chatbot* ini adalah mempermudah pengguna dalam mencari informasi seputar SMK YPE Nusantara Slawi sehingga mampu membantu dalam kegiatan promosi sekolah dan pusat media informasi seputar sekolah.

2. METODE PENELITIAN

Bahan penelitian yang digunakan dalam penelitian ini menggunakan data yang didapatkan dari SMK YPE Nusantara Slawi yang terdiri dari data umum sekolah, data jurusan, data guru dan data lainnya yang nantinya akan digunakan dalam aplikasi *chatbot* ini. Tahapan yang akan dilakukan di dalam penelitian ini dimulai dari pengumpulan dataset dan *tool* penelitian, *pre-processing* data, penentuan data *training* dan data *testing*, metode yang diusulkan dan evaluasi model diilustrasikan pada Gambar 1 berikut.



Gambar 1 Alur Penelitian

2.1 Dataset Dan Tools Penelitian

Dataset yang dijadikan acuan berasal dari SMK YPE Nusantara Slawi. Dataset ini mencakup beragam informasi, termasuk profil sekolah yang mencakup visi dan misi, profil lengkap jurusan yang ditawarkan, daftar dedikasi guru-guru berpengalaman, serta rangkuman mata pelajaran yang diajar dengan teliti. Semua data ini akan dianalisis dan dimanfaatkan menggunakan bahasa pemrograman Python. Kumpulan data yang telah diperoleh akan dimasukkan secara manual dan disimpan dalam format JSON (*JavaScript Object Notation*). Penelitian ini menghasilkan 116 Patterns dan Responses yang diilustrasikan pada Gambar 2.

```

{"tag": "profil",
 "patterns": ["Profil sekolah?", "profil", "profil smk ype nusantara slawi", "profil smk ype"],
 "responses": ["SMK YPE Nusantara Slawi merupakan salah satu Lembaga Pendidikan Menengah Kejuruan di Kabupaten Tegal
"],
 {"tag": "tahun",
 "patterns": [ "Kapan tahun berdiri?", "Pada tahun berapa?", "tahun berdiri", "berdiri", "kapan berdirinya smk ype nusantara slawi"],
 "responses": ["Kamu tahu gak sih bahwa SMK YPE Nusantara Slawi berdiri pada 1 Januari 1972 "]
},
 {"tag": "akreditasi",
 "patterns": [ "Apa akreditasi nya?", "Akreditasi nya adalah?", "Sudah terakreditasi?", "akreditasi smk ype nusantara slawi"],
 "responses": ["SMK YPE Nusantara Slawimemiliki Akreditasi B"]
},
 {"tag": "alamat",
 "patterns": [ "di mana alamat nya?", "alamat", "alamat smk ype nusantara slawi", "alamat smk ype", "alamat smk"],
 "responses": ["SMK YPE Nusantara Slawi beralamat di Jln. Letjen M.T. Haryono Kalisapu Slawi."]
},
 {"tag": "kepala",
 "patterns": [ "Siapa kepala sekolah nya?", "Nama kepala sekolah nya adalah?", "kepala sekolah", "kepsek", "kepala sekolah"],
 "responses": ["Kepala sekolah SMK YPE Nusantara Slawi adalah Bapak Wahyu Wahyu Handoko,SE"]
},
 {"tag": "keunggulan",
 "patterns": [ "Apa saja keunggulan?", "Ada berapa keunggulan nya?", "keunggulan smk ype nusantara slawi", "keunggulan"],
 "responses": ["Pelopor SMK Swasta pertama dan terpercaya di Kab.Tegal, Berdiri tahun 1972 dan sudah terakreditasi B"]
}

```

Gambar 2 Dataset dalam bentuk JSON

Dataset yang digunakan memiliki struktur sebagai berikut [12]:

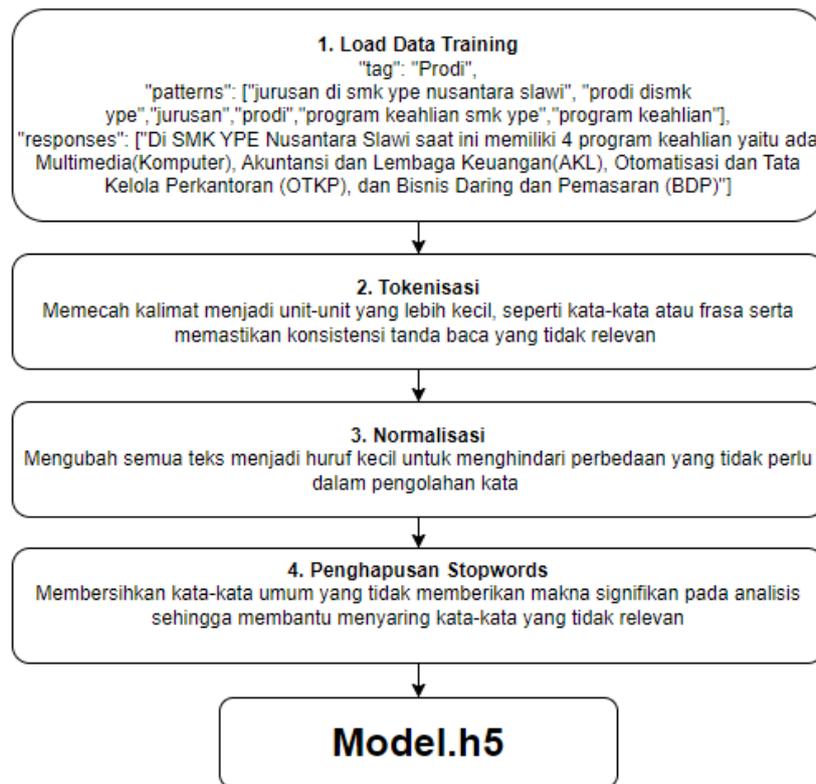
- Intents, yaitu sekumpulan data masukan dan keluaran yang digunakan untuk melatih chatbot.
- Tags, yang digunakan untuk mengelompokkan data teks yang sama sebagai output yang menjadi target dalam pelatihan jaringan neural.
- Patterns, merupakan bagian yang berisikan kata kunci atau pola input yang diharapkan dari pengguna.
- Responses, adalah bagian yang berisi pola output yang dihasilkan oleh chatbot dan akan dikirimkan kembali kepada pengguna.

2.2 Pre-processing Data

Pendekatan terbaik dengan memulai dari tahap *pre-processing* data yang sangat penting. Proses *pre-processing* ini mencakup langkah-langkah seperti pemisahan tanda baca serta eliminasi kalimat-kalimat duplikat. Setelah *pre-processing* selesai, barulah *dataset* yang telah dioptimalkan dimasukkan ke dalam model guna menjalani serangkaian eksperimen demi mencapai model yang paling optimal.

2.3 Penentuan Data Training dan Data Testing

Untuk mendapatkan model maka pada tahapan ini dilakukan dengan menentukan data training dan data testing. Data training dilakukan dengan memisahkan hasil keseluruhan *dataset* yang didapatkan yaitu 80% untuk data training dan 20% data testing. Dataset yang diperlukan untuk pelatihan model diperoleh dari sebuah file JSON yang berisi intents beserta patterns yang terkait dengan setiap intent dalam *newdataset.json*. Setiap pattern ditokenisasi menggunakan NLTK (*Natural Language Toolkit*). Kata-kata dalam setiap pattern dilematisasi (dibawa ke bentuk dasar) dan diubah menjadi huruf kecil. Kata-kata yang diabaikan seperti tanda tanya dan seru dihilangkan. Data kemudian diolah dalam format yang cocok untuk pelatihan model. Ilustrasi proses pelatihan dataset ditunjukkan pada Gambar 3.



Gambar 3 Proses Data Training

2.4. Metode Yang Diusulkan

Metode yang digunakan menggunakan *Natural Language Processing* (NLP), karena merupakan salah satu pendekatan yang sangat relevan dan kuat dalam mengolah serta menganalisis data teks secara komprehensif. Dengan memanfaatkan kemampuan NLP, aplikasi dapat menggali makna mendalam dari teks, mengenali pola-pola linguistik, dan bahkan memahami konteks yang lebih kompleks, yang semuanya berkontribusi pada pengembangan model yang lebih cerdas dan mampu memberikan hasil eksperimen yang lebih berkualitas. Tabel 1 merupakan penjelasan penggunaan *transfer learning* pada pengolahan data *chatbot* berbasis NLP dan jaringan saraf tiruan (*neural networks*).

Tabel 1 Transfer Learning Pada Chatbot NLP

Mempersiapkan data	Data yang diperlukan untuk pelatihan model diperoleh dari sebuah file JSON yang berisi intents (tujuan atau niat) beserta pola-pola (<i>patterns</i>) yang terkait dengan setiap intent. Setiap <i>pattern</i> ditokenisasi (dipecah menjadi kata-kata individual) menggunakan NLTK (<i>Natural Language Toolkit</i>). Kata-kata dalam setiap <i>pattern</i> dilematisasi (dibawa ke bentuk dasar) dan diubah menjadi huruf kecil. Kata-kata yang diabaikan seperti tanda tanya dan seru dihilangkan. Data kemudian diolah dalam format yang cocok untuk pelatihan model
Membangun Vocabulary	Kata-kata unik dalam semua <i>patterns</i> diambil dan diurutkan untuk membentuk kamus (<i>vocabulary</i>) dari kata-kata yang mungkin muncul dalam data. Kata-kata dilematisasi dan diurutkan dalam kamus.
Membangun Training Data	Data <i>patterns</i> dan <i>intents</i> dikonversi menjadi bentuk yang sesuai untuk pelatihan model. Setiap <i>pattern</i> diubah menjadi representasi " <i>bag of words</i> ," di mana setiap kata dalam kamus memiliki nilai 1 jika ada dalam <i>pattern</i> , dan 0 jika tidak ada. Label (<i>intent</i>) dari setiap <i>pattern</i> diubah

	menjadi vektor <i>one-hot encoding</i> .
Membangun Arsitektur Model	Model <i>neural networks</i> dibangun menggunakan Keras. Model terdiri dari tiga lapisan Dense (<i>connected layer</i>), dengan fungsi aktivasi ReLU di lapisan pertama dan kedua, dan <i>softmax</i> di lapisan keluaran. <i>Dropout</i> layer diterapkan setelah lapisan pertama dan kedua untuk mencegah <i>overfitting</i> .
Kompilasi dan Pelatihan Model	Model dikompilasi dengan fungsi kerugian <i>categorical cross-entropy</i> dan optimasi menggunakan <i>Stochastic Gradient Descent</i> (SGD) dengan <i>Nesterov accelerated gradient</i> . Data latihan disusun dalam batch dan dimasukkan ke dalam model selama beberapa <i>epoch</i> (iterasi) untuk melatih model. Hasil pelatihan disimpan dalam sebuah <i>variabel hist</i> .
Simpan Model	Model yang telah dilatih disimpan dalam format file <i>.h5</i> menggunakan metode <i>model.save</i> .

2.5. Evaluasi Model

Evaluasi model akan melibatkan berbagai metrik yang relevan, seperti akurasi, dan mungkin juga metrik-metrik khusus tergantung pada konteks tugas yang dihadapi. Hasil evaluasi ini akan memberikan wawasan mendalam tentang seberapa baik model berkinerja dalam memenuhi tujuan eksperimen dan mengidentifikasi area-area di mana peningkatan dapat dilakukan.

3. HASIL DAN PEMBAHASAN

Hasil pembahasan menunjukkan bahwa dataset yang digunakan dalam penelitian ini bersumber dari SMK YPE Nusantara Slawi. Dataset ini melibatkan berbagai informasi penting, seperti profil menyeluruh tentang sekolah yang mencakup visi dan misi, serta profil rinci dari setiap jurusan yang ditawarkan. Selain itu, dataset juga mencakup daftar guru-guru berpengalaman yang berkontribusi dalam pendidikan, dan juga merangkum mata pelajaran yang diajarkan secara cermat. Semua informasi berharga ini telah dianalisis dan dioptimalkan menggunakan bahasa pemrograman Python, yang telah terbukti sebagai alat yang handal dan efektif dalam proses ini.

Pada tahap *pre-processing* data dengan memanfaatkan *Natural Language Processing* (NLP), dataset dari SMK YPE Nusantara Slawi menjalani sejumlah proses penting. Tokenisasi dilakukan pada teks dataset untuk memecah kalimat menjadi unit-unit yang lebih kecil, seperti kata-kata atau frasa. Tanda baca yang tidak relevan dihapus untuk memastikan konsistensi dalam tokenisasi. Normalisasi dilakukan dengan mengubah semua teks menjadi huruf kecil untuk menghindari perbedaan yang tidak perlu dalam pengolahan kata. Penghapusan *Stopwords* merupakan proses ini adalah pembersihan kata-kata umum yang tidak memberikan makna signifikan pada analisis. Proses ini membantu menyaring kata-kata yang tidak relevan. Dalam usaha mengatasi duplikasi, dilakukan penghapusan kalimat yang serupa atau hampir identik. Proses ini bertujuan untuk menghindari pengaruh yang berlebihan dari data yang redundan dalam eksperimen selanjutnya. Selain itu, entitas-entitas penting seperti nama guru, nama mata pelajaran, dan nama jurusan diidentifikasi dan diambil menggunakan teknik entitas pengenalan (*entity recognition*) untuk mengamankan informasi yang lebih spesifik. Tahap *pre-processing* ini penting untuk mempersiapkan *dataset* secara optimal sebelum dianalisis lebih lanjut menggunakan pendekatan NLP, dan memastikan kualitas data yang diperlukan untuk eksperimen selanjutnya.

Tahap ini diawali dengan mengimpor pustaka yang dibutuhkan seperti pustaka *nlk*, *numpy*, *pickle*, *json*, dan *keras.models*. Setelah mengimpor pustaka kemudian dilanjutkan membaca data dari file dengan memecahkan setiap pola tanya menjadi unit-unit kecil yang disebut dengan token seperti ilustrasi yang tersaji pada Gambar 4.

```

import nltk
# nltk.download('popular')
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
import pickle
import numpy as np

from keras.models import load_model
model = load_model('static/model.h5')
import json
import random
intents = json.loads(open('static/newdataset.json').read())
words = pickle.load(open('static/texts.pkl','rb'))
classes = pickle.load(open('static/labels.pkl','rb'))

def clean_up_sentence(sentence):
    # tokenize the pattern - split words into array
    sentence_words = nltk.word_tokenize(sentence)
    # stem each word - create short form for word
    sentence_words = [lemmatizer.lemmatize(word.lower()) for word in sentence_words]
    return sentence_words

```

Gambar 4 Pre-processing Data

Setelah melalui tahap *pre-processing* yang cermat dengan pendekatan NLP, dataset dari SMK YPE Nusantara Slawi siap untuk menjalani tahap pelatihan (*training*) dan pengujian (*testing*) pada model yang akan dikembangkan. Dataset telah dibagi menjadi dua subset yaitu subset pelatihan dan subset pengujian. Pada subset pelatihan, model akan diberikan data-data yang telah dioptimalkan untuk memahami pola-pola dan karakteristik dalam teks. Terdapat proses Lemmatisasi yang bertujuan agar dapat dianalisis menjadi satu item sehingga dapat diidentifikasi dalam bentuk kamus kata yang diilustrasikan pada Gambar 5.

```

# lemmatize and lower each word and remove duplicates
words = [lemmatizer.lemmatize(w.lower()) for w in words if w not in ignore_words]
words = sorted(list(set(words)))
# sort classes
classes = sorted(list(set(classes)))
# documents = combination between patterns and intents
print (len(documents), "documents")
# classes = intents
print (len(classes), "classes", classes)
# words = all words, vocabulary
print (len(words), "unique lemmatized words", words)

```

Gambar 5 Proses Lemmatisasi

Model akan "belajar" dari dataset ini untuk mengenali hubungan antara atribut-atribut. Proses pelatihan ini akan memungkinkan model untuk mengembangkan pemahaman yang lebih mendalam tentang dataset dan tugas yang dihadapi yang diilustrasikan pada Gambar 6. Kemudian, subset pengujian digunakan untuk menguji kemampuan generalisasi model. Data yang belum pernah dilihat sebelumnya akan diberikan kepada model, dan model harus mampu menerapkan apa yang telah dipelajari selama pelatihan untuk melakukan prediksi dengan akurasi yang tinggi. Hasil dari pengujian ini akan memberikan wawasan tentang seberapa baik model bekerja dalam kondisi dunia nyata dan seberapa baik model dapat mengatasi data-data baru.

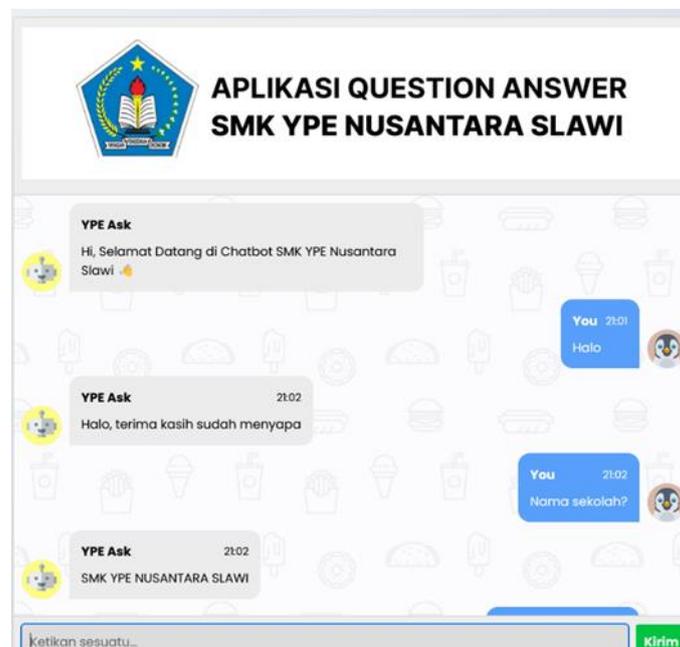
Penting untuk mencatat bahwa dalam proses ini, metrik evaluasi seperti akurasi, dan loss yang akan digunakan untuk mengukur kinerja model. Hasil evaluasi ini akan memberikan pandangan komprehensif tentang sejauh mana model dapat memenuhi tujuan eksperimen dan menghasilkan hasil yang handal dan berarti.

```
# create our training data
training = []
# create an empty array for our output
output_empty = [0] * len(classes)
# training set, bag of words for each sentence
for doc in documents:
    # initialize our bag of words
    bag = []
    # list of tokenized words for the pattern
    pattern_words = doc[0]
    # lemmatize each word - create base word, in attempt to represent related words
    pattern_words = [lemmatizer.lemmatize(word.lower()) for word in pattern_words]
    # create our bag of words array with 1, if word match found in current pattern
    for w in words:
        bag.append(1) if w in pattern_words else bag.append(0)

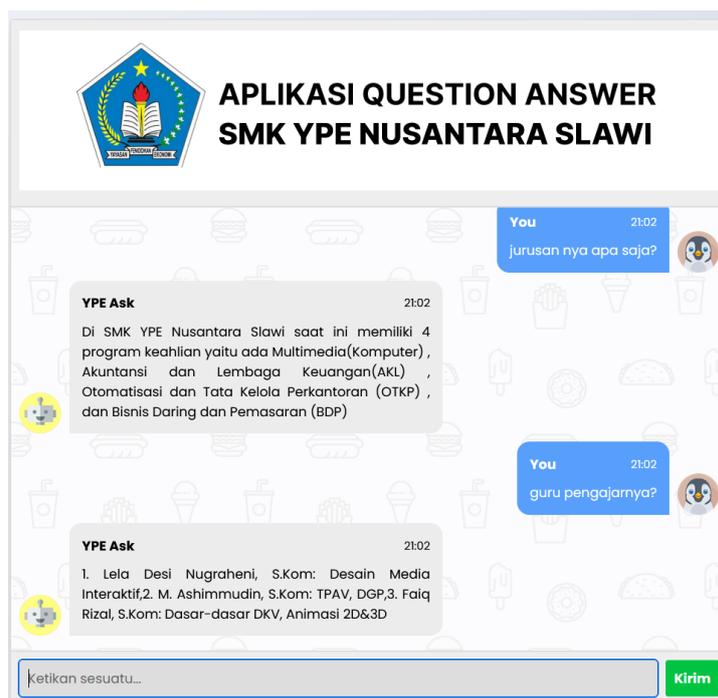
    # output is a '0' for each tag and '1' for current tag (for each pattern)
    output_row = list(output_empty)
    output_row[classes.index(doc[1])] = 1
```

Gambar 6 Proses *Data Training*

Dari hasil penelitian ini dapat dilihat tampilan dari *chatbot* SMK YPE Nusantara Slawi pada Gambar 7 dan Gambar 8.



Gambar 7 Tampilan utama halaman chatbot



Gambar 8 Tampilan respon pertanyaan chatbot

Pada tahap pembahasan ini, fokus ditempatkan pada hasil eksplorasi dan penerapan teknik *Natural Language Processing* (NLP) terhadap dataset yang telah disebutkan sebelumnya. Penggunaan NLP dalam pengembangan chatbot menjadi elemen sentral yang membantu mengurai dan memahami berbagai aspek dataset dengan lebih mendalam. Beberapa poin penting dalam hasil pembahasan ini:

- Melalui pendekatan NLP, dataset ini diolah untuk mendapatkan informasi yang lebih kaya dan terstruktur. Proses ekstraksi informasi melibatkan identifikasi entitas seperti profil sekolah, visi dan misi, profil jurusan, daftar guru, dan mata pelajaran. Penggunaan teknik entitas pengenalan dalam NLP membantu mengambil informasi relevan dari teks yang tersebar dalam dataset.
- Dengan bantuan NLP, dataset ini mampu membentuk pengetahuan lebih dalam tentang karakteristik sekolah, tujuan-tujuan strategisnya, dan ciri-ciri khas dari setiap jurusan yang ditawarkan. Pengetahuan ini dapat digunakan untuk memberikan wawasan yang lebih kaya dan mendalam kepada pengguna chatbot.
- Melalui penerapan NLP dalam chatbot, interaksi real-time dengan pengguna menjadi mungkin. Pengguna dapat mengajukan pertanyaan tentang visi dan misi sekolah, mengenai spesifikasi jurusan tertentu, atau mendapatkan informasi terbaru tentang daftar guru dan mata pelajaran yang sedang diajarkan. Ini memberikan pengalaman berinteraksi yang lebih responsif dan interaktif.
- Dengan memanfaatkan kemampuan NLP, chatbot dapat merespons pertanyaan dan permintaan pengguna dengan cara yang lebih kontekstual. Misalnya, jika pengguna bertanya tentang visi dan misi sekolah, chatbot dapat memberikan jawaban yang sesuai dengan konteks informasi dalam dataset.
- Penggunaan NLP memungkinkan chatbot untuk merangkum informasi kompleks menjadi format yang lebih ringkas dan mudah dimengerti. Ini memungkinkan pengguna untuk dengan cepat mendapatkan informasi yang mereka butuhkan tanpa harus membaca seluruh dataset secara rinci.
- Melalui implementasi NLP dalam pengembangan chatbot, dataset ini mampu memberikan layanan interaktif yang lebih canggih dan berarti kepada pengguna. Chatbot dapat mengenali maksud pengguna, mengekstrak informasi penting dari dataset, dan

memberikan jawaban yang relevan dan kontekstual. Dengan demikian, pendekatan ini memanfaatkan potensi NLP untuk meningkatkan aksesibilitas dan pengalaman pengguna dalam mendapatkan informasi dari dataset yang kaya dan beragam.

4. KESIMPULAN DAN SARAN

Penelitian ini menggunakan *dataset* yang berasal dari SMK YPE Nusantara Slawi untuk mengembangkan dan menerapkan pendekatan *Natural Language Processing* (NLP) dalam pengembangan *chatbot*. *Dataset* ini mencakup informasi tentang profil sekolah, visi dan misi, profil jurusan, daftar guru, dan mata pelajaran. *Dataset* tersebut telah melalui tahap *pre-processing* NLP yang meliputi tokenisasi, normalisasi, penghapusan stopwords, serta penghapusan duplikasi dan identifikasi entitas penting. Setelah *pre-processing*, *dataset* dibagi menjadi subset pelatihan dan subset pengujian untuk pelatihan dan pengujian model. Berdasarkan hasil penelitian ini menunjukkan bahwa pendekatan NLP dalam pengembangan *chatbot* memiliki beberapa implikasi penting yaitu

- Data yang semula kompleks dan terstruktur diolah menjadi informasi yang lebih kaya dan terstruktur melalui ekstraksi entitas menggunakan NLP.
- Pengetahuan lebih mendalam tentang sekolah dan jurusan dihasilkan, memungkinkan *chatbot* memberikan wawasan mendalam kepada pengguna.
- Interaksi real-time dengan pengguna menjadi mungkin, memungkinkan pengguna mendapatkan informasi yang mereka butuhkan dengan cepat.
- *Chatbot* mampu merespons pertanyaan pengguna dengan konteks yang sesuai, menghasilkan jawaban yang relevan.
- Informasi kompleks dapat dirangkum menjadi format yang lebih ringkas dan mudah dimengerti, meningkatkan aksesibilitas informasi

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Pusat Penelitian dan Pengabdian Masyarakat (P3M) Politeknik Harapan Bersama Tegal yang telah mendanai penelitian ini melalui skema Hibah Penelitian Institusi dengan nomor kontrak 037.16/P3M.PHB/V/2023.

DAFTAR PUSTAKA

- [1] R. Bangun, C. Informasi, and L. Pekerjaan, "Rancang Bangun *Chatbot* Informasi Lowongan Pekerjaan Berbasis Whatsapp dengan Metode NLP (Natural Language Processing)," vol. 5, no. 1, pp. 619–626, 2020.
- [2] O. Al and H. Tangerang, "Penerapan *Chatbot* Menggunakan Aimi Berbasis Android Pada Smk," no. November, pp. 27–33, 2019.
- [3] J. Hill, W. Randolph Ford, and I. G. Farreras, "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–*chatbot* conversations," *Computers in Human Behavior*, vol. 49, pp. 245–250, 2015, doi: <https://doi.org/10.1016/j.chb.2015.02.026>.
- [4] N. N. Khin and K. M. Soe, "University *Chatbot* using Artificial Intelligence Markup Language," in *2020 IEEE Conference on Computer Applications (ICCA)*, 2020, pp. 1–5. doi: 10.1109/ICCA49400.2020.9022814.
- [5] B. Rusmarasy, B. Priyambadha, and F. Pradana, "Pengembangan *Chat Bot* pada CoMa untuk memberikan motivasi kepada pengguna menggunakan AIML," vol. 3, no. 5, pp. 4484–4490, 2019.
- [6] J. Ilmiah, I. Komputa, E. N. S. C. P, I. Afrianto, J. Dipati, and U. No, "Wisata Kota Bandung Dengan Pendekatan *Natural Language Processing* Program Studi Teknik Informatika

- Fakultas Teknik dan Ilmu Komputer Jurnal Ilmiah Komputer dan Informatika (KOMPUTA) Teknik analisis data dalam pembuatan perangkat,” vol. 4, no. 1, pp. 49–54, 2015.
- [7] J. Aprian and A. Budi, “Jurnal Informatika dan Bisnis ISSN 2301-9670 Penerapan Natural Language Processing Dalam Aplikasi Chatbot Sebagai Media Pencarian Informasi Dengan Menggunakan React (Studi Kasus : Institut Bisnis Dan Informatika Kwik Kian Gie) Jurnal Informatika dan Bis”.
- [8] N. Munasatya and S. Novianto, “Natural Language Processing untuk Sentimen Analisis Presiden Jokowi Menggunakan Multi Layer Perceptron,” *Techno.Com*, vol. 19, no. 3, pp. 237–244, 2020, doi: 10.33633/tc.v19i3.3630.
- [9] G. Huang and K. Q. Weinberger, “Densely Connected Convolutional Networks”.
- [10] S. Sains, I. Gunawan, H. Asnal, M. N. Cahyadi, and Z. M. G, “Perancangan Chatbot untuk Layanan Informasi Sekolah (Studi Kasus SMK Dwi Sejahtera Pekanbaru),” 2022, doi: 10.33372/stn.v8i2.880.
- [11] A. Ho, J. Hancock, and A. S. Miner, “Psychological, Relational, and Emotional Effects of Self-Disclosure After Conversations With a Chatbot.,” *The Journal of communication*, vol. 68, no. 4, pp. 712–733, Aug. 2018, doi: 10.1093/joc/jqy026.
- [12] A. Hikmah *et al.*, “Implementasi Natural Language Processing Pada Chatbot Untuk Layanan Akademik,” vol. 10, no. 1, pp. 371–382, 2023.
- [13] S. Hidayat, A. Silvanie, and R. Subekti, “Chatbot Untuk Konsultasi Akademik Menggunakan Natural Language Processing,” vol. 6, no. 2, pp. 396–410, 2022, doi: 10.52362/jisamar.v6i2.780.
- [14] I. M. De Andrade and C. Tumelero, “Increasing customer service efficiency through artificial intelligence chatbot,” vol. 29, no. 3, pp. 238–251, 2022, doi: 10.1108/REG-07-2021-0120.
- [15] A. A. Hania and Y. T. Indonesia, “Mengenal Artificial Intelligence , Machine Learning , Neural Network , dan Deep Learning,” no. June, 2017.