

**PENERAPAN NEURAL NETWORK PADA CHATBOT  
TANYA JAWAB ILMU TAJWID**

**SKRIPSI**

**MAJIDAH ATMAYANA PURBA**

**171402120**



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS SUMATERA UTARA  
MEDAN  
2024**

**PENERAPAN NEURAL NETWORK PADA CHATBOT  
TANYA JAWAB ILMU TAJWID**

**SKRIPSI**

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah Sarjana  
Teknologi Informasi

**MAJIDAH ATMAYANA PURBA**

**171402120**



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS SUMATERA UTARA  
MEDAN  
2024**

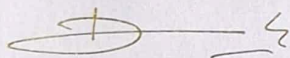
**PERSETUJUAN**

Judul : Penerapan Neural Network pada Chatbot Tanya Jawab Ilmu Tajwid  
Kategori : Skripsi  
Nama : Majidah Atmayana Purba  
Nomor Induk Mahasiswa : 171402120  
Program Studi : Sarjana (S-1) Teknologi Informasi  
Fakultas : Ilmu Komputer dan Teknologi Informasi  
Universitas Sumatera Utara

Medan, 12 Juli 2024

Komisi Pembimbing

Pembimbing 2



Baihaqi Siregar, S.Si., M.T.

NIP. 197901082012121002

Pembimbing 1



Ade Sarah Huzaifah, S. Kom., M. Kom

NIP. 198506302018032001

Diketahui/disetujui oleh

Program Studi Teknologi Informasi

Ketua,


Dedy Arisandi

NIP. 197908312009121002

**PERNYATAAN**

**PENERAPAN NEURAL NETWORK PADA CHATBOT  
TANYA JAWAB ILMU TAJWID**

**SKRIPSI**

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan,

Majidah Atmayana Purba

171402120

## UCAPAN TERIMA KASIH

Segala puji syukur kepada Allah SWT, karena atas rahmat-Nya penulis dapat menyelesaikan skripsi ini sebagai syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Skripsi ini penulis persembahkan kepada kedua orangtua penulis, Ayah dan Ibu dan juga adik-adik penulis Muhaini dan Ali yang selalu menjadi sumber semangat dan inspirasi dalam hidup.

Ucapan terima kasih penulis sampaikan kepada:

1. Bapak Dr. Muryanto Amin, S.Sos., M.Si. selaku Rektor Universitas Sumatera Utara.
2. Ibu Dr. Maya Silvi Lydia, B.Sc., M.Sc. selaku Dekan Fasilkom-TI Universitas Sumatera Utara.
3. Bapak Dedy Arisandy, S.T., M.T., selaku Ketua Program Studi S1 Teknologi Informasi Universitas Sumatera
4. Ibu Sarah Purnamawati, ST., MSc., selaku Sekretaris Program Studi S1 Teknologi Informasi Universitas Sumatera.
5. Ibu Ade Sarah Huzaifah, S.Kom, M. Kom. selaku Dosen Pembimbing I dan Bapak Baihaqi Sirega, S.Si., M.Kom., selaku Dosen Pembimbing 2 yang telah banyak meluangkan waktu, pemikiran, motivasi, kritik dan saran kepada penulis.
6. Ibu Sarah Purnamawati, ST., MSc. dan Bapak Fahrurrozi Lubis B. IT., M. ScIT selaku Dosen Pembimbing yang telah memberikan kritik dan saran yang membangun dalam proses penyempurnaan skripsi ini.
7. Staff dan pegawai Fasilkom-TI Universitas Sumatera Utara yang membantu segala urusan administrasi dalam menyelesaikan skripsi.
8. Sahabat seperjuangan yang sudah penulis anggap seperti keluarga sendiri yang tidak bisa disebutkan satu persatu, yang sudah bersama sejak awal perkuliahan, tempat berbagi cerita yang selalu memberikan dukungan dan energi positif kepada penulis dalam pengerjaan skripsi ini.

9. Penulis mengucapkan terima kasih kepada teman-teman angkatan 2017, khususnya KOM C yang telah menemani, memotivasi, memahami dan menerima sifat serta perilaku penulis, memberikan kritik dan saran yang baik selama proses perkuliahan maupun dalam masa pengerjaan skripsi.
10. Penulis juga mengucapkan terima kasih kepada rekan-rekan organisasi UKMI Al-Khuwarizmi Fasilkom-TI USU, UKMI Ad-Dakwah USU, yang selalu mendoakan dan memberikan semangat agar penulis senantiasa berjuang dalam kebaikan.
11. Terima kasih penulis ucapkan untuk semua pihak yang telah terlibat dalam pengumpulan data maupun pengujian sistem.

Medan, 12 Juli 2024

Penulis

## **PENERAPAN NEURAL NETWORK PADA CHATBOT TANYA JAWAB ILMU TAJWID**

### **ABSTRAK**

Penggunaan chatbot telah digunakan dalam berbagai bidang seperti ekonomi, pendidikan, kesehatan, dan sebagainya. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi *chatbot* sebagai alat bantu belajar ilmu tajwid bagi para pembelajar Al-Qur'an. Ilmu tajwid adalah sangat penting karena mempelajari cara membaca Al-Qur'an dengan benar sesuai dengan tuntunan Rasulullah saw. Metode penelitian yang digunakan adalah mengembangkan Neural Network untuk pencocokan string untuk memberikan respon yang sesuai dengan jawaban yang dibutuhkan. Pembangunan model *machine learning* melibatkan penggunaan *hyperparameter* tertentu, dikelola dengan PyTorch untuk mendefinisikan dataset dan *data loader*, dilatih dengan algoritma Adam, dan fungsi loss *CrossEntropyLoss*. Chat diuji dengan 30 pertanyaan dari user dan hasil menunjukkan bahwa chatbot dapat menjawab pertanyaan user dengan cukup baik dengan hasil F1-score sebesar 0,84.

Kata kunci : *chatbot*, *neural network*, ilmu tajwid.

**NEURAL NETWORK APPLICATION IN QUESTION ANSWERING :  
CHATBOT TAJWEED**

**ABSTRACT**

*The use of chatbots is widely used in various fields such as economics, education, health, and so on. This research aims to develop and evaluate a chatbot as a learning tool for tajweed for Qur'an learners. The science of tajweed is very important because it learns how to read the Qur'an correctly according to the guidance of the Prophet Muhammad PBUH. The research method used is to develop a Neural Network for string matching to provide a response that matches the required answer. The construction of the machine learning model involves the use of certain hyperparameters, managed with PyTorch to define the dataset and data loader, trained with Adam's algorithm, and CrossEntropyLoss loss function. The chat was tested with 30 questions from users and the results show that the chatbot can answer user questions quite well with an F1-score of 0.84.*

*Keywords : chatbot, neural network, tajweed.*



## DAFTAR ISI

PERSETUJUAN .....	Error! Bookmark not defined.
PERNYATAAN .....	iv
UCAPAN TERIMA KASIH .....	v
PENERAPAN NEURAL NETWORK PADA CHATBOT .....	vii
TANYA JAWAB ILMU TAJWID .....	vii
ABSTRAK .....	vii
ABSTRACT .....	viii
DAFTAR TABEL .....	xi
DAFTAR GAMBAR .....	xii
BAB 1 .....	1
PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah .....	3
1.5 Manfaat Penelitian .....	3
1.6 Metodologi Penelitian .....	4
1. Studi literatur .....	4
2. Analisis Permasalahan .....	4
3. Implementasi .....	4
4. Pengujian .....	4
5. Dokumentasi proses dan penyusunan laporan .....	4
1.7 Sistematika Penulisan .....	4
BAB 1 : PENDAHULUAN .....	4
BAB 2 : LANDASAN TEORI .....	5
BAB 3 : ANALISA DAN PERANCANGAN .....	5
BAB 4 : IMPLEMENTASI DAN PEMBAHASAN .....	5
BAB 5 : KESIMPULAN DAN SARAN .....	5
BAB 2 .....	6
LANDASAN TEORI .....	6
2.1 Ilmu Tajwid .....	6

2.2 <i>Natural Language Processing (NLP)</i> .....	6
2.4 <i>Chatbot</i> .....	9
2.5 <i>Natural Language Toolkit (NLTK)</i> .....	9
2.6 Pytorch .....	10
2.7 Tokenisasi .....	11
2.8 <i>Stemming</i> .....	11
2.9 <i>Bag of Words (BoW)</i> .....	12
2.10 Flask .....	12
2.11 Evaluasi Model .....	12
2.12 Penelitian Terdahulu .....	13
BAB 3 .....	16
ANALISIS DAN PERANCANGAN .....	16
3.1 Persiapan Data .....	16
3.2 Perancangan sistem .....	19
3.2.1 Arsitektur Umum .....	19
3.3 Antarmuka Sistem .....	29
BAB 4 .....	32
IMPLEMENTASI DAN PEMBAHASAN .....	32
4.1 Implementasi Sistem .....	32
4.1.1 Spesifikasi, <i>hardware</i> dan <i>software</i> .....	32
4.2 Implementasi Perancangan Antarmuka .....	32
4.3 Pengujian Sistem .....	36
4.4 Evaluasi Model .....	38
BAB 5 .....	41
KESIMPULAN DAN SARAN .....	41
5.1. Kesimpulan .....	41
5.2. Saran .....	41
DAFTAR PUSTAKA .....	42
LAMPIRAN .....	44

## DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu .....	13
Tabel 3. 1 Pembagian Data Latih dan Data Uji .....	16
Tabel 3. 2 Contoh Pertanyaan dan Jawaban .....	16
Tabel 3. 3 Contoh Data JSON .....	17
Tabel 3. 4 Contoh penerapan tokenisasi .....	20
Tabel 3. 5 Contoh penerapan stemming .....	20
Tabel 3. 6 Contoh encoding Bag of words .....	21
Tabel 3. 8 Pemodelan Neural Network dengan Modul NeuralNet .....	23
Tabel 4. 1 Contoh data uji .....	36

## DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Feed Forward Neural Network .....	7
Gambar 2. 2 Arsitektur FFNN dengan Perhitungan Bobot .....	8
Gambar 3. 1 Arsitektur Umum .....	19
Gambar 3. 2 Proses Training .....	27
Gambar 3. 3 Rancangan tampilan awal chatbot .....	30
Gambar 3. 4 rancangan tampilan halaman chat saat melakukan percakapan .....	31
Gambar 4 1 Tampilan halaman awal chatbot .....	33
Gambar 4 2 Tampilan kotak percakapan chatbot. ....	33
Gambar 4 4 Pertanyaan mengenai mad. ....	34
Gambar 4 5 Pertanyaan mengenai mad badal .....	35
Gambar 4 6 respon terhadap pertanyaan mengenai mad badal .....	35
Gambar 4 7 Tabel Confusion Matrix .....	38

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Al-Quran adalah tuntunan hidup bagi umat muslim. Sebagai pedoman hidup Al-Quran wajib dipelajari oleh umat muslim itu sendiri. Al-Quran diturunkan dalam bahasa Arab yang memiliki aturan tertentu dalam pengucapan huruf dan pertemuan antar huruf-hurufnya. Dalam membaca Al-Quran, langkah pertama yang dilakukan adalah belajar membaca aksara arab dan tatacara pelafalannya. Ilmu tatacara membaca Al-Quran disebut ilmu tajwid.

Kemampuan membaca Al-Quran dan mengerti ilmu tajwid merupakan dua hal yang tidak dapat dipisahkan (Safrina & Irfan, 2022). Sebab itu ilmu tajwid sangat penting untuk dipelajari oleh umat muslim. Untuk mendapat hikmah dari membaca Al-Quran, seorang muslim wajib membacanya dengan benar. Sedikit kesalahan saat membaca Al-Quran dapat mengubah makna dari ayat yang dibaca. Sebagai contoh pada potongan ayat **وَلَعَلَّكُمْ تَشْكُرُونَ** (Q.S. Al-Jatsiyah : 12). Ayat tersebut dibaca *la'allakum tasykuruun* yang berarti “*dan mudah-mudahan kamu bersyukur*”. Apabila pada kata *tasykuruun* huruf syin (ش) dibaca sebagai huruf sin (س) menjadi *taskuruun*, maka maknanya berubah menjadi “*dan mudah-mudahan kamu mabuk*”. Oleh karena itu, standar baku bacaan Al-Quran berada pada ilmu tajwid (Sa'dijah, C., 2021).

Menurut Kementerian Agama, hasil survey Indeks Literasi Al-Quran skor Indonesia berada di angka 66,038 dengan persentase yang bisa membaca Al-Quran dengan lancar sesuai tajwid sebesar 44,57%. Secara umum, ada sekitar 50% lebih penduduk muslim di Indonesia belum pandai membaca Al-Quran (Kemenag, 2023). Angka yang bernilai cukup besar jika dilihat dari total jumlah penduduk Indonesia yang berjumlah 275,5 juta.

Tidak ada batasan usia dalam belajar ilmu tajwid. Sebaiknya ilmu tajwid mulai dipelajari ketika masih berusia anak-anak. Selain itu, ilmu tajwid harus dipelajari dengan pendampingan dari ahli atau guru. Dikarenakan beberapa kondisi, tidak sedikit orang mempelajari ilmu tajwid dalam ragam usia termasuk ketika sudah dewasa. Faktor yang bisa menghambat seorang muslim untuk belajar ilmu tajwid seperti tidak mengetahui urgensi belajar ilmu tajwid, malu belajar saat sudah dewasa, atau belum menemukan guru yang tepat.

*Chatbot* adalah media interaksi untuk bertanya jawab tentang suatu topik. *Chatbot* mengizinkan pengguna berinteraksi dengan mesin dengan mulus seolah-olah berbicara dengan manusia yang lain (Gaikwad, 2019). *Chatbot* menawarkan percakapan yang nyaman, personal dan memberikan respon yang tepat sesuai dengan keinginan pengguna. Penggunaan *chatbot* sudah merambat ke berbagai layanan. Dalam bidang ekonomi, *chatbot* membantu penjual menjawab pertanyaan konsumen tanpa harus terhubung 24 jam. Dalam bidang kesehatan, *chatbot* dimanfaatkan untuk bertanya jawab seputar kesehatan. Begitu pula dalam bidang pendidikan, *chatbot* digunakan sebagai mesin yang dapat menjawab pertanyaan pengguna mengenai sebuah topik. *Chatbot* membantu manusia mendapatkan informasi tanpa menunggu lama.

Dengan teknologi *chatbot*, pengguna bisa bertanya jawab seolah mempunyai guru untuk mempelajari ilmu tajwid secara mandiri. (Pradana, 2022) membuat *chatbot* untuk memudahkan muslim dalam mencari informasi tentang tajwid menggunakan algoritma *Leveinsthein Distance*. Hasil penelitian ini mendapat akurasi sebesar 100% dengan pengujian *dataset* langsung dan 92% dengan pertanyaan sembarang.

Karena *chatbot* melakukan pekerjaan seperti manusia, maka ada beberapa teknologi yang dapat digunakan dalam membangunnya. *Artificial Intelligence* (AI) yang populer belakangan ini memiliki cabang ilmu bernama Jaringan Saraf Tiruan (*Artificial Neural Network*) atau dikenal juga dengan istilah *Deep Learning*, metode yang terinspirasi dari cara kerja otak manusia. Model-model pada *Deep Learning* memiliki arsitektur yang berbeda dan masing-masing bisa digunakan pada kasus spesifik (Gaikwad, 2019). Beberapa model ini seperti CNN (*Convolutional Neural Network*), RNN (*Reccurent Neural Network*), dan LSTM (*Long Short-Term Memory*) telah banyak diterapkan dalam pembuatan *chatbot*.

Dengan memadukan *chatbot* dengan teknologi jaringan saraf tiruan, pengguna dapat melakukan tanya jawab interaktif mengenai ilmu tajwid. *Chatbot* ilmu tajwid bisa menjadi langkah awal yang tepat untuk belajar ilmu tajwid secara mandiri. Diharapkan dengan belajar tajwid dengan *chatbot* bisa memadai pengetahuan dasar pengguna terhadap ilmu tajwid.

Berdasarkan uraian tersebut penulis bermaksud melakukan penelitian untuk menciptakan media belajar ilmu tajwid berupa *chatbot* dengan judul **“Penerapan Neural Network pada Chatbot Tanya Jawab Ilmu Tajwid”**.

## 1.2 Rumusan Masalah

Kurangnya pemahaman seorang muslim tentang ilmu tajwid bisa mengubah makna ayat Al-Quran yang dibacanya. Faktor yang menghambat seorang muslim untuk belajar ilmu tajwid diantaranya adalah kurangnya kesadaran terhadap urgensi belajar ilmu tajwid, malu untuk belajar saat sudah dewasa serta belum menemukan guru yang tepat. *Chatbot* tanya jawab bisa menjadi langkah awal untuk belajar tajwid secara mandiri. Diharapkan *chatbot* ini bisa media untuk mendapat pengetahuan dasar tajwid sehingga setelah pengguna mempunyai pengetahuan dasar dan diharapkan termotivasi untuk belajar lebih dalam (tahsin).

## 1.3 Tujuan Penelitian

Adapun tujuan dalam penelitian ini adalah menerapkan *Neural Network* dalam membangun *chatbot* tanya jawab ilmu tajwid sebagai sarana belajar ilmu tajwid dasar khususnya bagi pemula yang baru belajar ilmu tajwid.

## 1.4 Batasan Masalah

Batasan masalah yang diberikan pada penelitian ini bertujuan untuk menghindari pelebaran pokok masalah agar penelitian menjadi lebih terarah sehingga tujuan penelitian tercapai. Adapun batasan masalah yang diberikan pada penelitian ini adalah sebagai berikut :

1. *Chatbot* berinteraksi dengan bahasa Indonesia dan teks latin.
2. Aplikasi *chatbot* hanya dapat dijalankan secara *online* melalui web.
3. *Chatbot* hanya bisa memberikan jawaban berdasarkan pertanyaan *user*.
4. Batasan topik yang dapat dijawab chatbot hanya seputar mad, hukum nun mati dan hukum mim mati.
5. Data diambil dari pengumpulan informasi secara manual dari buku tentang ilmu tajwid yang disusun oleh H. Ahmad Muzammil MF, Al Hafizh dan tanya jawab dengan guru tahsin.

## 1.5 Manfaat Penelitian

Berikut ini adalah manfaat yang dapat diperoleh dari penelitian ini.

1. Menawarkan kemudahan untuk mendapat informasi yang tepat mengenai ilmu tajwid dengan cepat dan fleksibel.
2. Menjadi referensi dalam bidang *bot* untuk pengembangan seterusnya.

## 1.6 Metodologi Penelitian

Tahapan-tahapan yang dilakukan dalam menyusun penelitian ini adalah :

### 1. Studi literatur

Tahapan ini adalah melakukan pengumpulan dan mempelajari informasi yang berkaitan dengan *chatbot*, *natural language processing*, *neural network*, ilmu tajwid, dan pemrograman *python*. Informasi ini diperoleh dari jurnal, buku, artikel, skripsi dan sumber informasi yang lainnya.

### 2. Analisis Permasalahan

Adapun pada tahap ini adalah melakukan analisis permasalahan berdasarkan informasi yang telah dikumpulkan untuk mendapatkan pemahaman mengenai metode yang diterapkan yaitu *neural network* untuk dilatih dengan dataset tajwid.

### 3. Implementasi

Tahapan ini adalah melakukan implementasi *neural network* dalam pembuatan model *chatbot* menggunakan pemrograman *python*.

### 4. Pengujian

Tahapan ini adalah melakukan pengujian prakiraan terhadap model yang sudah dibuat untuk memastikan model berjalan sesuai yang diharapkan.

### 5. Dokumentasi proses dan penyusunan laporan

Tahap terakhir ini adalah melakukan dokumentasi hasil analisis dari penelitian penerapan metode *neural network* dalam pembuatan model *chat* ilmu tajwid.

## 1.7 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri dari lima bagian utama sebagai berikut :

### **BAB 1 : PENDAHULUAN**

Bab pendahuluan terdiri atas latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.



**BAB 2 : LANDASAN TEORI**

Bab ini menjelaskan teori apa saja yang digunakan untuk memahami permasalahan yang dibahas. Pada bab ini juga dijelaskan tentang pemrosesan bahasa alami dengan menerapkan *neural network* untuk pembuatan model *chatbot*.

**BAB 3 : ANALISA DAN PERANCANGAN**

Bab ini menjelaskan analisis dan penerapan *neural network* dalam pembuatan model *chatbot*. Bab ini membahas arsitektur umum dan *neural network* dengan menggunakan pemograman Python.

**BAB 4 : IMPLEMENTASI DAN PEMBAHASAN**

Bab ini menjelaskan pembahasan tentang implementasi dari analisis dan perancangan yang telah disusun pada Bab 3 dan pembahasan untuk memastikan metode telah sesuai dengan yang diharapkan. Selain itu terdapat pemaparan hasil dari pengujian sistem yang sudah dibangun.

**BAB 5 : KESIMPULAN DAN SARAN**

Bab ini berisi tentang kesimpulan hasil penelitian dan saran-saran untuk penelitian selanjutnya.

## **BAB 2**

### **LANDASAN TEORI**

#### **2.1 Ilmu Tajwid**

Ilmu tajwid sangat penting, karena membahas tentang aturan dan cara membaca Al-Quran dengan sebaik-baiknya sesuai dengan tuntunan Rasulullah saw. Ilmu tajwid meliputi mengeluarkan bunyi huruf dari asal tempat keluar huruf (makhraj) dan pengucapannya diharuskan sesuai dengan karakter bunyi (sifat) dan konsekuensi dari sifat yang dimiliki huruf tersebut. Selain itu, ketika membaca Al-Quran harus mengetahui dimana harus berhenti dan dimana harus memulai bacaannya kembali serta mengetahui sifat-sifat khusus dalam membaca Al-Quran. Ilmu tajwid penting dipelajari agar umat Islam bisa membaca Al-Quran sesuai dengan bacaan yang diajarkan Rasulullah SAW dan para sahabatnya, sebagaimana Al-Quran diturunkan (Solikhah dalam Syaifullah, 2021). Jika pemahaman terhadap ilmu tajwid minim, seorang muslim akan kesulitan dan melakukan banyak kesalahan saat membaca Al-Quran. Kesalahan memahami ilmu tajwid dapat berakibat fatal karena dapat mengubah makna dari ayat yang dibaca. Inilah sebabnya ilmu tajwid terus dipelajari secara mandiri setiap generasi umat islam. (Alfianto dalam Syaifullah, 2021).

Pengetahuan mengenai ilmu tajwid dasar yang dibahas dalam penelitian ini adalah jenis-jenis mad, hukum nun mati, hukum mim mati, dan idgham. Jenis-jenis mad meliputi mad asli, mad badal, mad silah qasirah, mad silah thawilah, mad wajib muttasil, mad jaiz munfasil, mad farqi, mad tamkin, mad iwadh, mad aridh lissukun, mad layyin, mad lazim kilmi mukhoffaf, mad lazim kilmi mutsaqqol, mad lazim harfi mukhoffaf dan mad lazim harfi mutsaqqol. Hukum nun mati yang dibahas meliputi ikhfah, iqlab, idgham bighunnah, dan idgham bilaghunnah. Hukum mim mati yang dibahas adalah ikhfah syafawi dan idgham mitslain. Pengetahuan ini akan dimasukkan ke dalam *chatbot* sebagai pengetahuan dasar atau *knowledge base*.

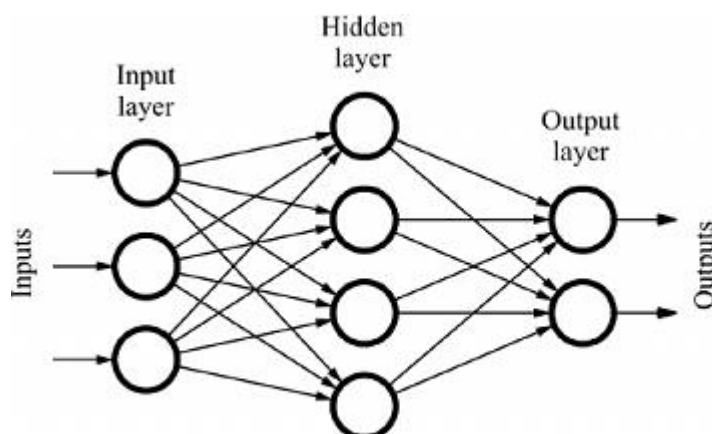
#### **2.2 Natural Language Processing (NLP)**

*Natural Language Processing* (NLP) merupakan suatu cabang dari kecerdasan buatan yang berfokus pada bagaimana memberi pemahaman pada komputer mengenai teks dan ucapan selayaknya manusia. NLP bekerja berdasarkan bagaimana manusia menggunakan bahasa yaitu belajar melalui pengalaman.

NLP adalah kombinasi antara ilmu komputasi bahasa dan ilmu lainnya seperti statistika, pembelajaran mesin, *deep learning*, dan lain sebagainya (Rayyan, 2022). NLP diterapkan dalam berbagai aplikasi, diantaranya adalah *Question Answering System* (Sistem Tanya Jawab) pada *browser*, *Machine Translation* atau mesin penerjemah bahasa, analisis sentimen terhadap sebuah topik, *Summarization* atau peringkasan teks, dan lain sebagainya.

### 2.3 Feed Forward Neural Network

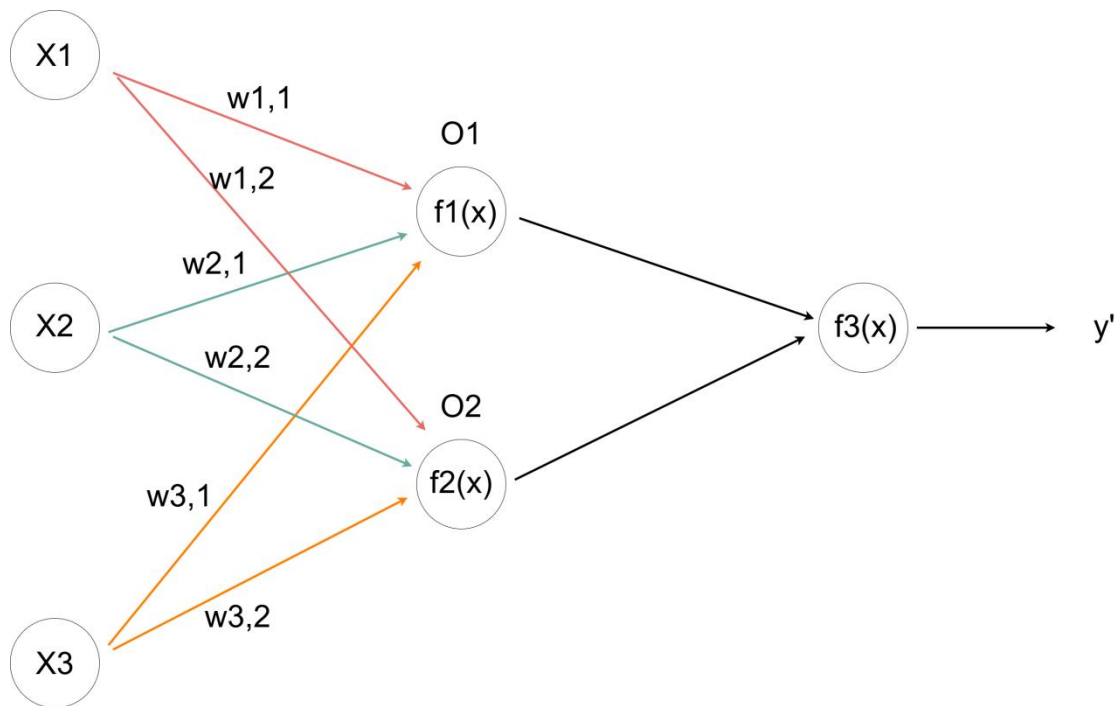
*Feed Forward Neural Network* (FFNN) adalah model jaringan saraf paling sederhana. FFNN terinspirasi dari struktur jaringan neuron manusia. FFNN berjalan satu arah dan tidak memiliki fungsi rekursif. FFNN memiliki struktur seperti berikut :



Gambar 2. 1 Arsitektur Feed Forward Neural Network

FFNN memiliki 3 layer, yaitu input layer, hidden layer dan output layer. Karena arsitektur ini memiliki sifat feed-forward, maka informasi berjalan hanya dari input searah ke hidden layer kemudian keluar ke output. Dengan alur kerja yang lurus ini, FFNN cocok untuk digunakan dalam tugas-tugas satu *one-way processing* data, seperti klasifikasi dan membuat prediksi. FFNN merupakan fondasi utama dalam neural network yang merupakan dasar dari neural network yang lebih kompleks seperti Recurrent Neural Network, Convolutional Neural Network, Transformer, dan lain sebagainya. Dua kunci utama dalam pengembangan FFNN adalah Multi Layer Perceptron dan Backpropagation.

Cara perhitungan FFNN pada dasarnya adalah merupakan perhitungan matriks. Dalam Gambar 1.2, arsitektur FFNN tersebut memiliki 3 nodes pada input layer. Ketiga nodes ini memiliki nilai masing-masing, biasa disimbolkan sebagai X1, X2 dan X3. Ketiganya mempunyai nilai masing lalu akan dikalikan dengan bobot yang searah dengan node pada hidden layer. Perhitungannya diilustrasikan seperti gambar berikut.



Gambar 2. 2 Arsitektur FFNN dengan Perhitungan Bobot

Maka jika bisa dilihat bahwa perhitungan output pada hidden layer adalah sebagai berikut :

$$O1 = f ( X1 * W1,1 + X2 * W1,2 + X3 * W1,3 )$$

$$O2 = f ( X2 * W2,1 + X2 * W2,2 + X3 * W2,3 )$$

Perkalian ini jika ditulis dalam bentuk perkalian matriks sebagai berikut :

$$\begin{bmatrix} O1 \\ O2 \end{bmatrix} = f \left( \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} \cdot \begin{bmatrix} W1,1 & W1,2 & W1,3 \\ W2,1 & W2,2 & W2,3 \end{bmatrix} \right)$$

Output 1 dan 2 merupakan hasil dari dot product antara X1, X2, dan X3 dengan bobot-bobot yang ditentukan. Begitu juga dengan perhitungan antara hidden

layer dan output. Untuk mendapatkan  $Y'$  atau prediksi adalah hasil perkalian dari matriks  $O$  dan matriks pada output layer. Ini adalah konsep dasar dari menghitung proses output pada FFNN. Setelah nodes ini menghasilkan output, maka informasi yang dibutuhkan adalah error untuk melakukan backpropagation untuk mendapat nilai yang diinginkan.

## 2.4 Chatbot

*Chatbot* atau *chatter bot* adalah sebuah program komputer yang digunakan sebagai media simulasi percakapan antara manusia dan komputer. *Chatbot* memungkinkan interaksi yang mulus antara manusia dengan *software* melalui pemrosesan bahasa alami (NLP) sehingga memberikan pengalaman seolah-olah berbicara dengan orang lain (Gaikwad, 2019). Bentuk *chatbot* adalah berupa program atau asisten digital yang mampu menjawab sebaris pertanyaan *user* dan belajar dari data dan informasi yang dikumpulkan untuk menghasilkan personalisasi bagi user (Rayyan, 2022).

Sistem *chatbot* telah ditemukan sejak tahun 1965, saat itu sistem ini disebut sistem percakapan atau *conversational system*. E.L.I.Z.A yang diciptakan oleh Joseph Weizenbaum adalah sistem percakapan pertama yang lulus tes Turing. Enam tahun kemudian PARRY ditemukan oleh Kenneth Colby, yaitu sebuah model percakapan yang disimulasikan seolah pasien berbicara dengan psikiater. A.L.I.C.E, diciptakan pada tahun 1995 oleh Wallace dalam AIML (*Artificial Intelligent Markup Language*), sebuah dialek XML (*Extensible Markup Language*) untuk membuat *chatbot*. (Gaikwad, 2019). Pada tahun-tahun berikutnya kemudian muncul *chatbot-chatbot* baru seperti SIRI dan ALEXA yang dapat merespon perintah menggunakan suara ataupun teks. Seiring dengan perkembangan ilmu pengetahuan, sistem *chatbot* berkembang tidak hanya dengan model berbasis aturan (*rule-based*), ditemukan juga sistem yang kecerdasan buatan yang terinspirasi dari cara kerja otak manusia. Sistem ini membuat *chatbot* belajar sendiri dan menghasilkan jawaban yang akurat.

## 2.5 Natural Language Toolkit (NLTK)

NLTK adalah sebuah library Python yang digunakan untuk mengolah dan menganalisis teks dalam bahasa alami. Library ini menyediakan berbagai alat dan teknik untuk tokenisasi, stemming, tagging, parsing, dan lainnya, yang sangat berguna dalam bidang linguistik komputasional dan pemrosesan bahasa alami (NLP). Dengan

NLTK, pengguna dapat melakukan berbagai tugas seperti pemecahan kalimat, pengenalan entitas bernama, analisis sintaksis, serta klasifikasi teks. NLTK juga dilengkapi dengan berbagai dataset dan korpus yang dapat digunakan untuk pelatihan model NLP, sehingga memudahkan para peneliti dan praktisi dalam mengembangkan aplikasi berbasis bahasa alami.

Selain menyediakan alat-alat dasar untuk pengolahan bahasa, NLTK juga memiliki modul-modul untuk pembelajaran mesin dan visualisasi data. Ini memungkinkan pengguna untuk membangun model-model pembelajaran mesin yang dapat belajar dari data teks dan membuat prediksi atau klasifikasi berdasarkan pola-pola yang ditemukan. Sebagai library yang open-source, NLTK memiliki komunitas yang aktif dan dokumentasi yang luas, sehingga sangat membantu bagi pemula maupun ahli dalam bidang NLP untuk memulai dan mengembangkan proyek-proyek mereka. Dengan segala kemampuannya, NLTK menjadi salah satu alat yang paling populer dan sering digunakan dalam dunia pemrosesan bahasa alami dan penelitian linguistik komputasional.

## 2.6 Pytorch

PyTorch adalah sebuah library dalam bahasa pemrograman python yang open-source yang digunakan untuk aplikasi deep learning dan machine learning, dikembangkan oleh Facebook's AI Research lab (FAIR). PyTorch terkenal karena kemudahannya dalam penggunaan dan fleksibilitasnya, yang memungkinkan para peneliti dan praktisi untuk dengan cepat mengembangkan dan menguji model-model deep learning. Salah satu fitur utama dari PyTorch adalah kemampuan tensor yang efisien dan dukungan untuk GPU, yang memungkinkan komputasi paralel yang cepat dan efisien. PyTorch juga mendukung dynamic computational graph, yang berarti grafik komputasi dapat dibangun secara dinamis dan memungkinkan debug dan modifikasi yang lebih mudah dibandingkan dengan library lain seperti TensorFlow yang menggunakan static computational graph.

Selain itu, PyTorch memiliki ekosistem yang kaya dan komunitas yang aktif. Library ini mencakup berbagai modul yang mendukung berbagai teknik machine learning, seperti `torch.nn` untuk neural networks, `torch.optim` untuk optimizers, dan `torchvision` untuk computer vision. PyTorch juga sering digunakan dalam penelitian karena kesederhanaannya dalam mendefinisikan dan melatih model, serta

kompatibilitasnya dengan berbagai alat penelitian lainnya. Keunggulan ini membuat PyTorch menjadi pilihan utama bagi banyak peneliti dan praktisi dalam bidang kecerdasan buatan dan machine learning.

## **2.7 Tokenisasi**

Tokensasi atau tokenizing adalah proses mengonversi teks menjadi unit-unit kecil yang lebih mudah diproses, yang disebut token. Token bisa berupa kata, frasa, atau simbol, tergantung pada konteks dan tujuan pemrosesan bahasa alami. Proses ini sangat penting dalam pemrosesan bahasa alami (NLP) karena memungkinkan komputer untuk memahami dan memanipulasi teks dengan cara yang lebih struktural. Tokenizing dapat dilakukan menggunakan berbagai teknik, mulai dari pemisahan berdasarkan spasi sederhana hingga penggunaan algoritma canggih yang mempertimbangkan konteks gramatikal dan semantik. Dalam prakteknya, tokenizing seringkali melibatkan penghilangan tanda baca, konversi ke huruf kecil, dan pemisahan kata berdasarkan delimiters seperti spasi, tanda baca, atau karakter khusus. Beberapa library populer untuk tokenizing termasuk NLTK (Natural Language Toolkit), spaCy, dan Tokenizer dari TensorFlow. Proses tokenisasi yang baik tidak hanya meningkatkan akurasi model NLP tetapi juga mempercepat waktu komputasi dan mempermudah pengembangan aplikasi berbasis teks. Dengan adanya alat-alat ini, para peneliti dan pengembang dapat lebih fokus pada pengembangan model dan analisis data tanpa harus khawatir tentang detail implementasi tokenisasi yang rumit.

## **2.8 Stemming**

Stemming adalah proses dalam pemrosesan bahasa alami (NLP) yang bertujuan untuk mengurangi kata-kata infleksional atau turunan ke bentuk dasarnya yang disebut stem. Stem biasanya tidak harus merupakan kata yang valid secara leksikal tetapi cukup untuk mewakili kata-kata yang terkait secara semantik. Misalnya, kata-kata seperti "running", "runner", dan "ran" akan direduksi menjadi "run" melalui proses stemming. Teknik ini penting dalam berbagai aplikasi NLP seperti mesin pencari, analisis teks, dan pengelompokan dokumen, karena membantu mengurangi redundansi dan meningkatkan akurasi dalam pemrosesan teks.

## **2.9 Bag of Words (BoW)**

BoW merupakan representasi tekstual yang digunakan dalam NLP dan teks mining. BoW mengubah teks menjadi koleksi kata-kata yang tidak terurut tanpa mempertimbangkan tata bahasa dan urutan kata. Setiap kata unik dalam teks dianggap sebagai fitur, dan nilai dari fitur tersebut adalah jumlah kemunculan kata dalam dokumen. Model BoW sederhana namun sangat efektif untuk tugas-tugas seperti klasifikasi teks, analisis sentimen, dan pengenalan topik. Kelebihan utama dari pendekatan ini adalah kesederhanaan dan kemampuannya untuk digunakan dengan berbagai algoritma pembelajaran mesin.

## **2.10 Flask**

Flask adalah sebuah micro web framework untuk Python yang dirancang untuk memudahkan pengembangan aplikasi web dengan menyediakan dasar-dasar yang dibutuhkan tanpa banyak kompleksitas. Flask dikenal karena kesederhanaannya, fleksibilitasnya, dan kemudahan dalam penggunaan, menjadikannya pilihan populer untuk proyek-proyek kecil hingga menengah. Framework ini mengikuti prinsip minimalisme dan tidak mengharuskan penggunaan alat atau pustaka tertentu, memungkinkan pengembang untuk menambahkan komponen sesuai kebutuhan mereka. Flask menawarkan routing yang mudah, template engine berbasis Jinja2, serta dukungan untuk pengujian unit, yang membuat proses pengembangan lebih efisien dan terstruktur.

## **2.11 Evaluasi Model**

Mengevaluasi model adalah langkah penting dalam menganalisis data. Evaluasi model digunakan untuk mengukur seberapa baik model bekerja. Ada beberapa metode yang bisa digunakan untuk memberikan gambaran objektif tentang kinerja model, seperti akurasi, presisi, *recall*, *specifity*, *F1-score* dan *confusion matrix*. Model-model evaluasi ini dapat disesuaikan dengan tugas spesifik yang dilakukan. Pada penelitian ini penulis menggunakan model evaluasi akurasi.



## 2.12 Penelitian Terdahulu

Untuk melengkapi pengetahuan tentang *chatbot* sebagai media tanya jawab tentang ilmu tajwid, berikut adalah tabel daftar penelitian yang pernah dilakukan tentang penerapan *neural network* dalam chatbot :

Tabel 2. 1 Penelitian Terdahulu				
No	Peneliti	Judul	Tahun	Keterangan
1	Pradana, ST	Penerapan Algoritma <i>Leveinsthein Distance</i> dalam Pembuatan <i>Chatbot</i> Informasi tentang Tajwid	2022	Aplikasi dikembangkan dengan algoritma <i>Leveinsthein Distance</i> yang dapat menghitung jarak dua string berbeda kemudian mencocokkannya dengan pertanyaan yang paling dekat. Hasil penelitian ini mendapat akurasi sebesar 100% dengan pengujian <i>dataset</i> langsung dan 92% dengan pertanyaan sembarang.
2	Purwitasari, N.A dan Soleh, M	Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Processing.	2019	Penulis membuat sistem bernama Peraturanku sebuah aplikasi yang dapat memudahkan pengguna untuk mencari informasi tentang peraturan dan perundang-undangan. Model ini mendapatkan tingkat akurasi tinggi sebesar 94,28% dan kesalahan 5,71%.
3	Fadli, M.F., Buntoro, G.A	Penerapan Algoritma	2022	Pengujian aplikasi Chatbot menggunakan data

	dan Masykur, F.	Neural Network pada Chatbot PMB Universitas Muhammadiyah Ponorogo Berbasis Web		sebanyak 251 kalimat pertanyaan tentang pendaftaran mahasiswa baru di Universitas Pakuan. Hasil pengujian menunjukkan bahwa Chatbot dapat menjawab pertanyaan mengenai pendaftaran mahasiswa baru dengan akurasi sebesar 88%, presisi sebesar 95% dan recall sebesar 92%.
4	Zuraiyah, Z.A., Utami, D.K., dan Herlambang, D.	Implementasi Chatbot Pada Pendaftaran Mahasiswa Baru Menggunakan Recurrent Neural Network	2019	Pengujian aplikasi Chatbot menggunakan data sebanyak 251 kalimat pertanyaan tentang pendaftaran mahasiswa baru di Universitas Pakuan. Hasil pengujian menunjukkan bahwa Chatbot dapat menjawab pertanyaan mengenai pendaftaran mahasiswa baru dengan akurasi sebesar 88%, presisi sebesar 95% dan recall sebesar 92%.
5	Ruben, I.	Implementasi Neural Network untuk Pembuatan Chatbot	2022	Penelitian ini membuat model pembelajaran mesin untuk chatbot serta aplikasi chatbot. Model ini

Menggunakan	mendapat akurasi sebesar
Dataset	98% dan loss value
Pertanyaan	sebesar 0,05275.
Mahasiswa	

## BAB 3

### ANALISIS DAN PERANCANGAN

#### 3.1 Persiapan Data

Dalam pembentukan dataset, data diambil dari buku panduan ilmu tajwid berjudul Panduan Tahsin Tilawah yang disusun H. Ahmad Muzammil MF, Al Hafizh dan tanya jawab dengan guru tahsin. Data dibagi menjadi data latih, data validation dan data uji. Data latih yang berekstensi json akan melakukan proses *training*, sedangkan data uji adalah data yang akan digunakan untuk menguji hasil *training* berupa model lewat *user interface*. Jumlah data yang dilatih adalah sebanyak 198 pertanyaan dan jawaban yang dibentuk dalam format json dengan hasil 35 tag dengan 198 pattern dan 35 respond. Sedangkan data uji berjumlah 20 pertanyaan acak tentang mad, hukum nun mati, dan hukum mim mati.

Tabel 3. 1 Pembagian Data Latih dan Data Uji

Data Latih	Data Uji
35 tag, 198 pattern dan 35 respon	30 pertanyaan.

Contoh pertanyaan :

Tabel 3. 2 Contoh Pertanyaan dan Jawaban

Pertanyaan	Jawaban
Apa itu mad?	Mad adalah hukum membaca panjang bacaan Al-Quran ketika ada huruf mad.
Iqlab adalah	Iqlab adalah ketika nun mati atau tanwin bertemu huruf ba. Hukum membacanya adalah ghunnah atau dengung sebanyak tiga harakat.
Hukum mim mati	Hukum mim mati ada 3 yaitu ikhfa syafawi dan idgam mtslain.

Apa saja hukum nun mati?	Hukum nun mati ada 4, yaitu Izhar, Iqlab, Idgham dan Ikhfa.
Ada berapa jenis mad?	Dalam ilmu tajwid terdapat 16 mad yang terbagi ke dalam 3 jenis berdasarkan panjangnya harakatnya.

Contoh data json :

Tabel 3. 3 Contoh Data JSON

<pre> {   "tag": "mad",   "patterns": [     "Apa itu mad?",     "Apa yang dimaksud dengan mad?",     "Definisi mad",     "Mad artinya",     "Mad adalah"   ],   "responses": [     "Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berbaris fathah diikuti dengan alif sukun, huruf berbaris kasrah diikuti ya sukun, dan huruf berbaris dhommah diikuti wau sukun."   ] }, {   "tag": "jumlah_mad",   "patterns": [     "Berapa mad yang ada dalam ilmu tajwid?",     "Ada berapa mad dalam tajwid?",     "Berapa jumlah mad?"   ],   "responses": [ </pre>
--

"Dalam ilmu tajwid terdapat 16 mad yang terbagi ke dalam 3 jenis berdasarkan panjangnya harakatnya."

]

},

{

"tag": "jenis\_mad",

"patterns": [

"Apa saja mad yang dalam tajwid?",

"Apa jenis-jenis mad dalam tajwid?",

"Jenis-jenis mad"

],

"responses": [

"Berdasarkan panjang harakatnya, mad bisa dibagi tiga yaitu panjang dua harakat, empat harakat dan enam harakat."

]

},

{

"tag": "mad\_dua\_harakat",

"patterns": [

"Kapan mad dibaca 2 harakat?",

"Apa saja mad yang dibaca 2 harakat?",

"Mad dua harakat",

"Mad yang panjangnya dua harakat"

],

"responses": [

"Mad harus dibaca dua harakat dan tidak boleh lebih ketika ada bacaan mad yang sesudahnya bukan huruf hamzah, tasydid ataupun huruf mati. Mad yang panjangnya dua harakat adalah Mad Thabi'i, Mad Badal, Mad Tamkin, Mad 'Iwadh, Mad Harfi, Mad Lazim Harfi Mutsaqqal dan Mad Shilah Qashirah."

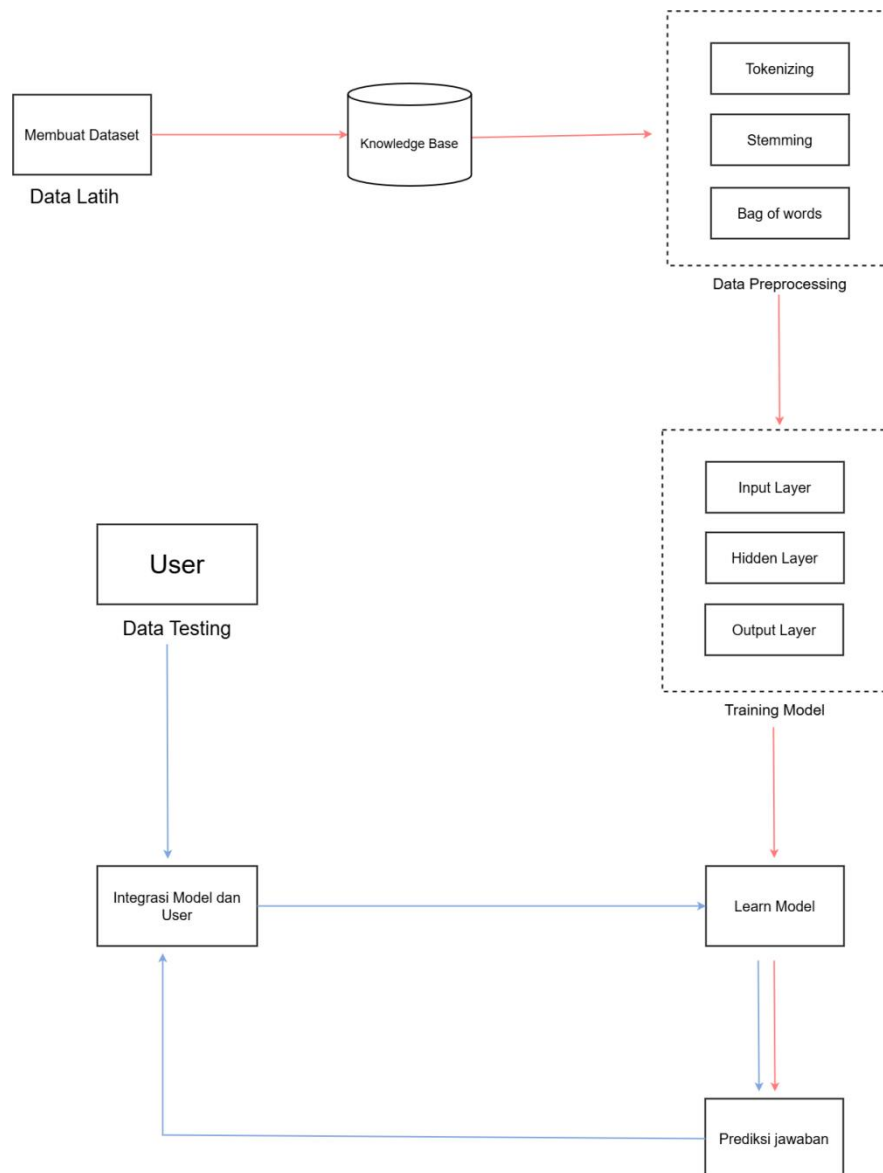
]

},

### 3.2 Perancangan sistem

#### 3.2.1 Arsitektur Umum

Berikut adalah arsitektur umum penerapan metode *neural network* pada model *chatbot* tanya jawab ilmu tajwid. Garis merah menunjukkan alur data latih dan garis biru menunjukkan alur data testing.



Gambar 3. 1 Arsitektur Umum

#### A. Membuat Dataset

Data latih diambil dari buku Panduan Tahsin dan Tilawah karya H. Ahmad Muzammil MF, Al Hafizh dan tanya jawab dengan seorang guru tahsin. Kemudian data tersebut dikumpulkan dalam bentuk tanya jawab. Data ini yang akan menjadi *knowledge base*

atau basis data dalam pengembangan chatbot ini. Kemudian data ini dikembangkan menjadi data yang berekstensi json (*Javascript Object Notation*) untuk digunakan dalam proses pelatihan.

## B. Preprocessing Data

Selanjutnya data di preprocessing untuk mempersiapkan data mentah (raw data) sebelum dimasukkan ke dalam model machine learning. Proses preprocessing ini bertujuan untuk mengoptimalkan kualitas data sehingga model dapat belajar dengan lebih efektif dan menghasilkan prediksi yang lebih baik.

### 1. Tokenisasi

Tokenisasi berfungsi untuk melakukan pemecahan terhadap teks menjadi unit-unit kecil yang disebut token. Tujuan utama dari tokenisasi adalah untuk menyederhanakan teks sehingga model dapat lebih mudah memprosesnya.

Tabel 3. 4 Contoh penerapan tokenisasi

Teks = Saya sedang makan karena saya lapar.

Proses tokenizing akan memecah teks menjadi kata-kata: ['Saya', 'sedang', 'makan', 'karena', 'saya', 'lapar', '.']

### 2. Stemming

Stemming adalah proses dalam pemrosesan teks yang bertujuan untuk menghilangkan afiks dari kata-kata agar hanya menyisakan bentuk dasar atau akar kata (stem). Tujuan utama dari stemming adalah untuk mengurangi kata-kata ke bentuk dasarnya sehingga kata-kata dengan akar yang sama dapat dianggap sebagai satu entitas. Proses ini membantu dalam mengurangi variasi kata yang serupa secara makna namun memiliki bentuk yang berbeda.

Tabel 3. 5 Contoh penerapan stemming

Teks = "saya sedang makan karena saya lapar"



Proses stemming akan membuat teks menjadi : ['saya', 'sedang', 'makan', 'karena', 'saya', 'lapar']

### 3. Bag of words

Hasil stemming diteruskan ke Bag of Words yaitu model representasi teks yang digunakan dalam pemrosesan bahasa alami (NLP) dan teks mining. Konsep dasar dalam *Bag of Words* adalah teks (misalnya, dokumen atau kalimat) diwakili sebagai tas (bag) dari kata-kata, tanpa memperhatikan urutan kata. *Bag of Words* akan memperhatikan frekuensi kemunculan kata dalam teks. Dalam *Bag of Words* terjadi proses encoding yaitu mengubah input teks menjadi representasi vektor. Setiap elemen dalam vektor menunjukkan frekuensi kemunculan kata tersebut dalam dokumen. Misalnya dalam kalimat : saya sedang makan karena saya lapar. Maka *Bag of Words* akan merepresentasikannya sebagai berikut.

Tabel 3. 6 Contoh encoding Bag of words

Kata	Bow vector					
Saya	1	0	0	0	0	0
sedang	0	1	0	0	0	0
makan	0	0	1	0	0	0
karena	0	0	0	1	0	0
saya	0	0	0	0	1	0
lapar	0	0	0	0	0	1

Setiap baris mewakili kalimat asli, dan setiap kolom mewakili kata-kata unik dalam semua kalimat. Nilai di dalam sel tabel menunjukkan berapa kali kata tersebut muncul dalam kalimat yang sesuai. Jika direpresentasikan dalam vektor, maka kalimat “saya sedang makan karena saya lapar” direpresentasikan sebagai berikut.

Tabel 3. 7 Contoh encoding Bag of words

saya	2
sedang	1

makan	1
karena	1
lapar	1

### C. Implementasi Model *Feed Forward Neural Network*

Training dataset tajwid menggunakan *Feed Forward Neural Network* (FFNN). FFNN adalah jenis jaringan saraf tiruan yang paling sederhana dan paling dasar. Nama "*feed forward*" berasal dari cara sinyal bergerak melalui jaringan: dari input ke output tanpa adanya *loop* atau *feedback*. Komponen utama FFNN ada 3 yaitu *input layers*, *hidden layers* dan *output layers*. Cara kerja FFNN yang pertama adalah *forward propagation*, yaitu proses yang melibatkan mengirimkan data input melalui jaringan dari lapisan input menuju lapisan output.

#### 1. *Input Layers*

*Input layers* yaitu layer pertama yang menerima input eksternal. Setiap neuron dalam lapisan ini mewakili satu fitur dari data input. Input ini adalah hasil *encoding* dari *bag of words* yang memiliki nilai masing-masing.

#### 2. *Hidden Layers*

*Hidden layers* merupakan lapisan yang terletak antara lapisan input dan lapisan output. FFNN bisa memiliki satu atau lebih hidden layers. Setiap neuron di *hidden layer* menerima input dari neuron di lapisan sebelumnya yaitu neuron input. Setiap neuron pada *hidden layer* memiliki bobot yang terkait dengan setiap input. Kemudian dilakukan perkalian antara nilai input dan bobot untuk menghasilkan nilai pada setiap neuron pada *hidden layer* dan penambahan bias (jika ada).

#### 3. *Output Layers*

*Output layers* yaitu lapisan ini menghasilkan output akhir dari jaringan. Proses perhitungan pada *hidden layer* diulang untuk setiap layer tersembunyi sampai output layer dicapai. Pada *output layer* ini terdapat fungsi aktivasi khusus yaitu fungsi ReLu. Output dari output layer ini merupakan hasil akhir dari *feed forward neural network*.

Tabel 3. 8 Pemodelan Neural Network dengan Modul NeuralNet

```

import torch.nn as nn

class NeuralNet(nn.Module):
    def __init__(self, input_size, hidden_size,
num_classes):
        super(NeuralNet, self).__init__()
        self.l1 = nn.Linear(input_size, hidden_size)
        self.l2 = nn.Linear(hidden_size, hidden_size)
        self.l3 = nn.Linear(hidden_size, num_classes)
        self.relu = nn.ReLU()

    def forward(self, x):
        out = self.relu(self.l1(x)) # Layer pertama
        out = self.relu(self.l2(out)) # Layer kedua
        out = self.l3(out) # Layer ketiga (output
layer), tanpa aktivasi
        return out

```

Torch adalah library utama yang digunakan untuk komputasi tensor dan jaringan saraf di PyTorch, sedangkan torch.nn menyediakan berbagai modul dan fungsi yang digunakan untuk membangun arsitektur jaringan saraf. self.l1, self.l2, dan self.l3 merupakan layer-layer linearnya :

- self.l1 = nn.Linear(input\_size, hidden\_size) mendefinisikan layer pertama yang mengubah input berukuran input\_size menjadi output berukuran hidden\_size.
- self.l2 = nn.Linear(hidden\_size, hidden\_size) mendefinisikan layer kedua yang mengubah input berukuran hidden\_size menjadi output berukuran hidden\_size.
- self.l3 = nn.Linear(hidden\_size, num\_classes) mendefinisikan layer ketiga yang mengubah input berukuran hidden\_size menjadi output berukuran num\_classes, sesuai dengan jumlah kelas yang ingin diprediksi.

- `self.relu = nn.ReLU()` adalah fungsi aktivasi ReLU yang digunakan untuk menambahkan non-linearitas ke dalam model. Fungsi ini diterapkan setelah setiap layer untuk memungkinkan model untuk mempelajari pola yang lebih kompleks.

Metode forward adalah metode yang mendefinisikan bagaimana data mengalir melalui jaringan saraf saat melakukan proses forward propagation dengan `x` sebagai input yang diterima oleh model.

- `out = self.relu(self.l1(x))` menghitung output dari layer pertama dengan mengalikan input `x` dengan bobot dari `self.l1`, kemudian menerapkan fungsi aktivasi ReLU.
- `out = self.relu(self.l2(out))` menghitung output dari layer kedua dengan mengalikan output dari layer pertama (`out`) dengan bobot dari `self.l2`, kemudian menerapkan fungsi aktivasi ReLU lagi.
- `out = self.l3(out)` menghitung output dari layer ketiga (output layer) dengan mengalikan output dari layer kedua (`out`) dengan bobot dari `self.l3`. Pada layer output, tidak diterapkan fungsi aktivasi tambahan karena dalam hal ini kita ingin mendapatkan nilai-nilai prediksi langsung.

#### 4. Fungsi Aktivasi ReLU

*Rectified Linear Unit* (ReLU) merupakan salah satu fungsi aktivasi yang berfungsi dalam meningkatkan percepatan dari proses training. Dibandingkan dari fungsi-fungsi lainnya seperti sigmoid/tan h, ReLU memiliki kelebihan untuk mempercepat komputasi karena tidak menggunakan operasi eksponensial. Fungsi ReLU juga menghilangkan *vanishing gradient* karena memiliki gradien konstan dari 1 untuk nilai positif. Ini memungkinkan pelatihan yang lebih lancar dan lebih cepat pada jaringan yang dalam.

#### 5. Cross Entropy Loss

*Cross Entropy Loss* adalah fungsi kerugian (*loss function*) yang umum digunakan untuk masalah klasifikasi multi-kelas dalam jaringan saraf. *Cross Entropy Loss* biasanya digunakan dalam kasus di mana output model adalah probabilitas di layer output. *Cross Entropy Loss* tidak memerlukan ekspresi softmax di layer output karena fungsi ini sudah mencakup operasi softmax secara bersamaan.

```
criterion = nn.CrossEntropyLoss()
```

#### 6. *Optimizer ADAM*

ADAM (*Adaptive Moment Estimation*) adalah algoritma optimisasi yang dapat digunakan untuk memperbarui bobot secara iteratif yang didasarkan pada data training. ADAM menggunakan estimasi gradien momen pertama dan kedua untuk mengadaptasi learning rate untuk setiap bobot jaringan saraf. *Hyperparameter* yang digunakan untuk menentukan optimasi adalah *learning rate*. *Learning rate* merupakan ukuran langkah yang diambil *optimizer* saat memperbarui parameter model. Nilai yang lebih kecil seperti membuat pembelajaran lebih lambat tapi stabil, sementara nilai yang lebih besar dapat membuat pembelajaran lebih cepat tapi berpotensi tidak stabil.

```
optimizer = torch.optim.Adam(model.parameters(),  
lr=learning_rate)
```

#### 7. Penentuan *Learning Rate*

*Learning rate* merupakan parameter yang menentukan laju pembelajaran algoritma yang digunakan. Nilai *learning rate* yang digunakan adalah 0.001 yang merupakan nilai default dari *optimizer* ADAM. Nilai yang kecil ini membuat pembelajaran lebih lambat tapi stabil.

```
learning_rate = 0.001
```

#### 8. Penentuan *Batch Size*

*Batch size* adalah penentuan pembagian jumlah data yang dilatih pada satu kali pelatihan iterasi (pengulangan). Contohnya pada 1000 data training, jika batch size diatur pada angka 100, maka algoritma akan mengambil 100 data pada setiap iterasi untuk di training. Proses ini membantu memperkecil memori yang digunakan saat *training* sehingga proses *training* menjadi lebih cepat. Batch size yang dipakai dalam penelitian ini adalah 8.

#### D. Proses *Training*

Setelah melalui data preparation, data mengalami preprocessing yaitu tokenisasi dan stemming. Token hasil tokenisasi kemudian diubah menjadi representasi *bag of words* menggunakan fungsi `bag_of_words` dari `nlTK_utils`. Representasi ini berbentuk vektor biner yang menunjukkan frekuensi kehadiran kata-kata dalam kalimat. Output stemming akan dimasukkan sebagai input pada *bag of words*. Selanjutnya *bag of words* akan menghitung frekuensi kata-kata yang muncul pada setiap input. Setiap kata yang muncul akan diberi nilai 1, jika tidak akan diberi nilai 0. Kemudian data ini akan dibagi menjadi 2 yaitu `x_train` dan `y_train`.

Selanjutnya adalah menginisiasi hyperparameter yang digunakan dalam training.

```
num_epochs = 1000
batch_size = 8
learning_rate = 0.001
input_size = len(X_train[0])
hidden_size = 8
output_size = len(tags)
```

Epoch merupakan hyperparameter yang menentukan berapa kali algoritma neural network bekerja melewati seluruh dataset baik secara forward maupun backward. Batch size adalah jumlah training sample yang digunakan dalam satu batch untuk sekali iterasi. Epoch yang digunakan dalam training adalah 1000 dengan batch size 8. Apabila terdapat 198 patterns, artinya dalam sekali iterasi terdapat 8 data yang ditraining dalam satu batch. Untuk mencapai satu epoch, dibutuhkan sebanyak 22 batch. Ini dilakukan berulang sehingga 1000 epoch dan mencapai final loss sebesar 0,00. Learning rate merupakan ukuran langkah yang diambil optimizer saat memperbarui parameter model. Nilai yang lebih kecil seperti 0.001 membuat pembelajaran lebih lambat tapi stabil, sementara nilai yang lebih besar dapat membuat pembelajaran lebih cepat tapi berpotensi tidak stabil.

Adapun `input_size` merupakan jumlah fitur dalam setiap sampel input. Ini diatur ke panjang dari `X_train[0]`, yang merupakan panjang dari vektor *bag of words* untuk setiap kalimat. Selanjutnya *hidden\_size*, yaitu jumlah neuron dalam layer

tersembunyi dari model neural network. Dalam contoh ini, ada 8 neuron di layer tersembunyi. Terakhir *output\_size*, yaitu jumlah kelas output yang harus diprediksi oleh model. Ini diatur ke jumlah tag unik dalam dataset.

Selanjutnya adalah membuat class dataset dan data loader.

```
dataset = ChatDataset()
train_loader = DataLoader(dataset=dataset,
                           batch_size=batch_size,
                           shuffle=True,
                           num_workers=0)
```

Dataset berfungsi untuk membuat class ChatDataset untuk memfasilitasi data input dan label. Sedangkan train\_loader untuk membuat DataLoader untuk mengelola batching dan shuffling data selama pelatihan. Selanjutnya, dilakukan pembuatan model beserta fungsi loss dan optimizer dengan NeuralNet untuk mendapatkan output logits.

```
(venv) root@chatbot-tahsin:~/chatbot-deployment# python3 train.py
137 patterns
25 tags: ["Mad thabi'i", 'goodbye', 'greeting', 'jenis_mad', 'jumlah_mad', 'mad', 'mad_aridh_lissukun', 'mad_badal', 'ma
d_dua_harakat', 'mad_empat_harakat', 'mad_enam_harakat', 'mad_farqi', 'mad_iwadh', 'mad_jaiz_munfashil', 'm
ad_layyin', 'mad_lazim_harfi_mukhoffaf', 'mad_lazim_harfi_mutsaqqol', 'mad_lazim_kilmi_mukhoffaf', 'mad_lazim_kilmi_muts
aqqol', 'mad_silah_qasirah', 'mad_silah_thowilah', 'mad_tamkin', 'mad_wajib_muttasil', 'salam']
77 unique stemmed words: ['', '2', '4', '6', 'ada', 'adalah', 'apa', 'arid', 'aridh', 'artinya', 'asli', 'assalamualaik
um', 'badal', 'bantu', 'berapa', 'bisa', 'dadah', 'dalam', 'definisi', 'dengan', 'dibaca', 'dimaksud', 'disebut', 'dua',
'empat', 'enam', 'farqi', 'hai', 'halo', 'harakat', 'harfi', 'i', 'ilmu', 'itu', 'iwadh', 'jaiz', 'jenis-jeni', 'jumlah
', 'jumpa', 'kalimi', 'kapan', 'killmi', 'kilmi', 'layyin', 'lazim', 'lein', 'lin', 'lissukun', 'mad', 'malam', 'mukhoff
af', 'munfashil', 'munfasil', 'musaqgol', 'mutsaqq', 'mutsaqqol', 'muttasil', 'pagi', 'panjangnya', 'pengertian', 'qasir
ah', 'saja', 'sampai', 'saya', 'selamat', 'siang', 'silah', 'tajwid', 'tamkin', 'thabi', 'thabii', 'thawilah', 'thowilah
', 'tinggal', 'towilah', 'wajib', 'yang']
77 25
Epoch [100/1000], Loss: 0.1231
Epoch [200/1000], Loss: 0.3306
Epoch [300/1000], Loss: 0.0022
Epoch [400/1000], Loss: 0.0037
Epoch [500/1000], Loss: 0.0000
Epoch [600/1000], Loss: 0.0001
Epoch [700/1000], Loss: 0.0001
Epoch [800/1000], Loss: 0.0000
Epoch [900/1000], Loss: 0.0000
Epoch [1000/1000], Loss: 0.0000
final loss: 0.0000
training complete. file saved to data.pth
(venv) root@chatbot-tahsin:~/chatbot-deployment#
```

Gambar 3. 2 Proses Training

Saat user mengirimkan input melalui antarmuka, input akan mengalami preprocessing berupa tokenisasi, stemming dan bag of words. Vektor bag of words diproses oleh model neural network yang telah dilatih sebelumnya (NeuralNet) untuk mendapatkan

output logits. Output logits digunakan untuk mengidentifikasi tag yang paling mungkin sesuai dengan input pengguna. Ini dilakukan dengan mengambil indeks nilai maksimum dari output model. Output logits kemudian diubah menjadi probabilitas menggunakan fungsi softmax. Probabilitas ini digunakan untuk menentukan model dalam memprediksi jawaban.

```
prob = probs[0][predicted.item()]
```

Dalam pemilihan respon, jika probabilitas prediksi lebih dari 0.75, chatbot memilih respon acak dari daftar respon yang sesuai dengan tag prediksi di intents.json. Jika tidak, chatbot memberikan respon default yang mengarahkan pengguna untuk memastikan pertanyaan mereka sesuai dengan domain yang diharapkan. Dengan penggunaan probabilitas untuk menentukan tanggapan yang paling relevan, chatbot dapat memberikan tanggapan yang lebih akurat dan kontekstual dibandingkan dengan metode pencocokan string sederhana.

#### E. Save Model

Bagian ini berisi perintah untuk menyimpan model yang sudah dilatih. Model disimpan dalam file bernama data.pth dan akan dipanggil saat pengguna melakukan interaksi dengan *chatbot* untuk melakukan prediksi jawaban.

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

#### F. Data testing

Data testing diuji langsung melalui user interface. Sistem diuji dengan pertanyaan yang berkaitan dengan mad, hukum nun mati, hukum mim mati dan idgham. Data yang diuji dapat berupa pertanyaan dan pernyataan. Data ini digunakan untuk menguji model yang sudah dilatih apakah model dapat merespon pertanyaan user dengan jawaban yang benar.

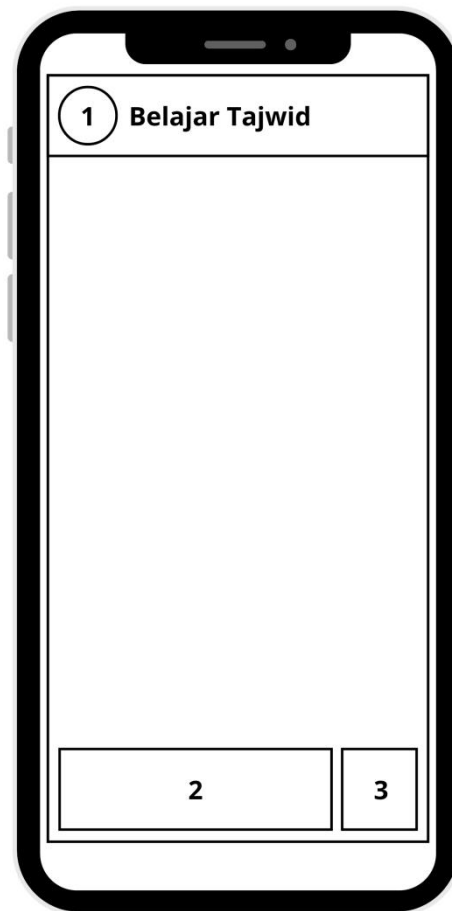


### G. Evaluasi Model

Tahap ini adalah tahap untuk menghitung kinerja model dalam format Confusion Matriks. Mengevaluasi chatbot dengan menggunakan confusion matrix memiliki beberapa alasan yang kuat, terutama karena chatbot sering kali diimplementasikan sebagai model klasifikasi dalam konteks pemrosesan bahasa alami. Confusion matrix memberikan gambaran yang jelas tentang performa chatbot dalam mengklasifikasikan input dari pengguna ke dalam berbagai kategori atau label. Ini mencakup jumlah prediksi yang benar dan salah untuk setiap kategori yang relevan. Confusion matrix dapat menghitung berbagai metrik evaluasi kinerja seperti akurasi (accuracy), precision, recall, dan F1-score. Metrik-metrik ini memberikan pemahaman yang mendalam tentang seberapa baik chatbot dalam menghasilkan respons yang tepat dan relevan.

### 3.3 Antarmuka Sistem

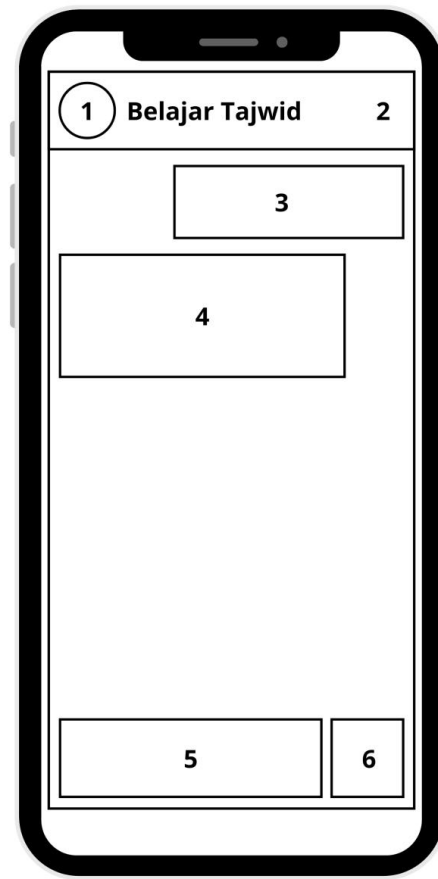
Rancangan antarmuka merupakan rancangan alat yang digunakan untuk menjalankan sistem yang akan dibangun dan menguji model yang digunakan. Antarmuka yang digunakan web. Tampilan yang dihasilkan dalam penggunaannya adalah sebagai berikut.



Gambar 3. 3 Rancangan tampilan awal chatbot

Keterangan :

1. Foto profil chatbot dan nama chatbot.
2. Kotak untuk melakukan percakapan atau mengirim pertanyaan.
3. Tombol kirim.



Gambar 3. 4 rancangan tampilan halaman chat saat melakukan percakapan

Keterangan :

4. Foto profil chatbot.
5. Nama chatbot.
6. Kotak chat pertanyaan : merupakan pertanyaan dari *user*.
7. Kotak chat jawaban : merupakan jawaban dari *chatbot*.
8. Kotak pesan : untuk menulis chat ke chatbot.
9. Tombol kirim : untuk mengirim pesan.

## **BAB 4**

### **IMPLEMENTASI DAN PEMBAHASAN**

Pada bab ini akan dijelaskan tentang proses penerapan *neural network* pada *chatbot* tanya jawab ilmu tajwid pada sistem sesuai dengan perancangan sistem yang telah dilakukan pada Bab 3 beserta pengujiannya.

#### **4.1 Implementasi Sistem**

Sesuai dengan implementasi yang telah dilakukan pada Bab 3, sistem diimplementasikan menggunakan bahasa pemrograman *python*.

##### **4.1.1 Spesifikasi, *hardware* dan *software***

Spesifikasi perangkat lunak yang digunakan untuk membangun sistem ini adalah sebagai berikut:

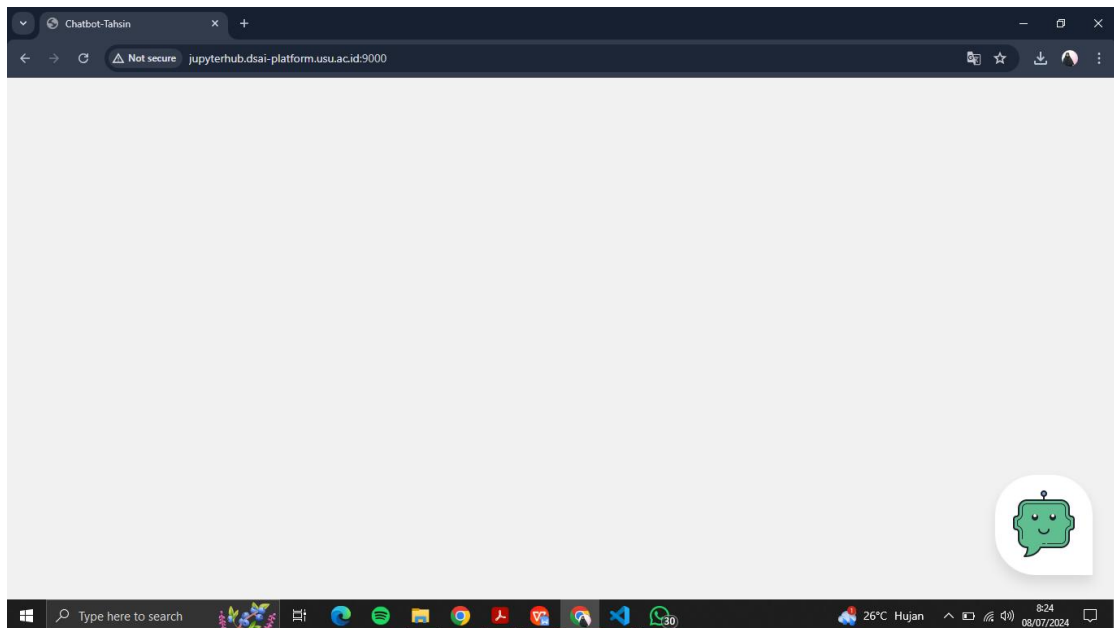
1. Processor Intel® Core™ i5-8250U CPU 1.6 1.8 GHz.
2. Memory (RAM): 12.00 GB DDR3.
3. Kapasitas hardisk 500 TB.
4. Sistem Operasi yang digunakan adalah Microsoft Windows 10 64 bit.
6. Microsoft Visual Studio Code
7. Pytorch

#### **4.2 Implementasi Perancangan Antarmuka**

Implementasi perancangan antarmuka berdasarkan perancangan antarmuka sistem yang telah dibahas pada Bab 3 adalah sebagai berikut:

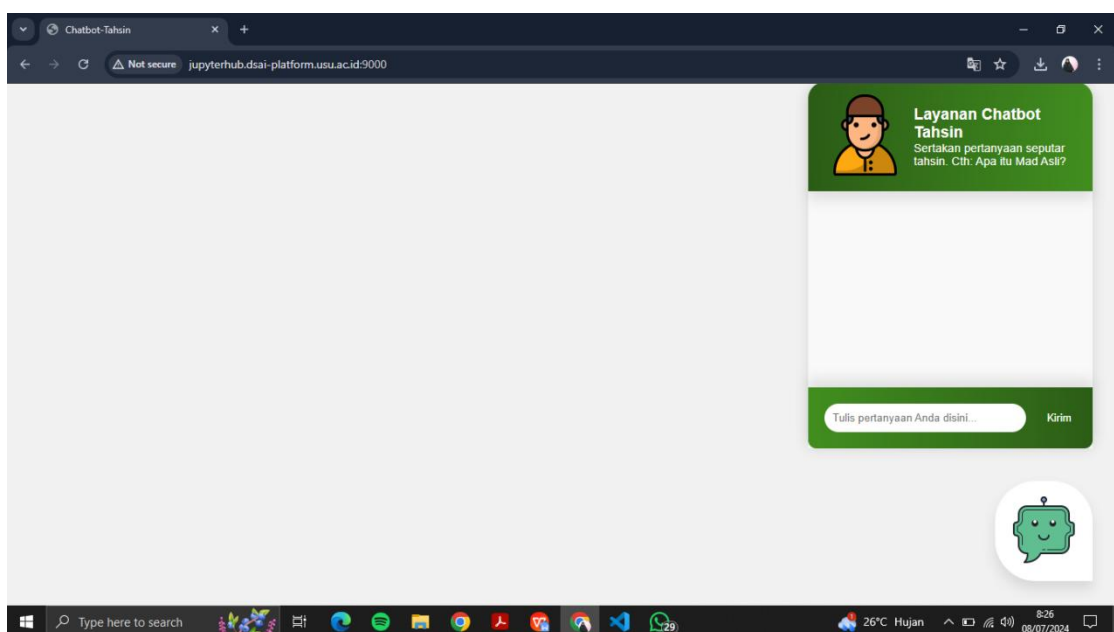
##### **a. Tampilan Chatbot**

Tampilan chatbot adalah halaman antarmuka yang siap untuk menjawab pertanyaan pengguna seputar tajwid terkhusus tentang mad. Pada halaman ini pengguna langsung bisa melakukan tanya jawab dengan chatbot dengan mengklik buble chat yang ada disebelah kiri.



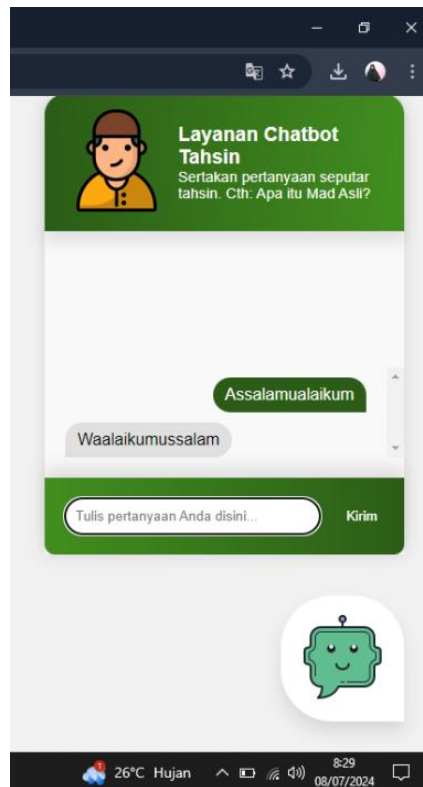
Gambar 4 1 Tampilan halaman awal chatbot

Setelah bubble chat diklik maka akan muncul kotak percakapan dengan chatbot.

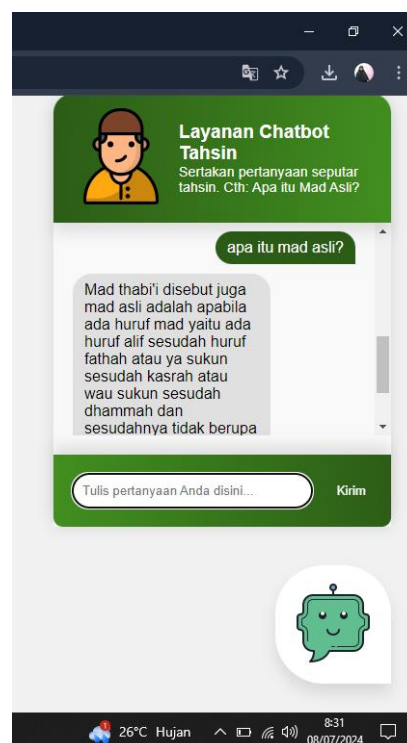


Gambar 4 2 Tampilan kotak percakapan chatbot.

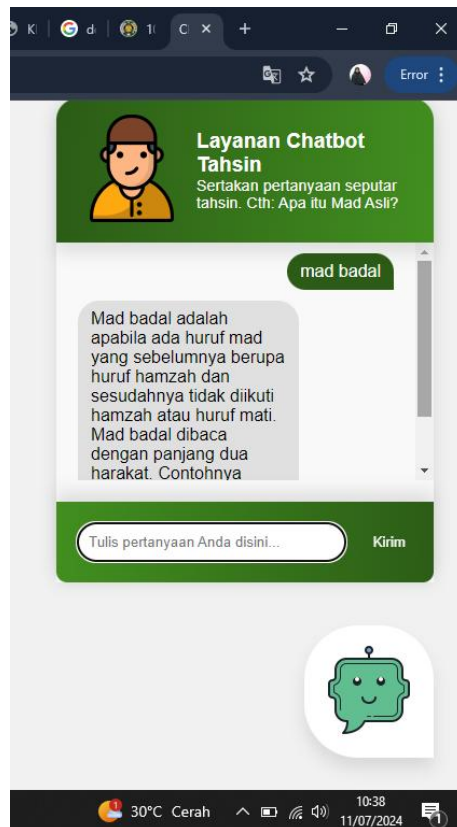
Setelah itu pengguna bisa langsung bertanya mengenai ilmu tajwid. Pada saat pembangunan chatbot ini, pengetahuan terbatas hanya pada seputar mad dalam tajwid. Berikut contoh pertanyaan tentang mad badal.



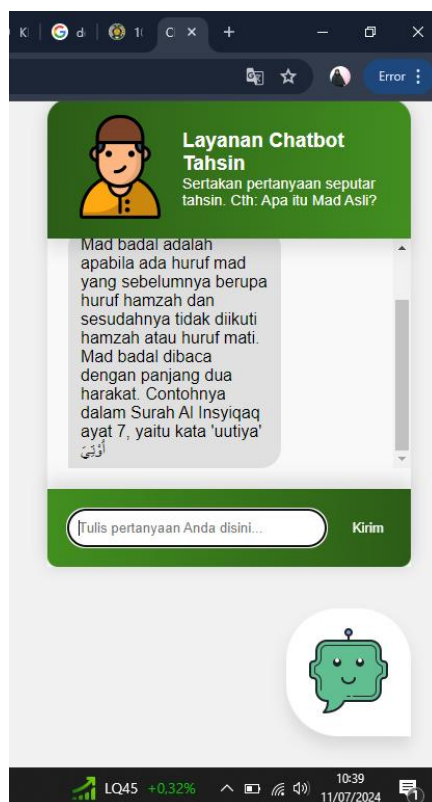
Gambar 4 3 Tampilan saat salam pembuka chatbot.



Gambar 4 4 Pertanyaan mengenai mad.



Gambar 4 5 Pertanyaan mengenai mad badal



Gambar 4 6 respon terhadap pertanyaan mengenai mad badal

### 4.3 Pengujian Sistem

Berikut adalah contoh pertanyaan yang sudah diuji pada chatbot. Data uji yang ditanyakan pada chatbot berjumlah 30 pertanyaan.

Tabel 4. 1 Contoh data uji

Pertanyaan	Jawaban Chatbot	Ketepatan antara pertanyaan dan jawaban
Apa yang dimaksud dengan mad asli?	Mad thabi'i disebut juga mad asli adalah apabila ada huruf mad yaitu ada huruf alif sesudah huruf fathah atau ya sukun sesudah kasrah atau wau sukun sesudah dhammah dan sesudahnya tidak berupa hamzah atau huruf mati atau huruf yang ditasydid. Mad Thabi'i dibaca dengan panjang 2 harakat. Mad asli sangat banyak contohnya, seperti suku kata maan pada rahmaan dan hiim pada hiim pada kalimat بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ	cocok
Ikhfa adalah	Ikhfa adalah kondisi apabila nun mati atau tanwin bertemu dengan lima belas huruf ikhfa. Ikhfa ini juga disebut Ikhfa haqiqi. Lima belas huruf ikhfa adalah ta, tsa, jim,	cocok



	dal, dzal, za, sin, syin, sho, dho, tho, dzo, fa, qof, kaf. Nun mati atau tanwin disamarkan pada huruf setelahnya sebanyak tiga harakat disertai ghunnah.	
Iqlab	Iqlab adalah kondisi apabila nun mati atau tanwin bertemu dengan huruf Ba. Cara membacanya adalah menukar nun mati atau tanwin menjadi mim disertai ghunnah dan merapatkan bibir selama 3 harakat. Iqlab dalam Al-Quran biasanya ditandai dengan huruf mim kecil berdiri.	cocok
Mad jaiz mutsaqqol	Berdasarkan panjang harakatnya, mad bisa dibagi tiga yaitu panjang dua harakat, empat harakat dan enam harakat.	Tidak cocok
Jenis-jenis mad	Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berharis fathah diikuti dengan alif sukun,	Tidak cocok

	huruf berbaris kasrah diikuti ya sukun, dan huruf berbaris dhommah diikuti wau sukun.	
--	---	--

#### 4.5 Evaluasi Model

Mengevaluasi model adalah langkah penting dalam menganalisis data. Evaluasi model digunakan untuk mengukur seberapa baik model bekerja. *Confusion matriks* merupakan salah satu model evaluasi yang digunakan dalam *Question Answering*. *Confusion matrix* menunjukkan perbandingan antara prediksi model dan nilai sebenarnya dalam bentuk matriks. Matriks ini membantu untuk mengidentifikasi seberapa baik model tersebut memprediksi setiap kelas dan di mana kesalahan prediksi terjadi. Confusion matrix terdiri dari empat komponen utama:

- True Positive (TP): Jumlah data yang sebenarnya positif dan diprediksi positif oleh model.
- True Negative (TN): Jumlah data yang sebenarnya negatif dan diprediksi negatif oleh model.
- False Positive (FP) (Type I Error): Jumlah data yang sebenarnya negatif tetapi diprediksi positif oleh model.
- False Negative (FN) (Type II Error): Jumlah data yang sebenarnya positif tetapi diprediksi negatif oleh model.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 4 7 Tabel Confusion Matrix

Dengan tabel ini, akurasi, presisi, sensitivitas dan F1-Score dapat dihitung. Akurasi adalah jumlah berapa banyak model membuat prediksi dengan benar. Akurasi dapat dihitung dengan :

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Presisi merupakan metrik evaluasi yang mengukur seberapa baik model membuat prediksi yang benar untuk kelas positif dari total prediksi positif yang dilakukan. Presisi dirumuskan dengan :

$$precision = \frac{TP}{TP + FP}$$

Sedangkan sensitivitas (recall) merupakan metrik evaluasi yang menggambarkan seberapa baik suatu model dalam mengidentifikasi kelas positif dengan benar. Sensitivitas dirumuskan dengan :

$$recall = \frac{TP}{TP + FN}$$

F-1 *score* merupakan metrik evaluasi yang mencerminkan keseimbangan antara presisi (precision) dan sensitivitas (recall). Maka dari itu rumus F1-Score adalah sebagai berikut.

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Recall + Precision}$$

Pada pengujian sistem ini, dilakukan tanya jawab dengan chatbot sebanyak 30 pertanyaan data uji. Chatbot berhasil menjawab 18 pertanyaan True Positive, 6 pertanyaan True Negative, 2 pertanyaan False Positive dan 4 pertanyaan False Negative.

Hasil akurasi yang didapat chatbot adalah :

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} = \frac{18 + 6}{18 + 2 + 4 + 6}$$

$$accuracy = \frac{24}{30} = 0.8$$

Hasil presisi yang didapat chatbot adalah :

$$precision = \frac{TP}{TP + FP} = \frac{18}{18 + 2} = \frac{18}{20} = 0.9$$

Hasil sensitivitas yang didapat chatbot adalah :

$$recall = \frac{TP}{TP + FN} = \frac{18}{18 + 4} = \frac{18}{22} = 0.81$$

Maka dari itu F1-score yang didapat adalah :

$$\begin{aligned} F1 - Score &= 2 \times \frac{Recall \times Precision}{Recall + Precision} = 2 \times \frac{0.9 \times 0.81}{0.9 + 0.81} \\ &= 2 \times \frac{0.729}{1.71} = 2 \times 0.42 = 0.84 \end{aligned}$$

## **BAB 5**

### **KESIMPULAN DAN SARAN**

Bab ini membahas tentang kesimpulan yang diperoleh dari implementasi penerapan *neural network* pada chatbot tanya jawab ilmu tajwid pada sesuai dengan perancangan sistem yang telah dilakukan dan saran-saran yang dapat digunakan sebagai referensi untuk pengembangan penelitian selanjutnya.

#### **5.1. Kesimpulan**

Dari hasil pengujian sistem pada Bab 4, diperoleh kesimpulan sebagai berikut:

2. Model *Question Answering* yang dilatih dengan FFNN dapat menjawab pertanyaan pengguna mengenai tajwid dengan akurasi sebesar 0.8, presisi sebesar 0.9, sensitivitas sebesar 0.81 dan F1-score sebesar 0.84.
3. Jumlah epoch memengaruhi proses training. Semakin tinggi nilai epoch yang digunakan saat training maka nilai loss yang dihasilkan semakin rendah. Namun harus diperhatikan dan dievaluasi karena dapat mengakibatkan overfitting.

#### **5.2. Saran**

Saran yang dapat diberikan penulis untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Melakukan analisis lebih dalam terhadap membangun dataset ilmu tajwid dengan metode pembelajaran mesin.
2. Menambahkan tahap *preprocessing* data seperti *filtering*, *stopword removal*, dan lain sebagainya.
3. Mengimplementasikan metode yang lebih baik sehingga percakapan dengan chatbot lebih natural.
4. Melengkapi topik tentang ilmu tajwid ke dalam model.

## DAFTAR PUSTAKA

- Diwandari, S. & Wibowo, A.P., 2022. Modul Praktikum Pemrosesan Teks. (*Online*). Universitas Teknologi Yogyakarta. <http://eprints.uty.ac.id/12402/> (29 April 2024).
- Fajri Ramadhani, N. (2023). Bimbingan Kemampuan Membaca Al-Qur'an melalui Ilmu Tajwid pada Santri Pengajian Gampong Merduati. *Jurnal Riset Dan Pengabdian Masyarakat*, 3(1), 1-6. <https://doi.org/10.22373/jrpm.v3i1.1037> (29 April 2024).
- Fadli, M.F., Buntoro, G.A dan Masykur, F. 2022. Penerapan Algoritma Neural Network pada Chatbot PMB Universitas Muhammadiyah Ponorogo Berbasis Web. Vol. 6 No. 1 (2022): *Jurnal Sistem dan Teknologi Informasi Komunikasi*. <https://doi.org/10.32524/jusitik.v6i1.786>. (7 Juli 2024).
- Faurina, R., Gazali, M.J., dan Herani, I.D.A. 2023. Implementasi Deep Feed-Forward Neural Network pada Perancangan Chatbot Berbasis Web di UPPIK RSUD M. YUNUS. *Komputika: Jurnal Sistem Komputer* Volume 12, Nomor 2, Oktober 2023, hlm.131 - 140. DOI: 10.34010/komputika.v12i2.8914 (10 Juli 2024).
- Gaikwad, S. 2019. *Chatbot with Personality Using Deep Learning*. Thesis. San Jose State University.
- Kurniawan, Bobby. 2021. *Chatbot Eksplorasi Ayat Al-Quran Menggunakan Algoritma Bidirectional Long Short-Term Memory (Bi-LSTM) dan Fuzzy String Matching*. Skripsi. Universitas Sumatera Utara.
- 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), 438–446. <https://doi.org/10.35870/jtik.v5i4.261>. (8 Juli 2024).
- Purwitasari, N.A dan Soleh, M. 2019. Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Processing. Vol. 6 No. 1 (2022): *Jurnal Ilmu Pengetahuan dan Teknologi (IPTEK)*. <https://doi.org/10.31543/jii.v6i1.192>. (8 Juli 2024).

- Pradana, S.T. 2022. Penerapan Algoritma *Levenshtein Distance* dalam pembuatan *Chatbot* informasi tentang Tajwid. Skripsi. UIN Sunan Gunung Djati Bandung.
- Sagita, N. 2022. *Chatbot* Layanan Informasi Kesehatan Organ Kewanitaan Menggunakan Pendekatan Metode *Long Short Term Memory* (LSTM). Skripsi. Universitas Sumatera Utara.
- Ruben, I. 2022. Implementasi Neural Network untuk Pembuatan Chatbot Menggunakan Dataset Pertanyaan Mahasiswa. KALBISIANA : Jurnal Mahasiswa Institut Teknologi dan Bisnis Kalbis Volume 8, No. 1, April 2022. (7 Juli 2024).
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is All You Need. Advances in Neural Information Processing Systems 30 (NIPS 2017).
- Zuraiyah, Z.A., Utami, D.K., dan Herlambang, D. 2019. Implementasi Chatbot Pada Pendaftaran Mahasiswa Baru Menggunakan Recurrent Neural Network. <http://dx.doi.org/10.35760/tr.2019.v24i2.2388>. (7 Juli 2024).

## LAMPIRAN

Berikut ini adalah tabel hasil chatbot diuji dengan 30 pertanyaan yang terdiri atas pertanyaan yang berkaitan dengan data dalam dataset.

Pertanyaan benar + Prediksi benar = True Positive

Pertanyaan benar + Prediksi salah = True Negative

Pertanyaan salah + Prediksi benar = False Positive

Pertanyaan salah + Prediksi salah = False Negative

No	Pertanyaan	Jawaban Chatbot	Keterangan
1	apa itu mad?	Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berharis fathah diikuti dengan alif sukun, huruf berharis kasrah diikuti ya sukun, dan huruf berharis dhommah diikuti wau sukun.	Cocok (TP)
2	berapa jumlah mad?	Dalam ilmu tajwid terdapat 16 mad yang terbagi ke dalam 3 jenis berdasarkan panjangnya harakatnya.	Cocok (TP)
3	berapa jenis mad?	Dalam ilmu tajwid terdapat 16 mad yang terbagi ke dalam 3 jenis berdasarkan panjangnya harakatnya.	Cocok (TP)
4	sebutkan jenis mad	Mad artinya membaca	Tidak cocok (TN)



		panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berharis fathah diikuti dengan alif sukun, huruf berharis kasrah diikuti ya sukun, dan huruf berharis dhommah diikuti wau sukun.	
5	jenis jenis mad	Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berharis fathah diikuti dengan alif sukun, huruf berharis kasrah diikuti ya sukun, dan huruf berharis dhommah diikuti wau sukun.	Tidak cocok (TN)
6	mad 2 harakat	Mad harus dibaca dua harakat dan tidak boleh lebih ketika ada bacaan mad yang sesudahnya bukan huruf hamzah, tasydid ataupun huruf mati. Mad yang panjangnya dua	Cocok (TP)

		harakat adalah Mad Thabi'i, Mad Badal, Mad Tamkin, Mad 'Twadh, Mad Harfi, Mad Lazim Harfi Mutsaqqal dan Mad Shilah Qashirah.	
7	mad 4 harakat	Mad harus dibaca empat harakat ketika ada bacaan mad atau ada tanda bendera diatas huruf bertemu dengan huruf hamzah. Mad yang panjangnya empat harakat adalah Mad Shilah Thowilah, Mad Wajib Muttasil, Mad Jaiz Munfashil, Mad Aridh Lissukun, dan Mad Layyin.	Cocok (TP)
8	mad aridh	Mad Aridh Lissukun ialah apabila ada huruf mad, sesudahnya berupa huruf sukun tidak asli karena waqaf. Mad Aridh Lissukun boleh dibaca dua, empat atau enam harakat tapi umumnya dibaca empat harakat. Contohnya pada surah Al Fil pada akhir setiap akhir ayatnya.	Cocok (TP)
9	berapa panjang mad silah?	Mad Shilah Qashirah ialah apabila ada ha dhamir yang sebelumnya berupa huruf	Cocok (TP)

		hidup, dan sesudahnya juga berupa huruf hidup yang bukan hamzah Qatha'. Mad Shilah Qasirah dibaca dengan panjang dua harakat. Contohnya kata innahuu pada Surah An-Nasr ayat 3	
10	mad yang panjangnya dibaca 6 harakat	Mad harus dibaca enam harakat ketika ada bacaan mad atau ada tanda bendera diatas huruf bertemu dengan huruf tasydid atau huruf mati. Mad yang panjangnya enam harakat adalah Mad Farqi, Mad Lazim Khilmi Mukhoffaf, Mad Lazim Kilmi Mutsaqqol, Mad Lazim Harfi Mutsaqqol.	Cocok (TP)
11	mad lazim mutaqol	Kata kunci yang kamu cari tidak ditemukan. Silakan coba dengan kata kunci lain. Contoh: jika kamu mencari tentang mad asli, maka sertakan kata 'mad asli'.	Tidak cocok (FN)
12	mad lazim mutsaqqol	Mad Lazim Kilmi Mutsaqqol adalah apabila ada huruf mad yang sesudahnya berupa huruf yang bertasydid dan terletak dalam satu kata. Mad Lazim Kilmi Mutsaqqol dibaca	cocok (TP)

		dengan panjang enam harakat. Contohnya pada Surah Al-Fatihah ayat 7 pada kata 'waladh dhaaaaaalliiiiin'	
13	mad lazim mukhoffaf	Kata kunci yang kamu cari tidak ditemukan. Silakan coba dengan kata kunci lain. Contoh: jika kamu mencari tentang mad asli, maka sertakan kata 'mad asli'.	Tidak cocok (TN)
14	mukhoffaf	Mad harus dibaca empat harakat ketika ada bacaan mad atau ada tanda bendera diatas huruf bertemu dengan huruf hamzah. Mad yang panjangnya empat harakat adalah Mad Shilah Thowilah, Mad Wajib Muttasil, Mad Jaiz Munfashil, Mad Aridh Lissukun, dan Mad Layyin.	Tidak cocok (FP)
15	mad silah towilah	Mad shilah thowilah alah apabila ada Ha Dhamir yang sebelumnya berupa huruf hidup, dan sesudahnya bertemu hamzah yang berharakat dhommah atau kasrah. Mad Shilah Thawilah dibaca dengan panjang 4 atau 5 harakat. Contohnya pada	Cocok (TP)

		surah Al-Humazah ayat 3 pada kata 'anna ma lahuuu akhladah;	
16	mad farq	Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad secara lengkap adalah membaca panjang ketika ada huruf berbaris fathah diikuti dengan alif sukun, huruf berbaris kasrah diikuti ya sukun, dan huruf berbaris dhommah diikuti wau sukun.	Cocok (FN)
17	mad farqi	Mad farqi adalah ketika mad badal bertemu tasydid. Mad Farqi dibaca dengan panjang enam harakat. Contohnya pada surah Al-An'am ayat 143 pada kata 'qul aaaaaadzdakaraini'	Cocok (TP)
18	mad yang dibaca 2 harakat	Mad harus dibaca dua harakat dan tidak boleh lebih ketika ada bacaan mad yang sesudahnya bukan huruf hamzah, tasydid ataupun huruf mati. Mad yang panjangnya dua	Cocok (TP)

		harakat adalah Mad Thabi'i, Mad Badal, Mad Tamkin, Mad 'Twadh, Mad Harfi, Mad Lazim Harfi Mutsaqqal dan Mad Shilah Qashirah.	
19	kapan mad dibaca 4 harakat?	Mad harus dibaca empat harakat ketika ada bacaan mad atau ada tanda bendera diatas huruf bertemu dengan huruf hamzah. Mad yang panjangnya empat harakat adalah Mad Shilah Thowilah, Mad Wajib Muttasil, Mad Jaiz Munfashil, Mad Aridh Lissukun, dan Mad Layyin.	Cocok (TP)
20	mengapa mad dibaca 4 harakat?	Mad harus dibaca empat harakat ketika ada bacaan mad atau ada tanda bendera diatas huruf bertemu dengan huruf hamzah. Mad yang panjangnya empat harakat adalah Mad Shilah Thowilah, Mad Wajib Muttasil, Mad Jaiz Munfashil, Mad Aridh Lissukun, dan Mad Layyin.	Cocok (TP)
21	jika ada huruf mad bertemu dengan hamzah	Mad artinya membaca panjang bacaan Al-Quran ketika ada huruf mad. Huruf mad ada tiga yaitu alif, ya dan wau. Definisi mad	tidak cocok (TN)

		<p>secara lengkap adalah membaca panjang ketika ada huruf berbaris fathah diikuti dengan alif sukun, huruf berbaris kasrah diikuti ya sukun, dan huruf berbaris dhommah diikuti wau sukun.</p> <p>jika ada huruf mad bertemu dengan hamzah</p>	
22	contoh mad lazim	<p>Kata kunci yang kamu cari tidak ditemukan. Silakan coba dengan kata kunci lain. Contoh: jika kamu mencari tentang mad asli, maka sertakan kata 'mad asli'.</p>	tidak cocok (FN)
23	sebutkan mad mad dalam tajwid	<p>Berdasarkan panjang harakatnya, mad bisa dibagi tiga yaitu panjang dua harakat, empat harakat dan enam harakat.</p>	cocok (TN)
24	mad lazim harfi	<p>Mad harfi adalah ketika membaca huruf Muqoththo'ah atau huruf Fawatihus Suwar. Huruf Muqoththo'ah atau huruf Fawatihus Suwar adalah huruf-huruf hijaiyah yang terdapat dalam beberapa awal surat dalam Al-Quran</p>	tidak cocok (FP)

		yang cara membacanya putus-putus per hurufnya. Huruf pada mad harfi ada lima huruf yaitu ha (ha kecil), ya, tho, ha (ha besar) dan ro. Mad harfi dibaca dengan panjang dua harakat. Contohnya huruf ya pada ayat Yaasiiin	
25	mad lazim harfi mukhoffaf	Mad Lazim Harfi Mukhoffaf yaitu terdapat pada huruf hijaiyah yang menjadi Fawatihus Suwar (pembuka surah) yang hanya dibaca dua harakat. Huruf mad ini ada 5 yaitu ha kecil, ya, tho, ha besar dan ra. Mad Lazim Harfi Mukhoffaf dibaca dengan panjang dua harakat. Contohnya, huruf ya pada Yasin, huruf ha pada hamim, huruf ra pada alif lam ra.	cocok (TP)
26	mad kilmi	Mad badal adalah apabila ada huruf mad yang sebelumnya berupa huruf hamzah dan sesudahnya tidak diikuti hamzah atau huruf mati. Mad badal dibaca dengan panjang dua harakat. Contohnya dalam Surah Al Insyiqaq ayat 7,	tidak cocok (FN)



		yaitu kata 'uutiya'	
27	tamkin	Mad tamkin adalah ketika bertemu dua huruf ya, yaitu ya pertama bertasydid dan berharakat kasrah, sedangkan ya kedua berharakat sukun. Contohnya pada surah Al-Muthaffifin ayat 18 pada kata 'illiyiin	Cocok (TP)
28	pengertian mad silah	Mad Shilah Qasirah ialah apabila ada ha dhamir yang sebelumnya berupa huruf hidup, dan sesudahnya juga berupa huruf hidup yang bukan hamzah Qatha'. Mad Shilah Qasirah dibaca dengan panjang dua harakat. Contohnya kata innahuu pada Surah An-Nasr ayat 3	Cocok (TP)
29	Pengertian mad lin	Mad Layyin atau Mad Lin adalah apabila ada huruf lin, yaitu ada huruf wau sukun atau ya sukun sebelumnya berupa huruf yang berharakat fathah, dan sesudahnya berupa huruf sukun (mati) tidak asli (baik berupa huruf hamzah atau bukan) yang disebabkan	Cocok (TP)

		<p>waqaf. Mad Layyin boleh dibaca dua, empat atau enam harakat tapi umumnya dibaca empat harakat. Contohnya pada Surah Quraishy pada setiap akhir ayatnya.</p>	
30	mad towilah	<p>Kata kunci yang kamu cari tidak ditemukan. Silakan coba dengan kata kunci lain. Contoh: jika kamu mencari tentang mad asli, maka sertakan kata 'mad asli'.</p>	tidak cocok (TN)