

**IMPLEMENTASI *CHATBOT* PADA *WEBSITE* PMB UIN MALANG
MENGUNAKAN *NEURAL NETWORK***

SKRIPSI

**Oleh :
IKHWATA ANDY PRATAMA
NIM. 18650049**



**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

**IMPLEMENTASI *CHATBOT* PADA *WEBSITE* PMB UIN MALANG
MENGUNAKAN *NEURAL NETWORK***

SKRIPSI

**Diajukan Kepada:
Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Malang
untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**Oleh :
IKHWATA ANDY PRATAMA
NIM. 18650049**

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

LEMBAR PERSETUJUAN

IMPLEMENTASI *CHATBOT* PADA *WEBSITE* PMB UIN MALANG MENGUNAKAN *NEURAL NETWORK*

SKRIPSI

Oleh :
IKHWATA ANDY PRATAMA
NIM. 18650049

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal : 03 Juni 2022

Dosen Pembimbing I



Puspa Miladin N.S.A.B, M.Kom
NIP. 19930828 20190 3 2018

Dosen Pembimbing II



Dr. Muhammad Faisal
NIP. 19740510 2005011 007

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Fachrul Kurniawan, M.MT
NIP. 19771020 200912 1 001

LEMBAR PENGESAHAN

IMPLEMENTASI *CHATBOT* PADA *WEBSITE* PMB UIN MALANG MENGUNAKAN *NEURAL NETWORK*

SKRIPSI

Oleh:
IKHWATA ANDY PRATAMA
NIM. 18650049

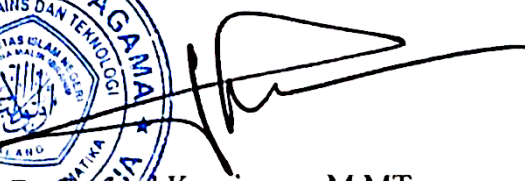
Telah Dipertahankan di Depan Dewan Penguji
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Pada Tanggal: 8 Juni 2022

Susunan Dewan Penguji

Penguji Utama	: <u>A'la Syauqi, M.Kom</u> NIP. 19771201 200801 1 007	()
Ketua Penguji	: <u>Johan Ericka Wahyu Prakasa, M.Kom</u> NIDT. 19831213 201903 1 004	()
Sekretaris Penguji	: <u>Puspa Miladin N. S. A. B., M.Kom</u> NIP. 19930828 20190 3 2018	()
Anggota Penguji	: <u>Dr. Muhammad Faisal</u> NIP. 19740510 2005011 007	()

Mengetahui dan Mengesahkan,
Ketua Jurusan Teknik Informatika




Dr. Fachrul Kurniawan, M.MT
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama : Ikhwata Andy Pratama
NIM : 18650049
Fakultas : Sains dan Teknologi
Jurusan : Teknik Informatika
Judul Skripsi : Implementasi *Chatbot* Pada *Website* PMB UIN
Malang Menggunakan *Neural Network*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 3 Juni 2022
Yang membuat pernyataan,



Ikhwata Andy Pratama
NIM. 18650049

MOTTO

“ Bersyukurlah tiap menit, hidupmu akan dimudahkan.”

HALAMAN PERSEMBAHAN

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ

Alhamdulillah, terimakasih telah berjuang sampai dengan titik ini.

KATA PENGANTAR

Assalamu alaikum, Wr. Wb.

Alhamdulillah, puji dan syukur kehadiran Allah SWT. Berkat kesehatan dan bimbingannya, penulis berkesempatan menyelesaikan skripsi ini. Sholawat dan salam semoga selalu tercurah kepada junjungan kita Nabi Muhammad SAW, semoga kita tergolong orang yang beriman.

Penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dalam penyusunan skripsi ini. Oleh karena itu, dengan ini saya mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Bapak Dr. Fachrul Kurniawan, M.MT., selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang senantiasa memberikan dorongan.
3. Ibu Puspa Miladin Nuraida A.Basid, M.Kom., selaku dosen pembimbing I yang telah memberi arahan serta membimbing sehingga penulis dapat menyelesaikan pengerjaan skripsi.
4. Bapak Dr. Muhammad Faisal., selaku dosen pembimbing II telah membimbing dan memberi arahan dan masukkan kepada penulis sehingga penulis dapat menyelesaikan skripsi.
5. Segenap Bapak dan Ibu Dosen Jurusan Teknik Informatika UIN Malang, yang telah memberi arah dan memberikan ilmunya.
6. Keluarga penulis terutama orang tua yang selalu membantu dan memberikan motivasi dan dukungan agar penulisan skripsi dapat selesai tepat waktu.
7. Jasiva Nandya Pratiwi yang sudah selalu memberikan motivasi, semangat, dan dorongan sehingga penulis dapat menyelesaikan skripsi tepat waktu.
8. Teman-teman Unity of Informatics Force (UFO) 2018, khususnya Muhammad Meganata Adam, Lingga Wahyu Rochim, Abd. Qohar, Muhamad Syifa', Rizki Fitriani, Inna Fatimatuzzahro, Naila Nahdiyah, Anisa Rizkiana Putri, Aulia Ananda Salsabila yang bahu membahu dan saling mendukung dalam berjuang bersama untuk menyelesaikan tugas mata kuliah dari awal masuk jurusan hingga menyelesaikan skripsi tepat waktu.
9. Teman-teman pengurus komunitas UINUX, dan pengurus himpunan mahasiswa jurusan periode 2019 dan 2020, yang sudah memberikan support dan pengalaman yang berharga.
10. Semua pihak yang telah membantu dalam menyelesaikan skripsi ini yang tidak dapat disebutkan satu persatu.
11. Penulis sendiri yang telah berusaha sebaik mungkin agar skripsi ini dapat selesai tepat waktu dan percaya dengan kemampuan tanpa harus melihat orang lain.

Skripsi yang telah ditulis ini masih jauh dari kata sempurna, oleh karena itu penulis sangat menghargai dan senang jika terdapat kritik dan saran. Semoga skripsi ini memberikan manfaat.

Wassalamu alaikum, Wr. Wb.

Malang, 3 Juni 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	ii
LEMBAR PERSETUJUAN.....	iii
LEMBAR PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN.....	v
MOTTO.....	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR	ii
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
ABSTRAK	viii
<i>ABSTRACT</i>	ix
المخلص.....	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Pernyataan Masalah.....	4
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	4
BAB II STUDI PUSTAKA.....	6
2.1 Penelitian Terdahulu	6
2.2 <i>Chatbot</i>	8
2.3 <i>Natural Language Processing</i>	9
2.4 <i>Neural Network (NN)</i>	11
BAB III METODE PENELITIAN.....	13
3.1 Pengumpulan Data	13
3.2 Desain Sistem	14
3.3 <i>Preprocessing</i>	15
3.4 <i>Neural Network</i>	22
3.5 <i>Pelatihan Neural Network</i>	33
BAB IV UJI COBA DAN PEMBAHASAN	36
4.1 Skenario Uji Coba	36
4.2 Tahap Pelatihan	36
4.3 <i>Output</i>	42
4.4 Implementasi Desain Antarmuka	44
4.4.1 Halaman Beranda.....	44
4.4.2 <i>Chatbot Messenger</i>	44
4.4.3 <i>Login Admin</i>	45
4.4.4 <i>Dashboard Admin</i>	46
4.4.5 <i>Tambah Data</i>	46
4.4.6 <i>Edit Data</i>	47
4.5 Hasil Uji Coba.....	47
4.6 Pembahasan.....	50
BAB V KESIMPULAN DAN SARAN.....	53

5.1 Kesimpulan.....	53
5.2 Saran.....	53
DAFTAR PUSTAKA	
LAMPIRAN	
Lampiran I	
Lampiran II	
Lampiran III	

DAFTAR GAMBAR

Gambar 3. 1 Struktur <i>Training Set</i>	13
Gambar 3. 2 Desain Sistem	14
Gambar 3. 3 <i>Flowchart Preprocessing</i>	15
Gambar 3. 4 Alur <i>Case Folding</i>	16
Gambar 3. 5 Contoh <i>Case Folding</i>	17
Gambar 3. 6 Alur Proses <i>Tokenizing</i>	17
Gambar 3. 7 Contoh <i>Tokenizing</i>	18
Gambar 3. 8 Alur Proses <i>Stemming</i>	19
Gambar 3. 9 Contoh <i>Stemming</i>	20
Gambar 3. 10 Alur Proses <i>Bag of Words</i>	20
Gambar 3. 11 Contoh Konversi Kata Menjadi <i>Biner</i>	22
Gambar 3. 12 Representasi <i>Neural Network</i>	22
Gambar 3. 13 Arsitektur <i>Neural Network</i>	26
Gambar 3. 14 Model <i>Forward Pass</i>	29
Gambar 3. 15 <i>Forward Pass Input Layer</i> Menuju <i>Hidden Layer 1</i>	31
Gambar 3. 16 <i>Forward Pass Hidden Layer 1</i> Menuju <i>Hidden Layer 2</i>	32
Gambar 3. 17 <i>Forward Pass Hidden Layer 2</i> Menuju <i>Output Layer</i>	33
Gambar 3. 18 <i>Flowchart Pelatihan Neural Network</i>	34
Gambar 4. 1 Potongan Kode Program Model <i>Neural Network</i>	39
Gambar 4. 2 Potongan Kode Program Melatih <i>Neural Network</i>	40
Gambar 4. 3 Grafik Error	40
Gambar 4. 4 Grafik Akurasi	41
Gambar 4. 5 Hasil Pelatihan	41
Gambar 4. 6 Potongan Kode Proses Penyimpanan Model Pelatihan	42
Gambar 4. 7 Kode Program Respon <i>Chatbot</i>	43
Gambar 4. 8 Tampilan Beranda	44
Gambar 4. 9 Tampilan <i>Chatbot</i>	45
Gambar 4. 10 Login Admin	45
Gambar 4. 11 Dashboard Admin	46
Gambar 4. 12 Tambah Data	46
Gambar 4. 13 Edit Data	47

DAFTAR TABEL

Tabel 3. 1 Proses <i>Bag of Words</i>	21
Tabel 3. 2 Arsitektur <i>Neural Network</i>	30
Tabel 4. 1 Inisiasi Nilai Model.....	36
Tabel 4. 2 Perbedaan Penggunaan Kata	43
Tabel 4. 3 Data Training	48
Tabel 4. 4 Data Pengujian	48
Tabel 4. 5 Hasil <i>Confusion Matrix</i>	49

ABSTRAK

Pratama, Ikhwata Andy. 2022. **Implementasi Chatbot Pada Website PMB UIN Malang Menggunakan Neural Network**. Skripsi. Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Puspa Miladin N.S.A.B, M.Kom (II) Dr. Muhammad Faisal.

Kata Kunci: *chatbot, neural network*, penerimaan mahasiswa baru

UIN Malang memiliki *website* yang khusus untuk pendaftaran mahasiswa baru yaitu *website* PMB UIN Malang. Pada *website* PMB UIN Malang menampilkan informasi pendaftaran program sarjana dan pascasarjana yang dapat diakses dengan mudah oleh pengguna. Tetapi pada *website* tersebut masih belum adanya komunikasi dua arah antara calon mahasiswa dan admin sehingga *website* tersebut kurang interaktif. Pesatnya perkembangan teknologi zaman sekarang dapat membantu layanan informasi yang dikerjakan manusia dapat dikerjakan oleh program komputer salah satunya dengan menggunakan AI. AI atau *Artificial Intelligence* adalah program komputer yang dapat melakukan tugas seperti manusia salah contoh adalah *chatbot*. *Chatbot* dapat berupa teks atau percakapan lisan. Terdapat beberapa metode untuk implementasi *chatbot* salah satunya yaitu *neural network*. *Neural network* adalah jaringan untuk meniru fungsi otak manusia dan sering disebut dengan jaringan syaraf tiruan. Terdapat beberapa tahapan dalam membuat model *neural network* yaitu *input*, *preprocessing*, dan *output*. *Input* adalah data pertanyaan yang dimasukkan oleh pengguna. *Preprocessing* menggunakan pendekatan *Natural Language Processing* (NLP) dalam prosesnya untuk mengolah kata. Terdapat beberapa proses pada *preprocessing* yaitu *case folding*, *tokenizing*, *stemming*, dan *bag of words*. Setelah proses *preprocessing*, data akan dilatih untuk dapat menjawab pertanyaan pengguna. Bahasa pemrograman yang digunakan adalah *Python* dan *Flask* agar dapat terhubung dengan *chatbot* yang sudah dibangun dan terintegrasi pada *website*. Pengujian menggunakan 25 data uji dan didapatkan nilai akurasi sebesar 88%, nilai *recall* sebesar 91%, nilai presisi sebesar 95%, dan nilai *f1 score* sebesar 93%.

ABSTRACT

Pratama, Ikhwata Andy. 2022. **Implementation of Chatbot on PMB UIN Malang Website Using Neural Network**. Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Maulana Malik Ibrahim Islamic State University Malang. Supervisor: (1) Puspa Miladin N.S.A.B, M.Kom (II) Dr. Muhammad Faisal.

Keyword: chatbot, neural network, admission of new students

UIN Malang has a special website for new student registration, namely the PMB UIN Malang website. The PMB UIN Malang website displays registration information for undergraduate and postgraduate programs that can be accessed easily by users. But on the website there is still no two-way communication between prospective students and the admin so that the website is less interactive. The rapid development of today's technology can help information services that are done by humans can be done by computer programs, one of which is by using AI. AI or Artificial Intelligence is a computer program that can perform tasks like humans, one example is a chatbot. Chatbots can be text or spoken conversations. There are several methods for implementing chatbots, one of which is a neural network. Neural network is a network to imitate the function of the human brain and is often referred to as an artificial neural network. There are several stages in making a neural network model, namely input, preprocessing, and output. Input is the question data entered by the user. Preprocessing uses the Natural Language Processing (NLP) approach in the process to process words. There are several processes in preprocessing, namely case folding, tokenizing, stemming, and bag of words. After the preprocessing process, the data will be trained to be able to answer user questions. The programming languages used are Python and Flask in order to connect with chatbots that have been built and integrated on the website. The test used 25 test data and obtained an accuracy value of 88%, a recall value of 91%, a precision value of 95%, and an f1 score of 93%.

الملخص

براتاما ، اخواتا اندي. 2022. تطبيق موقع الويب الخاص بنا على شبكة الإنترنت. سكريبيسي. جوروسان
تكنيك إنفورماتيك ، فاكولتاس سينز دان تكنولوجي ، جامعة الإسلام نيجري مولانا مالك إبراهيم مالانج.
..بيمبيمينج: (1) بوسبا ميلادين إن إس إيه بي ، إم كوم (الثاني) الدكتور محمد فيصل

كاتا كوني: تشاتبوت ، الشبكة العصبية ، بينيريمان ماهاسيسوا بارو

وين مالانج لديها موقع خاص لتسجيل الطلاب الجدد ، وهي بمب أوين مالانج الموقع. على موقع بمب أوين مالانج ، معلومات التسجيل لبرامج البكالوريوس والدراسات العليا يمكن الوصول إليها بسهولة من قبل المستخدمين. ولكن على الموقع لا يوجد حتى الآن اتصال ثنائي الاتجاه بين الطلاب والمسؤولين المحتملين بحيث يكون الموقع أقل تفاعلية. التطور السريع للتكنولوجيا اليوم يمكن أن تساعد خدمات المعلومات التي يقوم بها البشر يمكن أن يتم عن طريق برامج الكمبيوتر ، واحدة منها هو باستخدام منظمة العفو الدولية. الذكاء الاصطناعي أو الذكاء الاصطناعي هو برنامج كمبيوتر يمكنه أداء مهام مثل الإنسان. يمكن أن تكون روبوتات الدردشة عبارة عن محادثات نصية أو منطوقة. هناك عدة طرق لتنفيذ روبوتات المحادثة ، أحدها الشبكات العصبية. الشبكة العصبية هي شبكة لتقليد وظيفة الدماغ البشري وغالبا ما تسمى الشبكة العصبية الاصطناعية. هناك عدة مراحل في إنشاء نموذج الشبكة العصبية ، وهي الإدخال والمعالجة المسبقة والإخراج. المدخلات هي بيانات السؤال التي أدخلها المستخدم. تستخدم المعالجة المسبقة نهج معالجة اللغة الطبيعية (البرمجة اللغوية العصبية) في عملية معالجة الكلمات. هناك العديد من العمليات في المعالجة المسبقة ، وهي طي الحالة ، والترميز ، والنوع ، وحقيبة الكلمات. بعد عملية المعالجة المسبقة ، سيتم تدريب البيانات لتكون قادرة على الإجابة على أسئلة المستخدم. لغات البرمجة المستخدمة هي بايثون وقارورة من أجل التواصل مع روبوتات الدردشة التي تم إنشاؤها ودمجها على الموقع. استخدم الاختبار 25 بيانات اختبار وحصل على قيمة دقة 88 ٪ ، وقيمة استدعاء 91 ٪ ، وقيمة دقة 95 ٪ ، وقيمة درجة الفورمولا 1 93٪.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Website ialah kumpulan situs yang dapat diakses oleh publik yang saling tergabung di sebuah domain (Jonathan & Lestari, 2015). Informasi yang terdapat dalam *website* umumnya seperti teks, video, dan gambar untuk melayani berbagai macam kepentingan. Pada halaman utama atau *homepage* biasanya menampilkan informasi umum sebuah *website* seperti identitas, berita, dan fitur lain nya. Salah satu yang menggunakan *website* sebagai media informasi adalah *website* penerimaan mahasiswa baru (Guntoro et al., 2020).

UIN Maulana Malik Ibrahim Malang memiliki *website* yang khusus untuk pendaftaran mahasiswa baru yaitu *website* PMB UIN Malang. Pada *website* PMB UIN Malang menampilkan informasi pendaftaran program sarjana dan pascasarjana yang dapat diakses dengan mudah oleh pengguna. Tetapi pada *website* tersebut masih belum adanya komunikasi dua arah antara calon mahasiswa dan petugas PMB serta belum terdapat FAQ (*Frequently Asked Questions*) yang menyebabkan *website* tersebut masih kurang interaktif. Ada beberapa layanan dapat digunakan untuk menjawab pertanyaan pada *website* seperti *chatting*, telepon dan lain-lain (Guntoro et al., 2020). Tetapi layanan tersebut belum dapat dilakukan 24 jam karena keterbatasan waktu kerja. Sehingga pertanyaan yang disampaikan tidak dapat dijawab dengan instan dan harus menunggu sampai petugas sudah siap untuk menjawab pertanyaan (Cahya et al., 2021).

Pesatnya perkembangan teknologi di zaman sekarang dapat membantu layanan informasi yang dikerjakan manusia dapat dikerjakan oleh program komputer salah satunya dengan menggunakan *Artificial Intelligence*. *Artificial Intelligence* adalah program komputer yang mampu melakukan tugas seperti manusia salah contoh adalah *chatbot* (Khin & Soe, 2020). *Chatbot* merupakan kecerdasan buatan yang mampu berkomunikasi dengan manusia. *Chatbot* dapat berupa teks atau percakapan lisan (Lalwani et al., 2018). *Chatbot* dapat berguna untuk membantu merespon pertanyaan yang diajukan oleh calon mahasiswa dengan instan dan cepat serta tervalidasi kebenarannya.

Dalam surah Al Hujarat ayat 6 yang berbunyi :

يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَن تُصِيبُوا قَوْمًا
بِجَهَالَةٍ فَتُصْحَبُوا عَلَىٰ مَا فَعَلْتُمْ نَادِمِينَ

“Wahai orang-orang yang beriman! Jika seseorang yang fasik datang kepadamu membawa suatu berita, maka telitilah kebenarannya, agar kamu tidak mencelakakan suatu kaum karena kebodohan (kecerobohan), yang akhirnya kamu menyesali perbuatanmu itu.”

Menurut tafsir Ibnu Katsir, menjelaskan kita harus lebih teliti dalam menerima informasi dan jangan mudah percaya atas informasi yang ada karena Allah telah melarang orang beriman untuk menuruti orang yang salah.

Berdasarkan ayat di atas memiliki kesesuaian bahwa informasi yang ada harus jelas dan benar serta jika merasa bingung akan suatu informasi hendak bertanya kepada orang yang terpercaya atau orang yang tepat. Dalam penelitian ini, *website* harus memberikan informasi yang dapat dimengerti oleh pengguna agar tidak terjadi kebingungan atau kesalahan saat membaca informasi. Jika

pengguna merasa kebingungan saat membaca informasi dapat berinteraksi melalui *chatbot* karena informasi nya bisa dipercaya.

Berbagai macam teknologi yang dapat digunakan untuk implementasi *chatbot* yaitu *Machine Learning*, *Artificial Intelligence*, *Deep learning*, dan *Natural Language Processing*. *Machine Learning* dalam pengaplikasiannya digunakan sebagai mesin yang dapat membantu menganalisa dan mempelajari pola yang dibentuk untuk membuat *chatbot*. Untuk dapat mengerti dan memahami bahasa manusia dapat menggunakan *Natural Language Processing* (Fambayun et al., 2022). Bentuk komunikasi yang digunakan berbentuk suara atau ucapan dan berupa tulisan.

Salah satu algoritma yang digunakan pada *Machine Learning* untuk membuat *chatbot* yaitu *Neural Network*. *Neural Network* merupakan jaringan untuk meniru fungsi otak manusia dan sering disebut dengan jaringan syaraf tiruan. Otak manusia terdiri dari jutaan unit neuron yang dapat bekerja secara paralel. *Neuron* atau sel saraf terdiri dari node yang dapat mengkombinasikan inputan nya (Sibagariang et al., 2021). Pada penelitian dilakukan *training* model menggunakan modul TensorFlow. Pada penelitian yang menggunakan *Neural Network* (NN), *Latent Semantic Index*, dan *Jaccard*. *Jaccard* digunakan sebagai nilai actual dan menghasilkan metode *Neural Network* lebih mendekati metode *Jaccard* dalam pendeteksian kesamaan teks dibandingkan dengan metode *Latent Semantic Index* (Putra et al., 2021).

Berdasarkan uraian diatas, penulis akan menerapkan metode *Neural Network* untuk membangun *chatbot* pada *website* PMB UIN Malang.

1.2 Pernyataan Masalah

Pernyataan masalah yang diangkat adalah berapa nilai akurasi, presisi, *recall*, dan *f1-score* pada *chatbot website* PMB UIN Malang menggunakan algoritma *Neural Network*?

1.3 Tujuan Penelitian

Tujuan penelitian yang diangkat adalah untuk mengukur nilai akurasi, presisi, *recall*, dan *f1-score* pada *chatbot website* PMB UIN Malang menggunakan algoritma *Neural Network*.

1.4 Batasan Masalah

Batasan masalah pada skripsi berguna supaya pembahasan dapat dibatasi agar tidak keluar dari topik. Berikut batasan masalah dalam skripsi ini :

1. Kata pada *chatbot* menggunakan Ejaan Bahasa Indonesia
2. Objek dari penelitian ini menggunakan *website* PMB UIN Malang
3. Untuk daftar pertanyaan penulis menggunakan pertanyaan yang sering ditanyakan pada Google pada tahun 2022 mengenai pendaftaran mahasiswa baru di UIN Malang

1.5 Manfaat Penelitian

Berikut adalah manfaat penelitian yang diharapkan penulis dalam penelitian ini :

1. Manfaat bagi pengguna, dapat mempermudah calon mahasiswa baru untuk mengajukan pertanyaan mengenai penerimaan mahasiswa baru di UIN Malang.

2. Manfaat bagi peneliti, dapat mengetahui penerapan *chatbot* menggunakan algoritma *Neural Network*.
3. Manfaat bagi universitas, dapat mempermudah petugas menjawab pertanyaan pengguna tanpa harus *stand by* 24 jam.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terdahulu

Lahitani et al., (2016), mereka menggunakan metode untuk kesamaan menggunakan *cosine similarity*. Mereka mendapatkan ekstraksi fitur dari dokumen dengan langkah-langkah pre processing, lalu menghitung bobot istilah menggunakan TF-IDF dan menunjukkan tingkat kesamaan ukuran dalam bentuk yang berbasis peringkat. Dengan menerapkan prinsip kesamaan, pengenalan karakter *text* yang sama dari dokumen ahli untuk menguji dokumen akan memberikan penilaian yang objektif dan dapat mempercepat proses pengecekan dalam kategori tugas essay.

Bhartiya et al., (2019), mereka mengimplementasikan *Artificial Neural Network* (ANN) pada *chatbot* dan di implementasikan pada konseling universitas. *Chatbot* yang digunakan dapat merespon dengan baik seputar bidang teknik di universitas. Dengan *preprocessing* menggunakan *Natural Language Processing* (NLP) dengan tahapan *tokenization*, *stemming*, dan *regex extraction*. Data universitas dikembangkan dalam format JSON dan *feedforward neural model* untuk melatih dataset. Hasil pengujian yang didapat bahwa kinerja *chatbot* telah meningkat secara signifikan dari tahap pertama ke tahap kedua pengujian dengan skor probabilitas meningkat menjadi 0,72 dari 0,46 untuk jawaban yang benar.

Zuraiyah et al., (2019), mengimplementasikan *chatbot* menggunakan *Recurrent Neural Network* (RNN) untuk penerimaan mahasiswa baru di

Universitas Pakuan. *Chatbot* dibangun menggunakan Python dan Telegram API. Data yang digunakan untuk pengujian sebanyak 251 pertanyaan. Hasil yang didapat dari pengujian *chatbot* adalah akurasi yang didapat sebesar 88% , presisi 95%, dan *recall* 92%.

Elcholiqi & Musdholifah (2020), menjelaskan FAQ dalam website perusahaan berguna untuk menginformasikan layanan dan produknya. Tetapi FAQ dianggap kurang praktis dan interaktif. Terdapat alternatif salah satu nya yaitu dengan menggunakan chatbot. Chatbot digunakan untuk BTPN dalam menyediakan informasi untuk salah satu produk nya yaitu Jenius. Chatbot yang digunakan dalam Jenius memanfaatkan NLP agar sistem dapat memahami query dari pengguna dalam bentuk natural. Untuk mencari kemiripan query digunakan cosine similarity karena dianggap paling mirip dengan query pengguna. Tetapi dengan metode cosine similarity tidak memperhatikan struktur kalimat sehingga ditambahkan parse tree untuk pengecekan struktur kalimat dan memberi bobot pola. Hasil yang didapatkan dari 10 penguji menghasilkan tingkat kesesuaian jawaban dengan masukkan pengguna sebesar 84%.

Cahya Utama & Titi Komalasari (2021), menggunakan algoritma *Naïve Bayes Classifier* untuk meminimalisir waktu menjawab pernyataan pegawai baru dan diimplementasikan menggunakan *chatbot*. Dataset yang digunakan adalah *Frequently Asked Questions* dengan 10 pertanyaan dan jawaban. Diperoleh nilai akurasi menggunakan algoritma *Naïve Bayes Classifier* 93,33% dan nilai kesalahan 6,66%.

2.2 Chatbot

Chatbot adalah program manusia yang mensimulasikan dialog cerdas antara satu atau lebih pengguna menggunakan teks. *Chatbot* dapat digunakan pada percakapan yang berskala kecil atau dapat juga pada percakapan yang berskala sedang tergantung pada pertanyaan nya. *Chatbot* adalah salah contoh dari *Artificial Intelligence* (AI).

Sebagian besar *website* saat ini mengandalkan navigasi berbasis menu dan pencarian untuk memberikan informasi kepada pengguna. Namun *website* dengan jumlah konten yang banyak dapat mempersulit pengguna untuk menemukan informasi dengan cepat. Dalam kasus ini *chatbot* dapat memudahkan pengguna untuk mencari informasi. Pengguna memiliki opsi untuk mengobrol dengan *bot* dan mengajukan pertanyaan untuk mendapatkan jawaban. *Chatbot* memiliki respon yang telah diprogram sebelumnya dan dapat bekerja dengan baik untuk membuat percakapan yang relevan (S. Gupta et al., 2015).

Pelayanan yang diberikan oleh suatu perusahaan dengan media *website* tidak berjalan 24 jam. Hal ini membuat pelayanan kurang efektif karena terdapat keterbatasan jam kerja yang membuat pertanyaan-pertanyaan diluar jam kerja tidak mendapatkan respon. Untuk itu dibutuhkan *chatbot* yang menggantikan peran dari *customer service*. Penelitian oleh Hormansyah & Utama (2018), mereka mengimplementasikan *chatbot* menggunakan metode *TF-IDF* dan *Cosine Similarity* pada *website* layanan kesehatan Kota Malang. Pengimplementasian *chatbot* karena dapat memudahkan dalam mencari data

karena saat ini data harus dicari satu persatu. Dengan adanya *chatbot* pengunjung dapat melakukan tanya jawab dan mendapatkan informasi tentang pelayanan kesehatan di Kota Malang.

Perkembangan pariwisata semakin pesat setiap tahunnya seiring dengan banyak nya tempat wisata baru yang dibuka. Salah satu kota yang terkenal dengan perkembangan tempat wisata adalah kota Yogyakarta. Yogyakarta memiliki tempat wisata yang beragam dan membutuhkan informasi yang dapat diakses oleh wisatawan lokal dan mancanegara. Salisah et al., (2020), membuat *chatbot* untuk memudahkan wisatawan mencari informasi wisata yang ada di Yogyakarta. Algoritma yang digunakan untuk membangun chatbot menggunakan *Boyer-moore*. Algoritma *Boyer-moore* bekerja dengan cara membandingkan pola atau kata dari kanan ke kiri, sehingga algoritma tersebut menjadi algoritma pencarian yang baik. Hasil pengujian pada chatbot dapat dikatakan bekerja dengan baik.

2.3 Natural Language Processing

Natural Language Processing (NLP) adalah bidang kecerdasan buatan yang digunakan untuk menganalisis dan meniru bahasa manusia secara otomatis (Cambria & White, 2014). EYD atau Ejaan Yang Disempurnakan adalah aspek penting dalam pembuatan teks *essay*, jurnal, dan lainnya. Wangsanegara & Subaeki (2015) mereka melakukan penelitian dengan mengidentifikasi kesalahan dalam penulisan huruf kapital atau tanda baca dengan menggunakan implementasi *Natural Language Processing* (NLP) dan metode *Fuzzy Logic Tsukamoto*. Jumlah kata yang digunakan sebanyak 31.759

kata dan dibangun menggunakan MySQL dan PHP. Hasil dari pengujian dilakukan terhadap 20 abstrak skripsi dengan hasil kesesuaian identifikasi sebesar 70%.

Kota Bandung merupakan kota wisata dengan banyak tempat wisata yang ada. Informasi tentang tempat wisata pada kota Bandung dapat diakses pada *website* pariwisata kota Bandung. S. C. P & Afrianto (2015) melakukan wawancara kepada pengunjung *website* pariwisata kota Bandung dan mendapatkan hasil wawancara pengunjung harus mencari satu per satu tempat wisata pada *website* pariwisata kota Bandung. Dari permasalahan tersebut, mereka membuat sebuah aplikasi *chatbot* yang interaktif. Aplikasi *chatbot* menggunakan bahasa manusia dengan menggunakan metode *Natural Language Processing* (NLP) dan *Unified Model Language* (UML). Hasil pengujian yang didapat adalah wisatawan setuju bahwa aplikasi chatbot dapat membantu wisatawan mencari informasi tempat wisata di Kota Bandung.

Opini publik banyak sekali ditemukan pada media sosial, salah satunya pada twitter. Dalam penelitian yang dilakukan oleh Munasatya & Novianto (2020), mereka menggunakan *Natural Language Processing* (NLP) agar data dapat dipahami oleh komputer dan dimasukkan ke dalam model klasifikasi menggunakan model *Multi-Layer Perceptron* yang berguna untuk mendapatkan hasil prediksi akurasi. Objek penelitian ini menggunakan aplikasi *twitter* untuk Analisis Sentimen opini publik. Untuk mendapatkan visualisasi data menggunakan *Supervised Learning* dengan hasil pengujian yang didapat untuk nilai akurasi sebesar 93,26%.

2.4 *Neural Network* (NN)

Salah satu sistem penjurusan yang ada pada program studi adalah peminatan. Peminatan berfungsi untuk lebih fokus memilih satu bidang yang disukai sebagai penentu topik dari skripsi. Dalam penelitian oleh Riani et al., (2021), mereka melakukan penelitian terhadap program studi Sistem Informasi Universitas Telkom tentang penjurusan mahasiswa. Program studi Sistem Informasi menyediakan lima bidang peminatan yang nantinya dapat dipilih oleh mahasiswa semester tujuh. Mahasiswa masih kesulitan dalam memilih bidang peminatan yang dipilih, dan lebih menyukai bidang yang dianggap mudah. Oleh karena itu dilakukan penelitian untuk proses seleksi pemilihan peminatan mahasiswa program studi Sistem Informasi dengan menggunakan *Neural Network*. *Neural Network* digunakan karena mampu untuk mengklasifikasi data yang sangat kompleks. Hasil yang didapat adalah nilai AUC sebesar 0.9950 dan nilai akurasi sebesar 0.9863.

Indonesia tanaman tomat banyak dijumpai dan dibudidayakan oleh masyarakat karena iklim yang mendukung dan harga yang tomat yang stabil. Tetapi, terdapat beberapa kendala bagi para petani dalam budidaya tomat karena rentan diserang penyakit yang mengakibatkan petani kesulitan dalam mendeteksi penyakit tanaman tomat. Dalam penelitian yang dilakukan oleh Putri (2021), menggunakan metode *Neural Network* dan algoritma *backpropagation* dalam membantu petani untuk mengidentifikasi penyakit pada daun tanaman tomat. Hasil yang didapat untuk akurasi sebesar 78% , presisi sebesar 0,78 dan data klasifikasi bernilai benar sebanyak 117 data.

Indonesia adalah negara hukum dengan banyak hukum dan peraturan. Dalam penelitian Nia Agustina Purwitasari & Soleh (2022), Mereka membuat sistem untuk memudahkan masyarakat mendapatkan informasi tentang hukum Indonesia saat ini dengan menerapkan chatbot yang disebut "Peraturanku" untuk menghindari kebingungan tentang hukum dan peraturan. Mereka menggunakan *Neural Network* dengan pendekatan *Natural Language Processing* (NLP). Hasil uji coba dengan data sebanyak 35 percakapan dan berhasil menjawab percakapan dengan benar sebanyak 33 percakapan dengan nilai akurasi 94,28% dan kesalahan 5,71%.

BAB III

METODE PENELITIAN

3.1 Pengumpulan Data

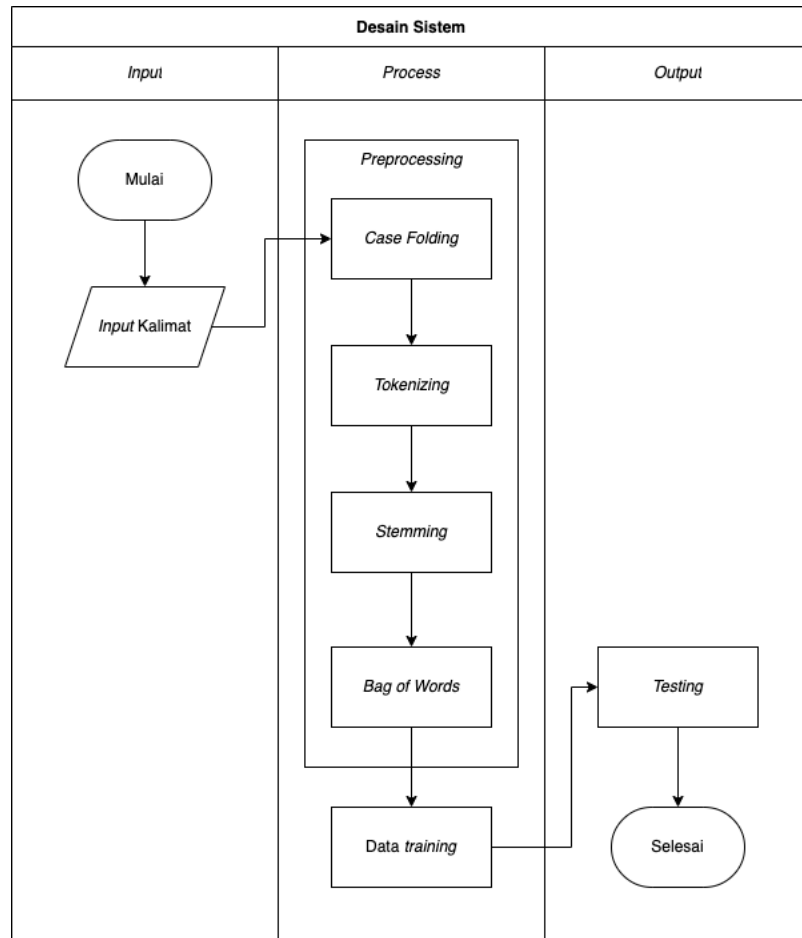
Pada penelitian ini pengumpulan data menggunakan referensi dari google dengan pertanyaan yang sering ditanyakan seputar PMB UIN Malang. Dengan mencari jawaban yang relevan pada *website* PMB UIN Malang. Terdapat 51 *tags*, 119 *patterns* serta 57 *response* yang digunakan pada dataset. Berikut merupakan contoh struktur *training set* atau *input* yang digunakan untuk membangun *chatbot* menggunakan metode *Neural Network*.

```
1  {"intents": [
2    {"tag": "error",
3      "patterns": [""],
4      "responses": ["Maaf, Kami Tidak Mengerti "]}
5  ],
6  {"tag": "salam",
7    "patterns": ["Halo", "Hai", "Hallo", "Salam", "Hola", "Permisi", "Mau tanya min", "Min"],
8    "responses": ["Hai, ada yang bisa dibantu?", "Halo! Apa yang bisa mimin bantu?"],
9    "context": [""]}
10 ],
11 {"tag": "salam2",
12   "patterns": ["Assalamualaikum", "Assalamu'alaikum", "Assalamu'alaikum Warahmatullahi", "Assalamu'alai
13   "responses": ["Waalaikummussalam"],
14   "context": [""]}
15 },
16 {"tag": "salam3",
17   "patterns": ["Selamat pagi", "Pagi", "Good morning", "Morning"],
18   "responses": ["Selamat pagi", "Good Morning"],
19   "context": [""]}
20 },
21 {"tag": "salam4",
22   "patterns": ["Selamat siang", "Siang", "Good afternoon"],
23   "responses": ["Selamat siang", "Good afternoon"],
24   "context": [""]}
25 },
26 {"tag": "salam5",
27   "patterns": ["Selamat malam", "Malam", "Good night"],
28   "responses": ["Selamat malam", "Good night"],
29   "context": [""]}
30 },
31 {"tag": "terimakasih",
32   "patterns": ["Terima kasih", "Makasih", "trims", "oke", "Sangat membantu"],
33   "responses": ["Senang bisa membantu", "Tentu, tidak masalah", "Terimakasih kembali"]
34 },
```

Gambar 3. 1 Struktur Training Set

3.2 Desain Sistem

Dalam sebuah penelitian dibutuhkan desain sistem agar tahapan penelitian dapat dilakukan secara terstruktur. Desain sistem dapat dilihat pada gambar 3.2.



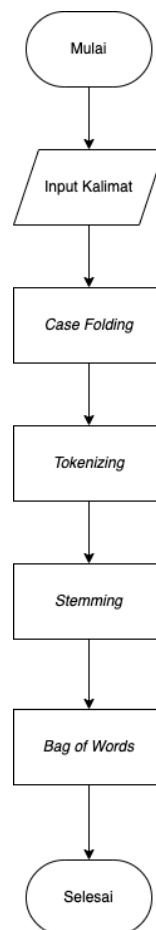
Gambar 3. 2 Desain Sistem

Pada tahap pertama terdapat *input* diberikan oleh pengguna berupa pertanyaan, kemudian kalimat yang telah dimasukkan oleh pengguna akan memasuki tahap *process* yang terdapat tahapan *preprocessing* dan *data training*. Pada tahap *preprocessing*, menggunakan pendekatan dari *Natural Language Processing* (NLP) yang terdiri dari *case folding*, *tokenizing*,

stemming, bag of words (Nia Agustina Purwitasari & Soleh, 2022). Setelah proses *preprocessing* selesai, proses data *training* yaitu proses melatih data dari hasil *preprocessing* menggunakan *neural network*. Proses terakhir yaitu *Testing* dilakukan untuk mendapatkan nilai akurasi dan *error* pada *chatbot*.

3.3 Preprocessing

Pada tahap *preprocessing* menggunakan pendekatan *Natural Language Processing*. Tahapan dalam *preprocessing* sebagai berikut:

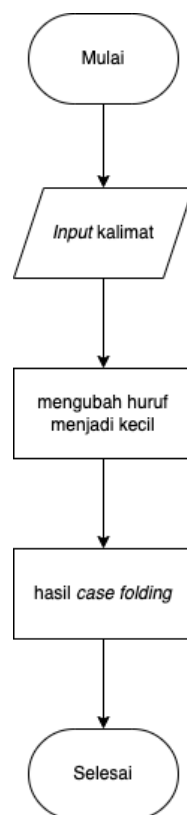


Gambar 3. 3 Flowchart Preprocessing

Tahapan dari proses *preprocessing* yang digunakan pada penelitian ini adalah sebagai berikut:

1. *Case Folding*

Case Folding yaitu cara mengubah semua huruf yang terdapat dalam dokumen menjadi huruf kecil. Huruf yang digunakan adalah huruf 'a' sampai huruf 'z' dan karakter selain huruf tersebut dianggap sebagai pembatas. Pada proses *case folding* menggunakan fungsi `.lower()` untuk mengubah huruf menjadi kecil semua.



Gambar 3. 4 Alur *Case Folding*

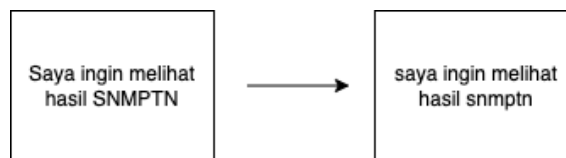
Keterangan :

1. Mulai
2. Pengguna memasukkan kalimat yang ingin ditanyakan pada program
3. Setelah kalimat berhasil dimasukkan, program akan menjalankan fungsi `.lower()` yang akan mengubah huruf besar menjadi kecil.

4. *Output* proses ini adalah mengubah huruf besar menjadi kecil semua.

5. Selesai

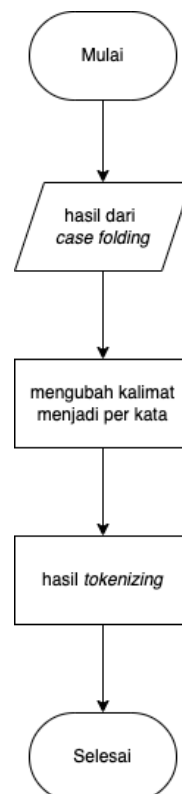
Berikut merupakan contoh dari proses *case folding* yang digunakan pada penelitian ini.



Gambar 3. 5 Contoh *Case Folding*

2. *Tokenizing*

Tokenizing merupakan cara memisahkan kata masukan berupa pertanyaan menjadi perkata (Mulyono & Sumijan, 2021). Untuk memisahkan pertanyaan menjadi per kata menggunakan fungsi *.split()*.

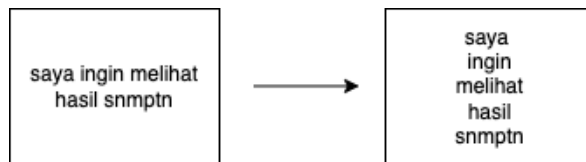


Gambar 3. 6 Alur Proses *Tokenizing*

Keterangan :

1. Mulai
2. Hasil *output* proses *case folding* digunakan untuk proses *tokenizing* sebagai *input*.
3. Setelah *output* proses *case folding* digunakan sebagai *input*, selanjutnya kalimat akan diubah menjadi kata dengan menggunakan fungsi *.split()*.
4. *Output* dari proses ini adalah mengubah kalimat menjadi kata.
5. Selesai

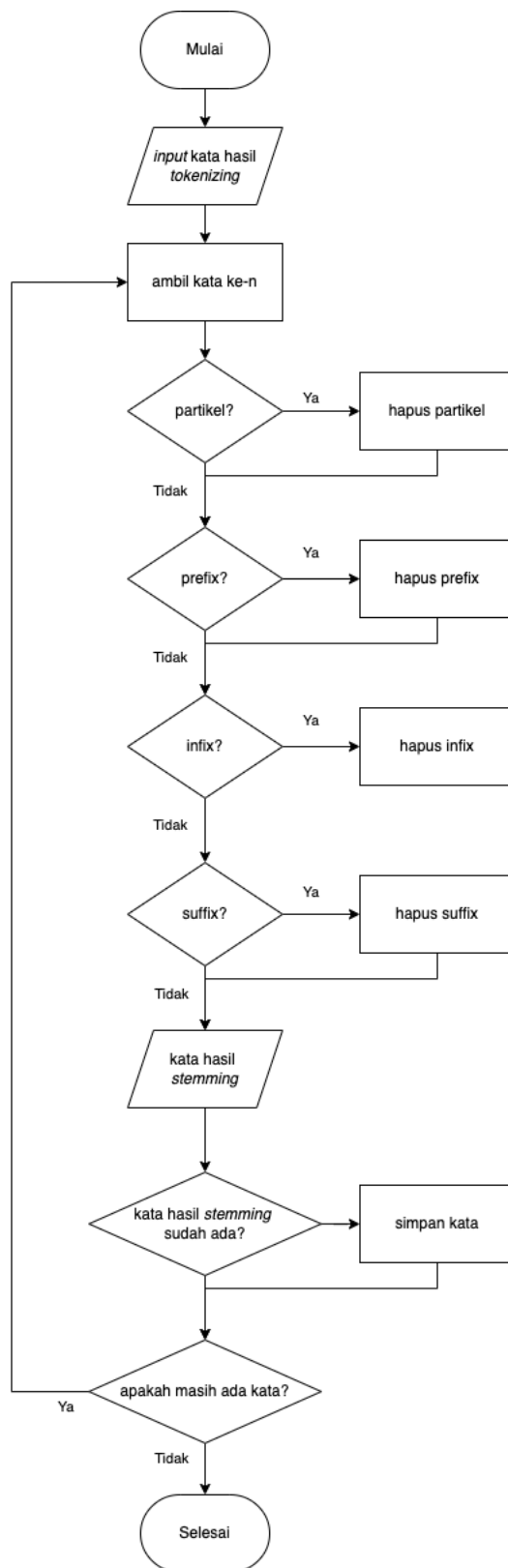
Setelah proses diatas selesai, berikut merupakan contoh dari proses *tokenizing* pada penelitian ini.



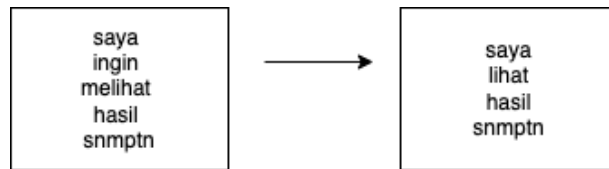
Gambar 3. 7 Contoh *Tokenizing*

3. *Stemming*

Stemming yaitu menghilangkan imbuhan pada kata sehingga menghasilkan hanya kata dasar (Fathoni et al., 2020). Terdapat 4 jenis imbuhan dalam Bahasa Indonesia seperti prefiks, infiks, sufiks, dan konfiks. Dalam satu kata dasar terdapat lebih dari satu imbuhan maka diperlukan urutan untuk menghilangkan kata imbuhan yang benar, jika salah dalam melakukan menghilangkan kata dasar maka kata dasar tersebut tidak dapat ditemukan. *Stemming* dapat digunakan untuk menemukan kata dasar dan menghasilkan struktur morfologi Bahasa Indonesia yang tepat.

Gambar 3. 8 Alur Proses *Stemming*

Proses diatas adalah untuk menghilangkan imbuhan yang terdapat pada kata dan akan dicari akar kata nya. Jika ada 2 kata dengan akar kata yang sama, 1 kata yang akan disimpan (Rezkanintio, 2021).



Gambar 3. 9 Contoh *Stemming*

4. *Bag of Words*

Bag of Words atau kantung kata adalah proses menghitung kemunculan setiap kata. Cara kerja dari *bag of words* yang pertama adalah data *training* disimpan ke dalam sistem dan akan diberikan nilai 1 jika kata tersebut terdapat pada kantung kata dan diberikan nilai 0 jika kata tersebut tidak terdapat pada kantung kata. Berikut alur dari proses *bag of words*.



Gambar 3. 10 Alur Proses *Bag of Words*

Keterangan :

1. Mulai program
2. *Output* proses *stemming* digunakan untuk proses *bag of words* sebagai *input*.
3. Setelah *output* dari proses *stemming* masuk sebagai *input*, selanjutnya setiap kata akan diberikan nilai 1 jika muncul pada kantung kata dan 0 jika kata tersebut tidak muncul. Kata tersebut akan diubah menjadi matriks berbentuk *array*. Data akan dibagi menjadi dua yaitu *x_train* dan *y_train*
4. Selesai

Setelah proses diatas selesai, berikut merupakan contoh dari pemberian nilai pada proses ini.

Tabel 3. 1 Proses *Bag of Words*

Kata	<i>Bag of Words Vector</i>					
	saya	lihat	covid	hasil	snmptn	mandiri
saya	1	0	0	0	0	0
lihat	0	1	0	0	0	0
hasil	0	0	0	1	0	0
snmptn	0	0	0	0	1	0

Setelah proses *preprocessing* selesai, langkah berikutnya adalah mengubah kata menjadi angka atau *biner*. Tujuan nya agar komputer dapat membaca pola atau kalimat yang akan digunakan pada saat pelatihan. Hasil dari proses *preprocessing* berupa *array* yang berisikan angka, angka tersebut adalah kata hasil *stemming* kata unik dan label yang selanjutnya akan di inisialisasikan sebagai *x_train* dan *y_train*.

setelah dikalikan hasil masukan tersebut dijumlahkan. Setiap koneksi pada *neuron* memiliki *weight* dengan nilai masing-masing akan berubah sampai *output* yang diinginkan sesuai dengan model *Neural Network*. Selanjutnya, hasil penjumlahan ditambahkan ke dalam bias *b*. bias ini digunakan agar perhitungan lebih akurat. Setelah proses penjumlahan selesai, dilanjutkan proses memasukkan *neuron* ke sebuah fungsi yaitu *activation function* yang nantinya akan dijadikan sebagai *output* dari *neuron*. *Activation function* digunakan untuk menentukan apakah suatu neuron aktif berdasarkan jumlah bobot dari *input*. Terdapat beberapa *activation function* yaitu *Sigmoid* dan *Softmax*.

1. *Sigmoid*

Sigmoid merupakan fungsi aktivasi yang menerima angka tunggal dan mengubah nilai *x* dengan rentang 0 hingga 1. Rumus dari *Sigmoid* :

$$S(x) = \frac{1}{1 + e^{-x}}. \quad (3.1)$$

2. *Softmax*

Softmax adalah fungsi aktivasi yang dapat digunakan untuk menentukan klasifikasi multi kelas perhitungan probabilitas. Dalam fungsi ini, diambil nilai probabilitas yang paling tinggi. Nilai probabilitas untuk *output* antara 0 hingga 1 (Husamuddin et al., 2020). Terdapat rumus dari *softmax* :

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K. \quad (3.2)$$

Dimana,

- z_j = Nilai *vector* ke - j
- k = Jumlah kelas

Neural Network dapat digunakan untuk aplikasi tertentu seperti pengenalan pola dan klasifikasi data selama proses pembelajaran. *Neural Network* memiliki kemampuan untuk mendapatkan makna dari data yang kompleks dan tidak akurat.. Terdapat beberapa keuntungan dengan menggunakan *Neural Network* antara lain (B. B. Gupta et al., 2012) :

1. *Adaptive learning*, adalah kemampuan untuk mempelajari cara melakukan tugas berdasarkan data yang disediakan untuk pelatihan.
2. *Self-organization*, adalah Jaringan Syaraf Tiruan (JST) yang membentuk kelompoknya sendiri dari informasi yang diterima selama periode penelitian.
3. *Real time operation*, yaitu perhitungan JST dapat dilakukan secara paralel dan perangkat keras dirancang dan dibuat untuk memanfaatkan fitur-fitur yang ada.
4. *Fault tolerance via redundant information coding*, jika bagian dari jaringan JST rusak, kinerjanya akan menurun. Namun, beberapa kapasitas jaringan dapat dipertahankan jika terjadi kerusakan jaringan yang parah.

Neural Network memiliki beberapa arsitektur model yaitu *Single Layer Perceptron* dan *Multi Layer Perceptron*.

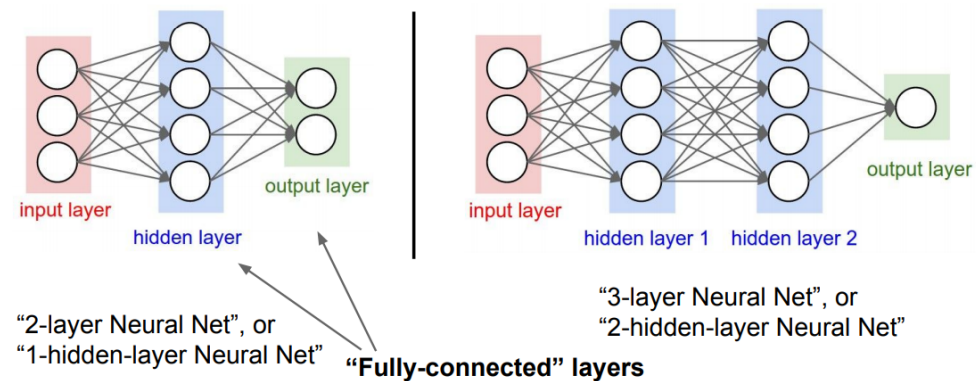
1. *Single Layer Perceptron*

Single Layer Perceptron atau Jaringan Layer Tunggal merupakan arsitektur *neural network* yang paling sederhana terdiri dari 1 *input layer*, 1 lapisan *bobot*, dan 1 *output layer*. Cara kerja dari *single layer* adalah *input layer* setelah mendapatkan masukan langsung menuju ke *output layer*.

2. *Multi Layer Perceptron*

Multi Layer Perceptron atau jaringan Layer Jamak merupakan arsitektur yang memiliki tiga jenis lapisan yaitu *input layer*, *hidden layer*, dan *output layer*. Arsitektur ini dapat memecahkan masalah yang lebih kompleks daripada *single layer perceptron*. Namun, proses *training* membutuhkan waktu yang cukup lama. Cara kerja dari *multi layer perceptron* adalah *input layer* mendapatkan masukan akan menuju pada *layer* kedua yaitu *hidden layer* dan dilakukan komputasi. Selanjutnya *output* dari *hidden layer* kedua digunakan sebagai *input* untuk *hidden layer* ketiga dan seterusnya.

Neural Network terdiri dari *neuron* atau unit yang berfungsi untuk memproses informasi yang saling terkoneksi dan tersusun berlapis-lapis.



Gambar 3. 13 Arsitektur *Neural Network*

Gambar diatas menunjukkan *Neural Network* memiliki sekumpulan *neuron* yang terorganisir dalam lapisan-lapisan atau *layers*. Gambar diatas dapat disebut *Multi Layer Perceptro* atau *Fully Connected Layer* dengan gambar pertama memiliki 3 layer yaitu *input layer* mempunyai 3 buah *neuron*, *hidden layer* mempunyai 4 buah *neuron*, dan *output layer* mempunyai 2 buah *node output*. Berikut penjelasan dari tiap *layers* :

1. *Input Layer* merupakan lapisan yang membawa data masuk ke dalam sistem kemudian dilakukan proses pada *layer* selanjutnya. Jumlah *input layer* menyesuaikan dengan jumlah masukan.
2. *Hidden Layer* merupakan *layer* yang terdapat diantara *input layer* dan *output layer*. *Hidden layer* menerima data dari *input layer* berupa data-data dengan pembobot atau *weight* yang berguna untuk menghasilkan *output neuron* menggunakan *activation function* pada proses selanjutnya.
3. *Output Layer* merupakan lapisan terakhir dari *Neural Network*. Pada *output layer* menghasilkan *output* dari sistem.

4. *Weight* dan *Bias* merupakan *neuron* yang saling berhubungan yang ditandai dengan tanda panah pada setiap *layer* nya. *Hidden layer* dan *output layer* memiliki tambahan *input* yang disebut dengan *bias* namun pada gambar diatas tidak disebutkan. Pada gambar pertama dapat dihitung nilai *weight* dan *bias* mendapatkan hasil 26 parameter pada proses *training* dengan perhitungan $(3 \times 4 \text{ weight}) + 4 \text{ bias}$ dan $(4 \times 2 \text{ weight}) + 2 \text{ bias}$. Pada gambar kedua terdapat $(3 \times 4 \text{ weight}) + 4 \text{ bias}$ dan $(4 \times 4 \text{ weight}) + 4 \text{ bias}$ dan $(4 \times 1 \text{ weight}) + 1 \text{ bias}$ dengan total 41 parameter.

Terdapat beberapa parameter untuk pengembangan *chatbot* dengan algoritma *neural network* yang berguna sebagai parameter data, yaitu :

1. *Epoch*

Epoch merupakan satu siklus perulangan proses dari *neural network* yang berarti satu *epoch* telah melakukan proses belajar dari pelatihan. Pelatihan dilakukan secara berulang agar mendapatkan nilai akurasi dan *error* dengan baik.

2. *Batch Size*

Batch size adalah banyak nya data yang akan dilakukan untuk proses pelatihan.

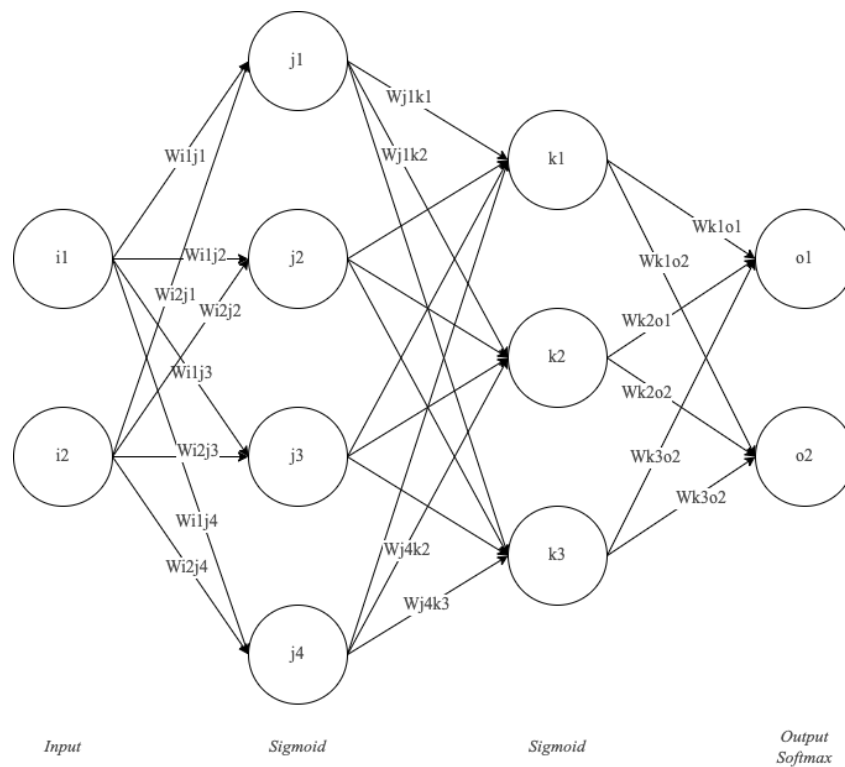
3. *Iteration*

Iteration merupakan banyaknya *batch* untuk menyelesaikan satu *epoch* yang berarti satu *epoch* telah melewati proses *feed forward* dan *backward*.

4. *Loss*

Loss merupakan nilai yang digunakan dalam perhitungan untuk membuat model semakin baik.

Pada penelitian ini menggunakan jenis klasifikasi *feedforward* dan *backpropagation* untuk koreksi nilai *output* dengan menggunakan 4 *layer* pada proses *training*. *Backpropagation* adalah algoritma pembelajaran yang berfungsi untuk meminimalisir *error* pada *output*. Terdapat 4 *layer* yang digunakan yaitu *layer* pertama merupakan *input layer* yang berfungsi untuk menerima data yang masuk ke dalam sistem. Data berasal dari tahapan *preprocessing* yang telah dilakukan sebelumnya. *Layer* kedua dan ketiga adalah *hidden layer* dengan menggunakan fungsi aktivasi *Sigmoid*. *Layer* keempat adalah *output* dengan menggunakan fungsi aktivasi *Softmax*. Penelitian sebelumnya telah berhasil menggunakan *neural network* sebagai implementasi dari *chatbot*. Dataset yang digunakan menggunakan Frequently Asked Questions (FAQ) yang terdapat pada website Universitas Nasional, dengan 25 pertanyaan dan jawaban yang dikelompokkan menjadi 16 kelas. Pengujian dilakukan dalam 110 percakapan, menjawab 107 percakapan dengan benar dan gagal menjawab 3 percakapan. Akurasi yang didapat sebesar 97,27% dan *error* sebesar 2,72% (Mustakim et al., 2021). Berikut contoh perhitungan menggunakan *forward pass*.



Gambar 3. 14 Model *Forward Pass*

Dimana,

- i = masukkan atau *input*
- w = *weight* atau bobot
- j = *node layer* pertama
- k = *node layer* kedua
- *bias* = nilai yang digunakan dalam *input* dan *output* untuk penyesuaian nilai *weight*

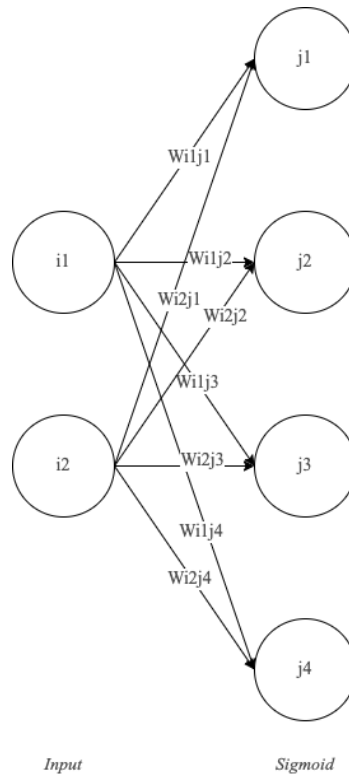
Untuk membuat model *neural network* dibutuhkan jumlah *layer* dan *activation function*. Berikut arsitektur *neural network* yang digunakan pada penelitian ini.

Tabel 3. 2 Arsitektur *Neural Network*

<i>Layer</i>	<i>Hidden Layer</i>	<i>Jumlah node hidden layer</i>	<i>Activation function hidden layer</i>	<i>Activation function output layer</i>
4	2	128	<i>Sigmoid</i>	<i>Softmax</i>

Pada penelitian ini menggunakan 4 *layer* yaitu *input layer*, 2 *hidden layer*, dan *output layer*. Pada 2 *hidden layer* menggunakan *activation function sigmoid* dengan jumlah *node* sebanyak 128. Pada *output layer* menggunakan *activation function softmax*. *Activation function Sigmoid* merupakan fungsi aktivasi yang menerima angka tunggal dan mengubah nilai x dengan rentang 0 hingga 1. *Activation function softmax* adalah fungsi yang dapat digunakan dalam menentukan klasifikasi multi kelas untuk perhitungan probabilitas. Dalam fungsi ini, diambil nilai probabilitas yang paling tinggi. Nilai probabilitas untuk *output* antara 0 hingga 1.

Nilai bobot digunakan nilai acak yang bernilai kecil dan untuk nilai bias menggunakan nilai 1. Setelah menentukan *weight* dan *bias*, maka dilanjutkan proses *forward pass* untuk perhitungan *input layer* dan *hidden layer* pertama.



Gambar 3. 15 *Forward Pass* Input Layer Menuju *Hidden Layer 1*

Proses selanjutnya adalah perhitungan *forward pass* dari *input layer* menuju *Hidden Layer* pertama dengan cara melakukan perkalian dan penjumlahan nilai matriks antara *input*, *weight* dan *bias*. Berikut perhitungan yang digunakan.

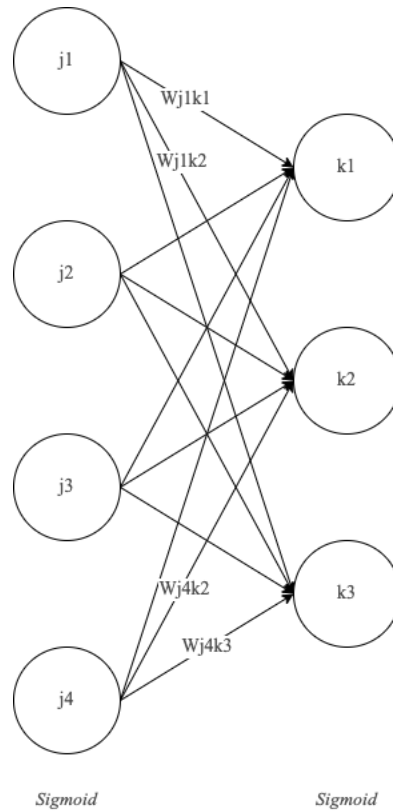
(3.3)

$$[j_{1in} \ j_{2in} \ j_{3in} \ j_{4in}] = [i_1 \ i_2] \times \begin{bmatrix} Wi1j1 & Wi1j2 & Wi1j3 & Wi1j4 \\ Wi2j1 & Wi2j2 & Wi2j3 & Wi2j4 \end{bmatrix} + [b_{1j1} \ b_{1j2} \ b_{1j3} \ b_{1j4}]$$

Setelah mendapatkan nilai *input* dari tiap *node*, nilai tersebut nantinya akan dihitung dan dikeluarkan menggunakan *activation function*. *Activation function* pada *hidden layer* pertama menggunakan *sigmoid*. Berikut perhitungan dari *activation function* menggunakan *sigmoid*.

$$Sigmoid ([j_{1in} \ j_{2in} \ j_{3in} \ j_{4in}]) = \left[\frac{1}{1+e^{-j_{1in}}} \quad \frac{1}{1+e^{-j_{2in}}} \quad \frac{1}{1+e^{-j_{3in}}} \quad \frac{1}{1+e^{-j_{4in}}} \right] \quad (3.4)$$

Selanjutnya, dilakukan *forward pass* pada *hidden layer* pertama menuju *hidden layer* kedua dengan nilai yang sudah didapatkan dari perhitungan *hidden layer* pertama menggunakan *activation function sigmoid*.



Gambar 3. 16 *Forward Pass Hidden Layer 1 Menuju Hidden Layer 2*

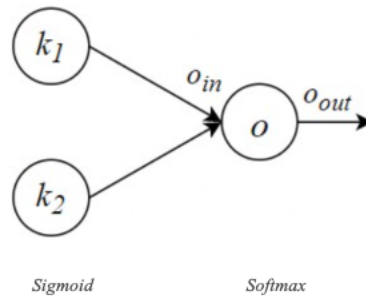
Berikut perhitungan dari *activation function* menggunakan *sigmoid* pada *hidden layer* dua.

$$[k_{1in} \ k_{2in} \ k_{3in}] = [j_{1out} \ j_{2out} \ j_{3out} \ j_{4out}] \times \begin{bmatrix} w_{j1k1} & w_{j1k2} & w_{j1k3} \\ w_{j2k1} & w_{j2k2} & w_{j2k3} \\ w_{j3k1} & w_{j3k2} & w_{j3k3} \\ w_{j4k1} & w_{j4k2} & w_{j4k3} \end{bmatrix} + [b_{jk1} \ b_{jk2} \ b_{jk3}] \quad (3.5)$$

Hasil perhitungan diatas akan menghasilkan *input* untuk selanjutnya diproses menggunakan *activation function sigmoid* pada *hidden layer* kedua.

$$\text{Sigmoid}([k_{1in} \ k_{2in} \ k_{3in}]) = \left[\frac{1}{1+e^{-k_{1in}}} \quad \frac{1}{1+e^{-k_{2in}}} \quad \frac{1}{1+e^{-k_{3in}}} \right] \quad (3.6)$$

Perhitungan selanjutnya adalah menghitung *hidden layer* dua dan *output*. Hasil dari *sigmoid* yang ada pada *hidden layer* dua diteruskan ke *output layer*.



Gambar 3. 17 Forward Pass Hidden Layer 2 Menuju Output Layer

Berikut perhitungan yang digunakan.

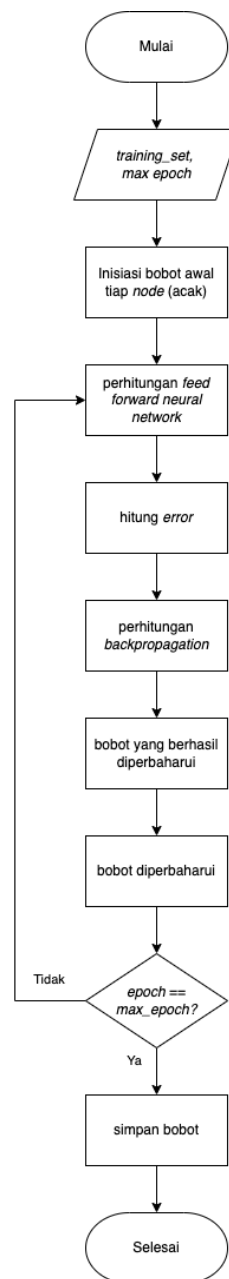
$$[o_{in}] = [k_{1out} \ k_{2out} \ k_{3out}] \times \begin{bmatrix} Wk1o1 & Wk1o2 \\ Wk2o1 & Wk2o2 \\ Wk3o1 & Wk3o3 \end{bmatrix} \quad (3.7)$$

Hasil perhitungan pada *output layer* menggunakan *activation function softmax*. Berikut perhitungan yang digunakan.

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K. \quad (3.8)$$

3.5 Pelatihan Neural Network

Nilai bobot atau *weight* perlu dilakukan optimasi agar dapat mengurangi nilai *error* pada *epoch*. Cara yang dapat dilakukan adalah dengan melakukan perulangan untuk memperkecil nilai *error*.



Gambar 3. 18 Flowchart Pelatihan Neural Network

Algoritma yang dapat digunakan untuk mengoptimalkan bobot dari node dan juga dapat untuk mengurangi *error* adalah algoritma *backpropagation*. Pada *output layer*, nilai prediksi diperoleh secara normal. Langkah selanjutnya adalah menggunakan *cross-entropy loss* untuk mencari loss. *Cross entropy loss* adalah fungsi yang digunakan untuk menghitung *error*

dari suatu model. Fungsi ini biasanya digunakan setelah fungsi *softmax* untuk menghitung *error*. Berikut rumus dari *Cross entropy loss*.

$$H(p, q) = - \sum_{i=1}^N p_i \log q_i \quad (3.8)$$

Dimana,

- $p = Tag$
- $q = Output$ dari *softmax*

Terdapat beberapa metode optimasi, salah satu nya yaitu ADAM (*Adaptive Moment Estimation*). ADAM dapat memperbarui bobot yang didasarkan pada data *training*. ADAM merupakan kombinasi dari RMSprop dan Stochastic Gradient Descent dengan momentum. ADAM bekerja dengan mempertahankan *learning rate* yang berfungsi untuk meningkatkan kinerja model.

BAB IV

UJI COBA DAN PEMBAHASAN

4.1 Skenario Uji Coba

Pada tahap ini dilakukan pengujian sistem yang sudah dibangun dengan cara menghitung nilai akurasi pada *chatbot*. Pada penelitian ini terdapat 51 *tags*, 119 *patterns* serta 57 *response* yang digunakan pada *dataset*. Pada data pengujian digunakan 25 *patterns* dari 20% total *patterns* yang ada pada *dataset* (Brownlee, 2020). Data yang digunakan untuk pengujian merupakan kalimat acak yang berdasarkan pada tiap *tags*.

4.2 Tahap Pelatihan

Tahap pelatihan membahas penerapan *Neural Network* dengan pendekatan *Natural Language Processing* pada *chatbot* PMB UIN Malang seperti yang sudah disebutkan pada Bab III. Sebuah model *chatbot* membutuhkan *training* atau pelatihan agar dapat mendapatkan hasil yang sesuai atau akurat. Pada tahap ini ditentukan jumlah *epoch*, jumlah data, dan *learning rate*.

Tabel 4. 1 Inisiasi Nilai Model

<i>Epoch</i>	Jumlah Data	<i>Learning rate</i>
500	8	0.001

Setelah menginisiasi nilai diatas, dilakukan proses pelatihan menggunakan *neural network* dengan proses *feed forward*. Proses *input* ke *hidden layer* pertama dapat dilihat pada gambar 3.15. *Input* pada proses ini adalah hasil dari konversi kata menjadi *biner* yang kemudian akan dilakukan

perhitungan menggunakan *sigmoid*. *Sigmoid* digunakan untuk memperkenalkan non-linearitas dan bernilai 0 – 1 pada *neural network*. Berikut perhitungan nilai *input* ke *hidden layer* pertama.

$$\begin{aligned}
 [j_{1in} \ j_{2in} \ j_{3in} \ j_{4in}] &= [i_1 \ i_2] \times \begin{bmatrix} W_{i1j1} & W_{i1j2} & W_{i1j3} & W_{i1j4} \\ W_{i2j1} & W_{i2j2} & W_{i2j3} & W_{i2j4} \end{bmatrix} + [b_{i1j1} \ b_{i1j2} \ b_{i1j3} \ b_{i1j4}] \\
 &= [1.0 \ 0.5] \times \begin{bmatrix} 0 & 1.0 & 0.25 & 1.0 \\ 1.0 & 0.5 & 0.5 & 0.5 \end{bmatrix} + [1.0 \ 1.0 \ 1.0 \ 1.0] \\
 &= [0.5 \ 1.25 \ 0.5 \ 1.25] + [1.0 \ 1.0 \ 1.0 \ 1.0] \\
 &= [1.5 \ 2.25 \ 1.5 \ 2.25]
 \end{aligned}$$

Output diatas digunakan untuk menjadi *input* tiap *node* pada *hidden layer* pertama. *Activation function* yang digunakan pada *hidden layer* pertama menggunakan *sigmoid*. Perhitungan *hidden layer* pertama sebagai berikut.

$$\begin{aligned}
 Sigmoid([j_{1in} \ j_{2in} \ j_{3in} \ j_{4in}]) &= \left[\frac{1}{1+e^{-j_{1in}}} \quad \frac{1}{1+e^{-j_{2in}}} \quad \frac{1}{1+e^{-j_{3in}}} \quad \frac{1}{1+e^{-j_{4in}}} \right] \\
 &= \frac{1}{1+e^{-1.5}} \quad \frac{1}{1+e^{-2.25}} \quad \frac{1}{1+e^{-1.5}} \quad \frac{1}{1+e^{-2.25}} \\
 &= 0.8175 \quad 0.9046 \quad 0.8175 \quad 0.9046
 \end{aligned}$$

Setelah menghitung nilai dari *activation sigmoid* pada *hidden layer* pertama, langkah berikutnya adalah proses *forward pass* pada *hidden layer* pertama menuju *hidden layer* kedua dapat dilihat pada gambar 3.16. Berikut perhitungan pada *hidden layer* kedua.

$$\begin{aligned}
 [k_{1in} \ k_{2in} \ k_{3in}] &= [j_{1out} \ j_{2out} \ j_{3out} \ j_{4out}] \times \begin{bmatrix} W_{j1k1} & W_{j1k2} & W_{j1k3} \\ W_{j2k1} & W_{j2k2} & W_{j2k3} \\ W_{j3k1} & W_{j3k2} & W_{j3k3} \\ W_{j4k1} & W_{j4k2} & W_{j4k3} \end{bmatrix} + [b_{jk1} \ b_{jk2} \ b_{jk3}] \\
 &= [0.8175 \ 0.9046 \ 0.8175 \ 0.9046] \times \begin{bmatrix} 1.0 & 0.75 & 0.5 \\ 0 & 0.5 & 0 \\ 0.25 & 0 & 1.0 \\ 0.5 & 1.0 & 0.25 \end{bmatrix} + [1.0 \ 1.0 \ 1.0] \\
 &= [0.62 \ 1.98 \ 1.45] + [1.0 \ 1.0 \ 1.0] \\
 &= [1.62 \ 2.98 \ 2.45]
 \end{aligned}$$

Output diatas digunakan sebagai *input* tiap *node* pada *hidden layer* kedua. *Activation function* yang digunakan pada *hidden layer* ini adalah *sigmoid*. Perhitungan pada *hidden layer* kedua sebagai berikut.

$$\begin{aligned} \text{Sigmoid}([k_{1in} \quad k_{2in} \quad k_{3in}]) &= \left[\frac{1}{1+e^{-k_{1in}}} \quad \frac{1}{1+e^{-k_{2in}}} \quad \frac{1}{1+e^{-k_{3in}}} \right] \\ &= \left[\frac{1}{1+e^{-1.62}} \quad \frac{1}{1+e^{-2.98}} \quad \frac{1}{1+e^{-2.45}} \right] \\ &= [0.8347 \quad 0.9516 \quad 0.9205] \end{aligned}$$

Setelah menghitung nilai dari *activation sigmoid* pada *hidden layer* kedua, langkah selanjutnya adalah proses *forward pass* menuju *output* yang dapat dilihat pada gambar 3.17. Berikut perhitungan pada *output*.

$$\begin{aligned} [O_{in}] &= [k_{1out} \quad k_{2out} \quad k_{3out}] \times \begin{bmatrix} Wk1o1 & Wk1o2 \\ Wk2o1 & Wk2o2 \\ Wk3o1 & Wk3o3 \end{bmatrix} \\ &= [0.8347 \quad 0.9516 \quad 0.9205] \times \begin{bmatrix} 0.5 & 0.25 \\ 0 & 0.5 \\ 0.25 & 0 \end{bmatrix} \\ &= [0.6474 \quad 0.6844] \end{aligned}$$

Output diatas digunakan sebagai *input node* pada *output layer*. *Activation function* yang digunakan pada *output layer* adalah *softmax*.

$$\begin{aligned} \sigma(\mathbf{z})_j &= \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K. \\ &= \frac{e^{-0.6474}}{1.0277} \text{ Kelas 1} \\ &= \frac{0.5234}{1.0277} \text{ Kelas 1} \\ &= 0.5092 \text{ Kelas 1} \\ &= \frac{e^{-0.6844}}{1.0277} \text{ Kelas 2} \\ &= \frac{0.5043}{1.0277} \text{ Kelas 2} \\ &= 0.4907 \text{ Kelas 2} \end{aligned}$$

Selanjutnya, setelah mendapatkan nilai dari perhitungan *activation function softmax*, langkah berikutnya adalah menghitung *loss* dari *output softmax* menggunakan *cross entropy loss*. Hasil dari kelas 1 menunjukkan nilai *output* paling besar dan akan diasumsikan memiliki nilai 1, sedangkan kelas 2 menunjukkan *output* paling kecil dan akan diasumsikan memiliki nilai 0.

$$\begin{aligned}
 H(p, q) &= - \sum_{i=1}^N p_i \log q_i \\
 &= -(1 \times \log(0.5092) + 0 \times \log(0.4907)) \\
 &= 0.2931
 \end{aligned}$$

Potongan kode program pembuatan model menggunakan *neural network* dapat dilihat pada gambar 4.1.

```

class NeuralNetwork(nn.Module):
    def __init__(self, ukuran_input, ukuran_hidden, jumlah_kelas):
        super(NeuralNetwork, self).__init__()
        # menggunakan layer Linear / Fully-Connected Layer
        # mendefinisikan ada sebuah bobot yang memetakan layer dengan neuron sebanyak ukuran_input ke layer
        self.l1 = nn.Linear(ukuran_input, ukuran_hidden)
        self.l2 = nn.Linear(ukuran_hidden, ukuran_hidden)
        self.l3 = nn.Linear(ukuran_hidden, jumlah_kelas)
        self.sigmoid = nn.Sigmoid()

    def forward(self, x):
        out = self.l1(x)
        out = self.sigmoid(out)
        out = self.l2(out)
        out = self.sigmoid(out)
        out = self.l3(out)

        return out

```

Gambar 4. 1 Potongan Kode Program Model *Neural Network*

Pelatihan model *neural network* akan dihitung nilai akurasi dan nilai *error* pada proses pelatihan. Pada proses pelatihan model ini menggunakan *feed forward* dan *backpropagation* yang berguna untuk meminimalkan kesalahan dalam *output* yang dihasilkan. Berikut potongan kode program pada proses pelatihan.

```

for epoch in range(jumlah_epochs):
    train_acc, correct, train_loss, target_count = 0, 0, 0, 0
    for (kata, label) in trainset:
        kata = kata.type(torch.FloatTensor)
        label = label.to(dtype=torch.long)

        # Feed Forward
        outputs = model(kata)
        label = torch.max(label, 1)[1]
        loss = hitung_error(outputs, label)
        grafik_error.append(loss.item())
        # Untuk melakukan sebuah operasi back propagation, kita perlu memanggil fungsi .backward()
        # mereset gradien dengan memanggil optimizer.zero_grad() dan fungsi optimizer.step() akan mengupdate model.

        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

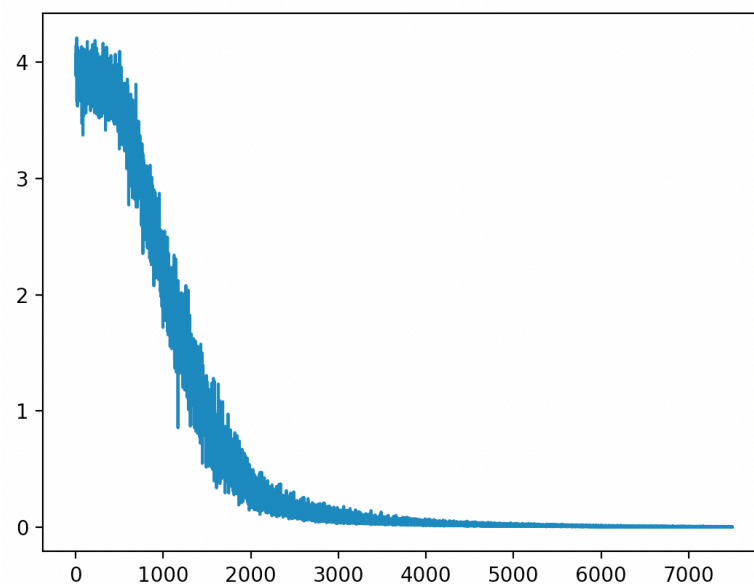
    # accuracy
    _, predicted = torch.max(outputs.data, 1)
    grafik_akurasi.append((predicted == label).sum().item() / predicted.size(0))

# 4f jumlah desimal loss
print(f'Epoch [{epoch + 1}/{jumlah_epochs}], Loss: {loss.item():.4f}, Accuracy: {np.mean(grafik_akurasi)*100:.2f}')

```

Gambar 4. 2 Potongan Kode Program Melatih *Neural Network*

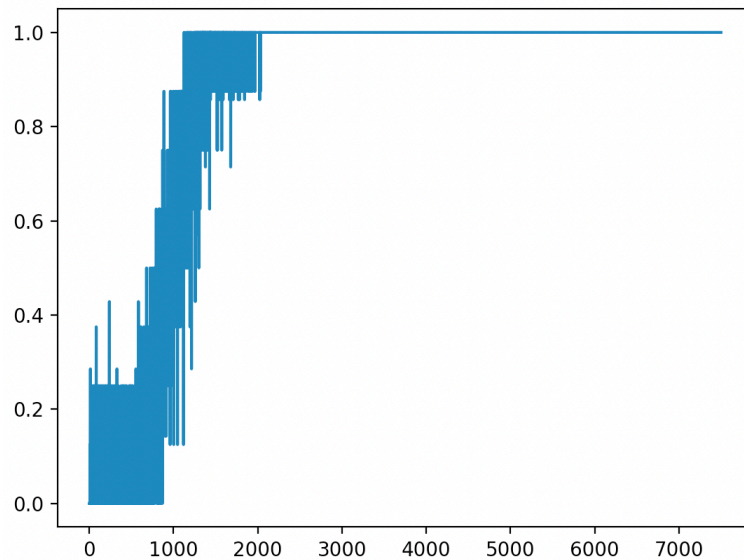
Setelah didapatkan nilai *error* menggunakan *cross entropy loss*, Nilai *error* yang didapatkan akan di *training* sampai 500 *epoch* hingga didapatkan nilai *error* paling kecil. Berikut grafik *error* yang didapatkan.



Gambar 4. 3 Grafik *Error*

Grafik *error* diatas menunjukkan penurunan *error* dari angka terbesar menuju *error* paling kecil sampai dengan *index error* 7000. Setelah mengetahui grafik *error* pada pelatihan *chatbot*, selanjutnya mencari grafik

nilai akurasi pada tahap pelatihan *chatbot*. Grafik akurasi pada dilihat pada gambar 4.4.



Gambar 4. 4 Grafik Akurasi

Nilai akurasi yang didapatkan *chatbot* meningkat. Proses pelatihan pada *chatbot* dilakukan sebanyak 500 *epoch* dapat mendapatkan hasil dengan *final loss* 0.0023 dan akurasi sebesar 87.85.

```
Epoch [491/500], Loss: 0.0018, Accuracy: 87.62
Epoch [492/500], Loss: 0.0089, Accuracy: 87.65
Epoch [493/500], Loss: 0.0025, Accuracy: 87.67
Epoch [494/500], Loss: 0.0017, Accuracy: 87.70
Epoch [495/500], Loss: 0.0022, Accuracy: 87.72
Epoch [496/500], Loss: 0.0024, Accuracy: 87.75
Epoch [497/500], Loss: 0.0018, Accuracy: 87.77
Epoch [498/500], Loss: 0.0017, Accuracy: 87.80
Epoch [499/500], Loss: 0.0021, Accuracy: 87.82
Epoch [500/500], Loss: 0.0023, Accuracy: 87.85
final loss: 0.0023
pelatihan selesai. file disimpan data.pth
```

Gambar 4. 5 Hasil Pelatihan

Pelatihan diatas akan disimpan pada file data.pth yang nantinya akan digunakan untuk menjadi dasar untuk menjawab pertanyaan pada *chatbot*. Berikut potongan kode pada proses penyimpanan model pelatihan untuk selanjutnya disimpan pada file data.pth.

```

# model.state_dict () untuk menyimpan model yang terlatih
data = {
    "model_terlatih": model.state_dict(),
    "ukuran_input": ukuran_input,
    "ukuran_hidden": ukuran_hidden,
    "ukuran_output": ukuran_output,
    "semua_kata": semua_kata,
    "tags": tags
}

FILE = "data.pth"
torch.save(data, FILE)

print(f'pelatihan selesai. file disimpan {FILE}')
with open('intent.json', 'r') as json_data:
    intents = json.load(json_data)

FILE = "data.pth"
data = torch.load(FILE)

```

Gambar 4. 6 Potongan Kode Proses Penyimpanan Model Pelatihan

4.3 Output

Tahapan terakhir yaitu tahap *output* berisikan jawaban berdasarkan pertanyaan dari pengguna. Proses mendapatkan hasil jawaban dari *chatbot* yaitu dengan mencari nilai terbesar dari setiap kata yang masuk dan dicocokkan dengan *tag* pada *intent*. Berikut adalah kode program respon *chatbot*.

```

def chatbot_response(msg):

    kalimat = tokenize(msg)
    X = bag_of_words(kalimat, semua_kata)

    # reshape digunakan untuk membuat atau merubah baris kolom
    X = X.reshape(1, X.shape[0])

    X = torch.from_numpy(X)

    output = model(X)

    _, prediksi = torch.max(output, dim=1)

    tag = tags[prediksi.item()]
    print(tag)

    # fungsi aktivasi relu
    kemungkinan = torch.softmax(output, dim=1)
    mungkin = kemungkinan[0][prediksi.item()]
    print(mungkin)

    for intent in intents['intents']:
        if tag == intent["tag"]:
            respon = random.choice(intent['responses'])

    return respon

```

Gambar 4. 7 Kode Program Respon *Chatbot*

Chatbot memberikan jawaban yang memiliki akurasi paling tinggi untuk diberikan kepada pengguna. Hasil pengujian membuktikan beberapa faktor yang mempengaruhi keakuratan *chatbot*. Beberapa faktor adalah kata yang mirip, pertanyaan kurang lengkap, dan pertanyaan yang tidak ada pada *intent* sehingga menyebabkan akurasi yang didapatkan oleh *chatbot* rendah. Pada tabel 4.2 dapat dilihat perbedaan akurasi dari *respon chatbot*.

Tabel 4. 2 Perbedaan Penggunaan Kata

No	Input	Tag	Akurasi
1	saya ingin lihat pengumuman utbk	utbk3	0.5831
2	kapan pengumuman utbk	utbk3	0.9980
3	berapa fakultas di uin malang	fakultas	0.9828
4	saya ingin tahu total fakultas yang ada	fakultas	0.8310

4.4 Implementasi Desain Antarmuka

Pada tahap implementasi digunakan bahasa pemrograman *python* untuk membuat model dari *neural network* dan *framework flask* berfungsi untuk membuat tampilan dari *website* yang terintegrasi dengan model *chatbot*. Bahasa pemrograman lain yang digunakan adalah HTML, dan CSS untuk membangun kerangka dan memperindah tampilan *website*.

4.4.1 Halaman Beranda

Halaman beranda menampilkan informasi mengenai *website* PMB UIN Malang dan informasi lain dapat diakses pada *navbar*. Berikut contoh halaman beranda pada *website* PMB UIN Malang.



Gambar 4. 8 Tampilan Beranda

4.4.2 Chatbot Messenger

Chatbot Messenger pada *website* PMB UIN Malang digunakan sebagai sarana untuk menjembatani pengguna dengan *bot* agar dapat berinteraksi dengan mudah. Tugas dari *chatbot* ini adalah menjawab pertanyaan dari pengguna seputar PMB UIN Malang. Untuk dapat terhubung dengan *chatbot*

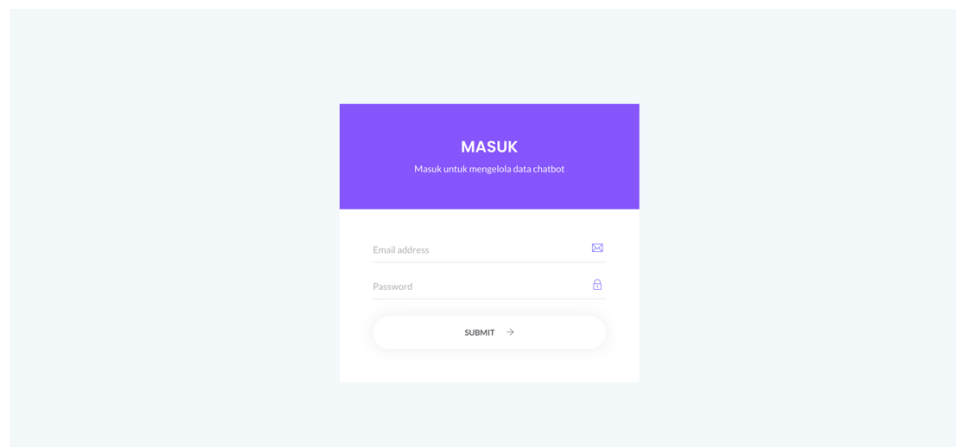
pada *website* digunakan *framework flask*. *Chatbot* yang sudah dibangun diberi nama Vatin (*Virtual Assistant* PMB UIN Malang).



Gambar 4. 9 Tampilan Chatbot

4.4.3 Login Admin

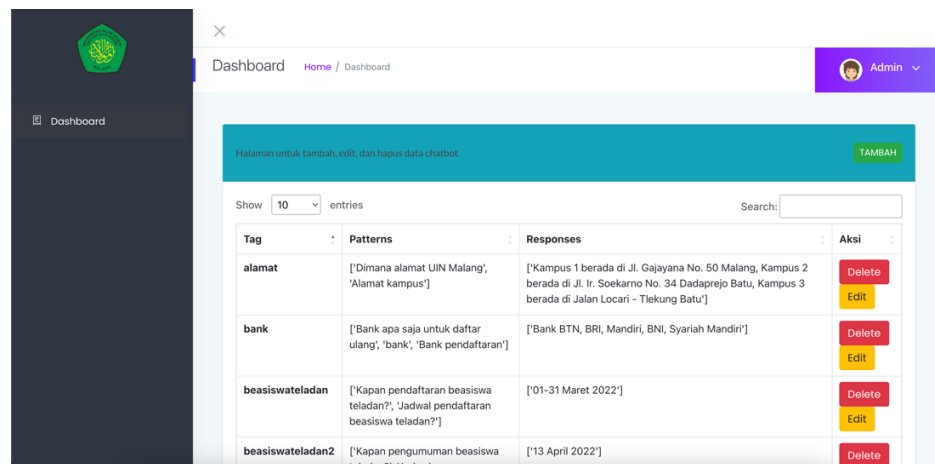
Pada halaman beranda *website* PMB UIN Malang terdapat menu login yang dapat diakses oleh admin. Admin dapat langsung masuk menggunakan email dan password yang telah tersedia.



Gambar 4. 10 Login Admin

4.4.4 Dashboard Admin

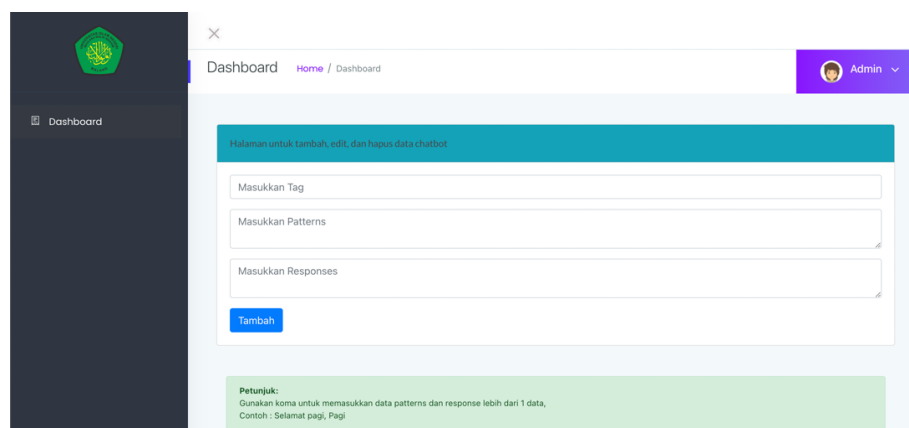
Pada halaman dashboard admin terdapat menu utama yaitu daftar data yang digunakan pada *chatbot* dan memiliki aksi untuk menambah, memperbaharui, dan menghapus data.



Gambar 4. 11 Dashboard Admin

4.4.5 Tambah Data

Pada halaman tambah data, admin dapat memasukkan *tag*, *pattern*, dan *responses* untuk menambahkan data pada *chatbot*. Untuk menambahkan data yang lebih dari dua dapat menggunakan koma karena data akan di *convert* menjadi *JSON*.



Gambar 4. 12 Tambah Data

4.4.6 Edit Data

Pada halaman edit data, admin mempunyai akses untuk mengubah data sesuai dengan kebutuhan. Terdapat 3 *form input* yaitu *tag*, *pattern*, dan *responses* dan menampilkan data sesuai dengan data yang dipilih. Proses mengedit data yang lebih dari dua dapat menggunakan koma karena data akan di *convert* menjadi *JSON*.

Gambar 4. 13 Edit Data

4.5 Hasil Uji Coba

Pada penelitian ini terdapat 51 *tags*, 119 *patterns* serta 57 *response* yang digunakan pada *dataset*. Pada data pengujian digunakan 25 *patterns* dari 20% total *patterns* yang ada pada *dataset* (Brownlee, 2020). Nilai akurasi kemudian diukur menggunakan *pattern*/kalimat masukan pada data pengujian. Pada tabel 4.3 menunjukkan potongan *pattern* yang ada di data *training*, dan ditampilkan secara lengkap dalam Lampiran I. Tabel 4.4 menunjukkan potongan kalimat yang ada di data pengujian, dan dapat dilihat secara lengkap dalam Lampiran II.

Tabel 4. 3 Data *Training*

No	Kalimat Masukan	Tag
1	Woy	error
2	Halo min	salam
3	Assalamualaikum min	salam2
4	Selamat pagi admin	salam3
5	siap terimakasih infonya min	terimakasih
6	Apa saja jalur masuk di uin malang	pendaftaran
7	Apa yang dimaksud KIP	kip
8	Wajib atau tidak tinggal di mahad	mahad
9	Pendaftaran SNMPTN mulai kapan	snmptn
10	pengumuman snmptn tanggal berapa	snmptn2

Tabel 4. 4 Data Pengujian

No	Kalimat Masukan	Tag
1	Woy	error
2	Halo min	salam
3	Assalamualaikum min	salam2
4	Selamat pagi admin	salam3
5	siap terimakasih infonya min	terimakasih
6	Apa saja jalur masuk di uin malang	pendaftaran
7	Apa yang dimaksud KIP	kip
8	Wajib atau tidak tinggal di mahad	mahad
9	Pendaftaran SNMPTN mulai kapan	snmptn
10	pengumuman snmptn tanggal berapa	snmptn2

Setelah pengujian dilakukan, langkah selanjutnya adalah menghitung kinerja sistem dengan menggunakan *confusion matrix*. Data hasil pengujian dapat dilihat dalam Lampiran III. Penelitian ini dikategorikan sebagai *binary classification* karena hanya memprediksi apakah data pengujian cocok dengan data *training*. Nilai akurasi, *recall*, presisi dan *F1 score* akan dihitung menggunakan persamaan 4.1, 4.2, dan 4.3 ,4.4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.1)$$

$$Recall = \frac{TP}{TP+FN} \quad (4.2)$$

$$Presisi = \frac{TP}{TP+FP} \quad (4.3)$$

$$F1\ score = \frac{2*recall*presisi}{recall+presisi} \quad (4.4)$$

Dimana,

1. *True Positive* (TP) merupakan hasil *query* yang kelasnya benar dan diprediksi benar oleh sistem.
2. *True Negative* (TN) merupakan hasil *query* yang tidak ada kelas dan diprediksi salah oleh sistem.
3. *False Positive* (FP) merupakan hasil *query* yang tidak ada kelas dan diprediksi benar oleh sistem.
4. *False Negative* (FN) merupakan hasil *query* yang kelasnya benar dan diprediksi salah oleh sistem.

Jumlah pengujian pada penelitian ini sebanyak 25 *test set*. Didapatkan hasil pengujian menggunakan *confusion matrix* seperti tabel 4.5.

Tabel 4. 5 Hasil *Confusion Matrix*

	<i>Positive</i> (Aktual)	<i>Negative</i> (Aktual)
<i>Positive</i> (Prediksi)	21	1
<i>Negative</i> (Prediksi)	2	1

Berdasarkan nilai *confusion matrix*, selanjutnya dilakukan perhitungan untuk mencari nilai akurasi, *recall*, presisi dan *F1 score*.

$$\text{Akurasi} = \frac{21+1}{21+1+1+2} = \frac{22}{25} = 0,88 \times 100\% = 88\%$$

$$\text{Recall} = \frac{21}{21+2} = \frac{21}{23} = 0,91 \times 100\% = 91\%$$

$$\text{Presisi} = \frac{21}{21+1} = \frac{21}{22} = 0,95 \times 100\% = 95\%$$

$$F1 \text{ score} = \frac{2 \times 0,91 \times 0,95}{0,91+0,95} = \frac{1,73}{1,86} = 0,93 \times 100\% = 93\%$$

Berdasarkan perhitungan kinerja sistem diatas, dapat diketahui nilai hasil pengukuran akurasi, *recall*, presisi, dan *F1 score* pada *chatbot* menggunakan *Neural Network* adalah 88% untuk nilai akurasi, 91% untuk nilai *recall*, 95% untuk nilai presisi, dan 93% untuk nilai *F1 score*.

4.6 Pembahasan

Setelah proses pengujian selesai, pada penelitian ini dapat menggunakan metode *Neural Network* dengan pendekatan *Natural Language Processing* yang diimplementasikan pada *chatbot*. Metode *Neural Network* dengan pendekatan *Natural Language Processing* dapat memberikan respon yang sesuai dengan tingkat akurasi yang didapat sebesar 88%, nilai *recall* sebesar 91%, nilai presisi sebesar 95%, dan nilai *F1 score* sebesar 93% dengan data pengujian sebanyak 25 kalimat masukan. Dalam proses pelatihan, jumlah *epoch* mempengaruhi akurasi sistem. Jika nilai *epoch* kecil maka nilai bobot atau *weight* setiap tiap node tidak akan diperbaharui dan akan terjadi *error* yang besar. Jika jumlah *epoch* terlalu besar juga dapat mengakibatkan *overfitting*. *Overfitting* adalah keadaan dimana model terlalu baik dalam memprediksi data *training*, tetapi tidak terlalu baik memprediksi data pada saat pengujian. Selain

itu, jumlah sampel data pertanyaan juga mempengaruhi tingkat akurasi sistem, semakin banyak data yang dimiliki, semakin akurat sistem yang dibangun.

Dalam surah Al-Furqan ayat 32 yang berbunyi :

وَقَالَ الَّذِينَ كَفَرُوا لَوْلَا نُزِّلَ عَلَيْهِ الْقُرْآنُ جُمْلَةً وَاحِدَةً كَذَلِكَ لِنُثَبِّتَ بِهِ فُؤَادَكَ وَرَتَّلْنَاهُ تَرْتِيلًا

“Dan orang-orang kafir berkata, “Mengapa Al-Qur'an itu tidak diturunkan kepadanya sekaligus?” Demikianlah, agar Kami memperteguh hatimu (Muhammad) dengannya dan Kami membacakannya secara tartil (berangsur-angsur, perlahan dan benar).”

Menurut tafsir Al-Muyassar, ayat ini menjelaskan bahwa Al-Qur'an diturunkan secara bertahap kepada Nabi Muhammad SAW, agar perasaan batin yang dimiliki menjadi kuat dan dapat menambah ketenangan. Selain itu, tujuan lainnya yaitu agar dapat memahami dan menjalankan tugas dengan baik serta dapat menjelaskan dengan teliti dan perlahan.

Berdasarkan ayat diatas memiliki kesesuaian bahwa *neural network* yang dibangun harus melewati tahapan pelatihan dan dilatih satu persatu . Selanjutnya akan dilakukan uji coba hingga menemukan akurasi yang paling baik, agar informasi yang diberikan oleh *chatbot* lebih akurat.

Chatbot yang sudah dibangun diharapkan dapat membantu dalam menjawab pertanyaan pengguna. Adanya *chatbot* dapat mempercepat jawaban yang diterima tanpa harus menunggu jawaban dari admin, karena akan otomatis dibalas oleh *bot*. Waktu merupakan hal yang seharusnya digunakan sebaik mungkin. Dalam surah Al-Ashr ayat 1 yang berbunyi :

وَالْعَصْرِ

“Demi Masa”

Menurut tafsir Jalalain dalam ayat ini Allah subhanahu wa ta'ala bersumpah dengan mengacu pada waktu. Dalam hal ini waktu adalah isyarat bagi Nabi Muhammad SAW dan orang mukmin, bahwa waktu itu penting dan berharga. Kaitannya dengan sistem yang telah dibangun adalah dengan adanya sistem ini dapat membantu dalam mencari informasi secara cepat dan akurat tanpa membuang waktu.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pengujian yang diperoleh, dapat disimpulkan bahwa *Neural Network* dapat digunakan untuk mengimplementasikan *chatbot* pada *website* PMB UIN Malang dengan memiliki tingkat akurasi sebesar 88%, nilai *recall* sebesar 91%, nilai presisi sebesar 95%, dan nilai *f1 score* sebesar 93% dengan data pengujian sebanyak 25 kalimat masukan. Pada proses pengujian jumlah *epoch* mempengaruhi akurasi pada sistem. Jika nilai *epoch* kecil maka nilai bobot atau *weight* setiap tiap node tidak akan diperbaharui dan akan terjadi *error* yang besar.. Jika jumlah *epoch* terlalu besar juga dapat mengakibatkan *overfitting*.

5.2 Saran

Berikut beberapa saran untuk penelitian selanjutnya agar pembuatan *chatbot* pada *website* PMB UIN Malang semakin baik sebagai berikut.

1. Diperlukan lebih banyak data agar dapat menjawab pertanyaan pengguna, dimana data berpengaruh besar agar mendapatkan akurasi yang baik.
2. Saran untuk penelitian selanjutnya agar menggunakan data selain bahasa Indonesia.

DAFTAR PUSTAKA

- Bhartiya, N., Jangid, N., Jannu, S., Shukla, P., & Chapaneri, R. (2019). Artificial Neural Network Based University Chatbot System. *2019 IEEE Bombay Section Signature Conference, IBSSC 2019*, 2019Januar, 1–6. <https://doi.org/10.1109/IBSSC47189.2019.8973095>
- Brownlee, J. (2020). *Train-Test Split for Evaluating Machine Learning Algorithms*. Python Machine Learning.
- Cahya Utama, R., & Titi Komalasari, R. (2021). *STRING (Satuan Tulisan Riset dan Inovasi Teknologi) APLIKASI CHATBOT BERBASIS TEKS MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER FAQ GRABADS*. 6(1).
- Cahya, N., Triayudi, A., & Benrahman, B. (2021). Implementasi Framework Codeigniter Pada Perancangan Chatbot Interaktif Menerapkan Metode Waterfall. *Jurnal Media Informatika Budidarma*, 5(1), 273. <https://doi.org/10.30865/mib.v5i1.2623>
- Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 48–57. <https://doi.org/10.1109/MCI.2014.2307227>
- Elcholiqi, A., & Musdholifah, A. (2020). Chatbot in Bahasa Indonesia using NLP to Provide Banking Information. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(1), 91. <https://doi.org/10.22146/ijccs.41289>
- Fambayun, F. D., Buntoro, G. A., Masykur, F., Teknik, F., Ponorogo, U. M., & Network, N. (2022). *PENERAPAN ALGORITMA NEURAL NETWORK*. 14(1), 40–46.
- Fathoni, F., Afrianti, E., & Heroza, R. I. (2020). Klasifikasi Teks dengan Naïve Bayes Classifier (NBC) untuk Pengelompokan Keterangan Laporan dan Durasi Recovery Time Laporan Gangguan Listrik PT. PLN (Persero) WS2JB Area Palembang. *JSI: Jurnal Sistem Informasi (E-Journal)*, 12(1), 1955–1961. <https://doi.org/10.36706/jsi.v12i1.9586>
- Guntoro, G., Loneli Costaner, & Lisnawita, L. (2020). Aplikasi Chatbot untuk Layanan Informasi dan Akademik Kampus Berbasis Artificial Intelligence Markup Language (AIML). *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 11(2), 291–300. <https://doi.org/10.31849/digitalzone.v11i2.5049>
- Gupta, B. B., Joshi, R. C., & Misra, M. (2012). ANN based scheme to predict number of zombies in a DDoS attack. *International Journal of Network Security*, 14(2), 61–70.
- Gupta, S., Borkar, D., Mello, C. De, & Patil, S. (2015). An E-Commerce Website based Chatbot. *International Journal of Computer Science and Information Technologies*, 6(2), 1483–1485.
- Hormansyah, D. S., & Utama, Y. P. (2018). Aplikasi Chatbot Berbasis Web Pada Sistem Informasi Layanan Publik Kesehatan Di Malang Dengan

- Menggunakan Metode Tf-Idf. *Jurnal Informatika Polinema*, 4(3), 224. <https://doi.org/10.33795/jip.v4i3.211>
- Husamuddin, H., Prasetyo, D. B., & Rustamadj, H. C. (2020). Otomatisasi Layanan Frequently Ask Questions Berbasis Natural Language Processing Pada Telegram Bot. *Telematika*, 17(2), 145. <https://doi.org/10.31315/telematika.v1i1.3383>
- Jonathan, W., & Lestari, S. (2015). Sistem Informasi Ukm Berbasis Website Pada Desa Sumber Jaya. *Z.A. Pagar Alam*, 01(1), 16–22.
- Khin, N. N., & Soe, K. M. (2020). University Chatbot using Artificial Intelligence Markup Language. *2020 IEEE Conference on Computer Applications, ICCA 2020*, 1–5. <https://doi.org/10.1109/ICCA49400.2020.9022814>
- Lahitani, A. R., Permanasari, A. E., & Setiawan, N. A. (2016). Cosine similarity to determine similarity measure: Study case in online essay assessment. *Proceedings of 2016 4th International Conference on Cyber and IT Service Management, CITSM 2016*. <https://doi.org/10.1109/CITSM.2016.7577578>
- Lalwani, T., Bhalotia, S., Pal, A., Bisen, S., & Rathod, V. (2018). *Implementation of a Chatbot System using AI and NLP*. 2580011(3), 26–30. <https://doi.org/10.21276/ijircst.2018.6.3.2>
- Muliyono, M., & Sumijan, S. (2021). Identifikasi Chatbot dalam Meningkatkan Pelayanan Online Menggunakan Metode Natural Language Processing. *Jurnal Informatika Ekonomi Bisnis*, 3, 142–147. <https://doi.org/10.37034/infeb.v3i4.102>
- Munasatya, N., & Novianto, S. (2020). Natural Language Processing untuk Sentimen Analisis Presiden Jokowi Menggunakan Multi Layer Perceptron. *Techno.Com*, 19(3), 237–244. <https://doi.org/10.33633/tc.v19i3.3630>
- Mustakim, F., Fauziah, F., & Hayati, N. (2021). Algoritma Artificial Neural Network pada Text-based Chatbot Frequently Asked Question (FAQ) Web Kuliah Universitas Nasional. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 5(4), 438. <https://doi.org/10.35870/jtik.v5i4.261>
- Nia Agustina Purwitasari, & Soleh, M. (2022). *View of Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Processing.pdf*.
- Putra, I. M. S., Putu Jhonarendra, & Ni Kadek Dwi Rusjyanthi. (2021). Deteksi Kesamaan Teks Jawaban pada Sistem Test Essay Online dengan Pendekatan Neural Network. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1070–1082. <https://doi.org/10.29207/resti.v5i6.3544>
- Putri, A. W. (2021). *IMPLEMENTASI ARTIFICIAL NEURAL NETWORK (ANN) BACKPROPAGATION UNTUK KLASIFIKASI JENIS PENYAKIT PADA DAUN TANAMAN TOMAT*. 09(02), 344–350.
- Rezkanintio, A. N. (2021). *Klasifikasi Intent Pada Chatbot Menggunakan Metode Query Expansion Dan Neural Network*.

- Riani, N. A., Andreswari, R., Fauzi, R., & Informasi, S. (2021). *IMPLEMENTASI ALGORITMA ARTIFICIAL NEURAL NETWORK*. 4307(3), 241–247.
- S. C. P, E. N., & Afrianto, I. (2015). Rancang Bangun Aplikasi Chatbot Informasi Objek Wisata Kota Bandung Dengan Pendekatan Natural Language Processing. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, 4(1), 49–54. <https://doi.org/10.34010/komputa.v4i1.2410>
- Salisah, T., Sari, B. P., Yulianto, Y., & Hartanto, A. D. (2020). Implementasi Algoritma Boyer-Moore Pada Chatbot Wisata Yogyakarta. *Technomedia Journal*, 5(1), 54–66. <https://doi.org/10.33050/tmj.v5i1.1189>
- Sibagariang, S., Riyadi, A., Dzikri, A., Suandi, F., Sirait, K. T., & Setiawan, F. (2021). Prediksi Prospek Kerja Alumni Dengan Algoritma Neural Network. *CESS (Journal of Computer Engineering, System and Science)*, 6(1), 91. <https://doi.org/10.24114/cess.v6i1.22237>
- Wangsanegara, N. K., & Subaeki, B. (2015). Implementasi Natural Language Processing Dalam Pengukuran Ketepatan Ejaan Yang Disempurnakan (Eyd) Pada Abstrak Skripsi Menggunakan Algoritma Fuzzy Logic. *Jurnal Teknik Informatika*, 8(2). <https://doi.org/10.15408/jti.v8i2.3185>
- Zurairiyah, T. A., Utami, D. K., Herlambang, D., Studi, P., Komputer, I., Matematika, F., Ilmu, D., Alam, P., Pakuan, U., Bogor, K., & Barat, J. (2019). *MAHASISWA BARU MENGGUNAKAN RECURRENT NEURAL NETWORK*. 91–101.

LAMPIRAN

Lampiran I

Train Set

No	Kalimat Masukan	Tag	Jawaban
1	Hai	salam	Maaf, Kami tidak mengerti
2	Halo, hai, hallo, salam, hola, permisi, mau tanya min, min	salam	Hai, ada yang bisa dibantu? , Halo! Apa yang bisa mimin bantu?
3	Assalamualaikum, Assalamu'alaikum, Assalamu'alaikum Warahmatullahi, Assalamu'alaikum Warahmatullahi Wabarakatuh	salam2	Waalaikummussalam
4	Selamat pagi, pagi, good morning, morning	salam3	Selamat pagi, good morning
5	Selamat siang, siang, good afternoon	salam4	Selamat siang, good afternoon
6	Selamat malam, malam, good night	salam5	Selamat malam, good night
7	Terima kasih, makasih, trims, oke, sangat membantu	terimakasih	Senang bisa membantu, tentu tidak masalah, terimakasih kembali
8	Jalur masuk uin malang	pendaftaran	SNMPTN, SBMPTN, SPAN PTKIN, UM-PTKIN, Mandiri, Beasiswa Teladan, Pascasarjana
9	Petunjuk teknis KIP, apa itu KIP	kip	Silahkan akses website PMB UIN Malang dan menuju menu kip
10	Apakah di uin malang wajib mahad, asrama	mahad	Wajib tinggal di mahad pada semester 1 dan 2
11	Kapan pendaftaran snmptn, jadwal pendaftaran snmptn	snmptn	14-28 februari 2022
12	Kapan pengumuman snmptn, jadwal pengumuman snmptn	snmptn2	29 maret 2022
13	Alur daftar ulang snmptn	snmptn3	Silahkan akses pada halaman website

			https://verifikasi.uin-malang.ac.id/
14	kapan pengisian biodata snmptn, atas pengisian biodata snmptn	snmptn4	pengisian biodata jalur snmptn pada tanggal 29 Maret - 10 April 2022
15	kapan validasi snmptn, kapan daftar ulang snmptn	snmptn5	pengisian validasi snmptn pada tanggal 13 Maret - 30 April 2022
16	Saya ingin melihat hasil snmptn, cek hasil snmptn	snmptn6	Silahkan akses menu SNMPTN dan cek kelulusan anda
17	kapan pendaftaran beasiswa teladan, Jadwal pendaftaran beasiswa teladan?	beasiswateladan	01-31 Maret 2022
18	Kapan pengumuman beasiswa teladan, Jadwal pengumuman beasiswa teladan	beasiswateladan2	13 April 2022
19	Kapan pelaksanaan seleksi beasiswa teladan, Kapan pelaksanaan seleksi beasiswa teladan	beasiswateladan3	6-7 April 2022
20	Apa kategori penerima beasiswa teladan, Syarat beasiswa teladan	beasiswateladan4	Mempunyai hafalan al-quran 30 Juz, Memiliki kemampuan qira'atul kutub, dan berprestasi dalam bidang olahraga dan seni tingkat nasional atau internasional
21	Saya ingin melihat hasil beasiswa teladan, cek hasil beasiswa teladan	beasiswateladan5	Silahkan akses menu beasiswa teladan dan cek kelulusan anda
22	Kapan pendaftaran span ptkin, Jadwal pendaftaran span ptkin	span	4-31 Maret 2022
23	Kapan pelaksanaan seleksi span ptkin, Jadwal pelaksanaan seleksi span ptkin	span2	4-7 April 2022
24	Kapan pengumuman span ptkin, Jadwal	span3	15 April 2022

	pengumuman span ptkin		
25	Saya ingin melihat hasil SPAN PTKIN, cek hasil SPAN PTKIN	span4	Silahkan akses menu SPAN PTKIN dan cek kelulusan anda
26	Kapan pendaftaran mandiri, Jadwal pendaftaran mandiri	mandiri	1 Juni - 18 Juli 2022
27	Kapan pelaksanaan seleksi mandiri, Jadwal pelaksanaan seleksi mandiri	mandiri2	20-22 Juli 2022
28	Kapan pengumuman mandiri, Jadwal pengumuman mandiri	mandiri3	27 Juli 2022
29	Saya ingin melihat hasil mandiri, cek hasil mandiri	mandiri4	Silahkan akses menu mandiri dan cek kelulusan anda
30	Kapan pendaftaran UTBK, Jadwal pendaftaran UTBK	utbk	23 Maret - 15 April 2022
31	Kapan pelaksanaan seleksi UTBK, Jadwal pelaksanaan seleksi UTBK	utbk2	17-23 Mei 2022
32	Kapan pengumuman UTBK, Jadwal pengumuman UTBK	utbk3	23 Juni 2022
33	Saya ingin melihat hasil UTBK, cek hasil UTBK	utbk4	Silahkan akses menu UTBK dan cek kelulusan anda
34	Kapan pendaftaran UMPTKIN, Jadwal pendaftaran UMPTKIN	umptkin	25 April - 4 Juni 2022
35	Kapan pelaksanaan seleksi UMPTKIN, Jadwal pelaksanaan seleksi UMPTKIN	umptkin2	14-17 Juni 2022
36	Kapan pengumuman UMPTKIN, jadwal pengumuman UMPTKIN	umptkin3	30 Juni 2022

37	Saya ingin melihat hasil UMPTKIN, cek hasil UMPTKIN	umptkin4	Silahkan akses menu UMPTKIN dan cek kelulusan anda
38	Kapan pendaftaran pascasarjana, Jadwal pendaftaran pascasarjana	pascasarjana	1 Juni - 18 Juli 2022
39	Kapan pelaksanaan seleksi pascasarjana, Jadwal pelaksanaan seleksi pascasarjana	pascasarjana2	20-22 Juli 2022
40	Kapan pengumuman pascasarjana, Jadwal pengumuman pascasarjana	pascasarjana3	30 Juli 2022
41	Saya ingin melihat hasil pascasarjana, cek hasil pascasarjana	pascasarjana4	Silahkan akses menu pascasarjana dan cek kelulusan anda
42	Berapa biaya pendaftaran, Berapa harga pendaftaran, harga masuk	biayapendaftaran	Biaya pendaftaran sebesar Rp300.000
43	Ujian dilaksanakan online atau daring, online, daring	pelaksanaan	Ujian dilaksanakan secara online
44	berapa banyak jurusan, jurusan	jurusan	UIN Malang memiliki 41 jurusan
45	Jurusan apa saja yang ada di UIN Malang, Daftar jurusan	jurusan2	Daftar jurusan dapat dilihat pada http://pmb.uin-malang.ac.id/index.php/category/programstudi/
46	berapa banyak fakultas, fakultas	fakultas	UIN Malang memiliki 7 fakultas
47	Nomor whatsapp, Nomor yang dapat dihubungi, whatsapp	whatsapp	Nomor admin PMB UIN Malang +6281553402097
48	Berapa biaya mahad, Berapa biaya asrama, biaya mahad	biayamahad	Biaya mahad 1 tahun Rp7.500.000 kecuali kedokteran Rp10.000.000 per tahun
49	Bank apa saja untuk daftar ulang, bank, Bank pendaftaran	bank	Bank BTN, BRI, Mandiri, BNI, Syariah Mandiri
50	Siapa nama rektor UIN Malang, rektor	rektor	Prof. Dr. H. M. Zainuddin, MA

51	Dimana alamat UIN Malang, Alamat kampus	alamat	Kampus 1 berada di Jl. Gajayana No. 50 Malang, Kampus 2 berada di Jl. Ir. Soekarno No. 34 Dadaprejo Batu, Kampus 3 berada di Jalan Locari - Tlekung Batu
----	---	--------	--

Lampiran II

Pengujian

No	Kalimat Masukan	Tag
1	Woy	error
2	Halo min	salam
3	Assalamualaikum min	salam2
4	Selamat pagi admin	salam3
5	siap terimakasih infonya min	terimakasih
6	Apa saja jalur masuk di uin malang	pendaftaran
7	Apa yang dimaksud KIP	kip
8	Wajib atau tidak tinggal di mahad	mahad
9	Pendaftaran SNMPTN mulai kapan	snmptn
10	pengumuman snmptn tanggal berapa	snmptn2
11	Alur daftar ulang snmptn	snmptn3
12	Beasiswa teladan kapan dibuka pendaftaran nya	beasiswateladan
13	Jadwal pengumuman beasiswa teladan	beasiswateladan2
14	apa saja syarat untuk mendapatkan beasiswa teladan	beasiswateladan4
15	Kapan pengumuman span ptkin keluar	span3
16	Bagaimana cara cek hasil tes mandiri	mandiri4
17	Kapan jadwal pendaftaran utbk tahun ini	utbk
18	Jadwal pengumuman umptkin kapan keluar	umptkin3
19	Saya ingin melihat hasil seleksi pascasarjana	pascasarjana4
20	Berapa biaya masuk untuk pendaftaran di uin malang	biayapendaftaran
21	Ujian dilaksanakan secara luring atau daring	pelaksanaan
22	Ada berapa banyak jurusan di uin malang	jurusan
23	berapa banyak jumlah fakultas di uin malang	fakultas
24	Nomor yang dapat saya hubungi	whatsapp
25	hei, nama kamu siapa?	error

Lampiran III

Hasil Pengujian

No	Kalimat Masukan	Tag	Prediksi	Target
1	Woy	error	Benar	Benar
2	Saya perlu informasi min	salam	Salah	Benar
3	Assalamualaikum min	salam	Benar	Salah
4	Selamat pagi admin	salam3	Benar	Benar
5	siap terimakasih infonya min	salam	Benar	Salah

6	Apa saja jalur masuk di uin malang	pendaftaran	Benar	Benar
7	Apa yang dimaksud KIP	kip	Benar	Benar
8	Wajib atau tidak tinggal di mahad	mahad	Benar	Benar
9	Pendaftaran SNMPTN mulai kapan	snmptn	Benar	Benar
10	pengumuman snmptn tanggal berapa	snmptn2	Benar	Benar
11	Alur daftar ulang snmptn	snmptn3	Benar	Benar
12	Beasiswa teladan kapan dibuka pendaftaran nya	beasiswateladan	Benar	Benar
13	Jadwal pengumuman beasiswa teladan	beasiswateladan2	Benar	Benar
14	apa saja syarat untuk mendapatkan beasiswa teladan	beasiswateladan4	Benar	Benar
15	Kapan pengumuman span ptkin keluar	span3	Benar	Benar
16	Bagaimana cara cek hasil tes mandiri	mandiri4	Benar	Benar
17	Kapan jadwal pendaftaran utbk tahun ini	utbk	Benar	Benar
18	Jadwal pengumuman umptkin kapan keluar	umptkin3	Benar	Benar
19	Saya ingin melihat hasil seleksi pascasarjana	pascasarjana4	Benar	Benar
20	Berapa biaya masuk untuk pendaftaran di uin malang	biayapendaftaran	Benar	Benar
21	Ujian dilaksanakan secara luring atau daring	pelaksanaan	Benar	Benar
22	Ada berapa banyak jurusan di uin malang	jurusan	Benar	Benar
23	berapa banyak jumlah fakultas di uin malang	fakultas	Benar	Benar
24	Nomor yang dapat saya hubungi	whatsapp	Benar	Benar
25	Hei, nama kamu siapa?	rektor	Salah	Salah