

# AI CHATBOT IMPLEMENTATION FOR NURUL JADID UNIVERSITY WEBSITE USING LSTM ALGORITHM

**Sudriyanto<sup>1,\*</sup>, Kamil Malik<sup>2</sup>, Jamal<sup>3</sup>**

<sup>1,2</sup>Teknik Informatika, Universitas Nurul Jadid, Indonesia

<sup>3</sup>Sistem Informasi Akuntansi, Akademi Manajemen Informatika dan Komputer Taruna, Indonesia

EMAIL: ([sudriyanto@unuja.ac.id](mailto:sudriyanto@unuja.ac.id), [nomor1001@gmail.com](mailto:nomor1001@gmail.com), [Tanpade@gmail.com](mailto:Tanpade@gmail.com))

Diterima : 28 Maret 2025. Disetujui : 20 Mei Tahun. Dipublikasikan : 12 Juni 2025.

**ABSTRACT** - The rapid advancement of technology has brought significant changes in various aspects of life, including the education sector. As an educational institution, Nurul Jadid University must adopt the latest technology to enhance efficiency and service, particularly in responding to the increasing volume of inquiries and information needs from the public and parents before enrolling their children. A chatbot, as part of Natural Language Processing (NLP) based on Artificial Intelligence (AI), is designed to interact with users through text or voice, providing fast, accurate, and continuous service. The Long Short-Term Memory (LSTM) algorithm in deep learning is utilized for text data prediction and classification. In this research, the data consists of tags, patterns, and responses obtained manually from the official Nurul Jadid University website and then preprocessed to develop the chatbot model. The core component of this model is the embedding layer, which assigns vector values to each word in the processed text data. The model training results indicate an accuracy of 99.32% and a loss of 12.57%,

demonstrating that the model performs well without overfitting or underfitting, making it suitable for testing and deployment. Thus, the LSTM-based chatbot serves as an effective virtual assistant to help the public, prospective students, and current students access information more easily and efficiently.

**Keywords :** Chatbot, Deep Learning, Long Short-Term Memory, Natural Language Processing, Universitas

**ABSTRAK** - Perkembangan teknologi yang pesat telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk di sektor pendidikan. Universitas Nurul Jadid sebagai institusi pendidikan perlu mengadopsi teknologi terkini guna meningkatkan efisiensi serta pelayanan, khususnya dalam menjawab tingginya volume pertanyaan dan kebutuhan informasi dari masyarakat atau orang tua calon mahasiswa sebelum proses pendaftaran. Chatbot merupakan salah satu penerapan Natural Language Processing (NLP) berbasis Artificial Intelligence (AI) yang dirancang untuk berinteraksi dengan

pengguna melalui teks atau suara, sehingga dapat memberikan layanan secara cepat, akurat, dan sepanjang waktu. Algoritma Long Short-Term Memory (LSTM) dalam deep learning digunakan untuk melakukan prediksi serta klasifikasi data teks. Dalam penelitian ini, data yang digunakan terdiri dari tag, pola (pattern), dan respons yang diperoleh secara manual dari referensi situs web resmi Universitas Nurul Jadid, kemudian diproses lebih lanjut (preprocessing) untuk membangun model chatbot. Bagian utama dari model ini adalah lapisan embedding, yang berfungsi memberikan nilai vektor pada setiap kata dalam data teks yang telah dimasukkan. Hasil pelatihan model menunjukkan tingkat akurasi sebesar 99,32% dan nilai loss sebesar 12,57%. Hal ini menunjukkan bahwa model telah bekerja dengan baik tanpa mengalami overfitting atau underfitting, sehingga layak untuk diuji dan diterapkan. Dengan demikian, chatbot berbasis LSTM ini dapat menjadi asisten virtual yang efektif dalam membantu masyarakat, calon mahasiswa, maupun mahasiswa dalam memperoleh informasi secara lebih mudah dan efisien.

**Kata kunci :** *Chatbot, Deep Learning, Long Short-Term Memory, Natural Language Processing, Universitas*

## I. PENDAHULUAN

Di era digital saat ini, perkembangan teknologi informasi mengalami kemajuan yang sangat pesat dan dinamis, sehingga membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam dunia pendidikan [1], [2]. Salah satu institusi pendidikan yang perlu mengadopsi teknologi terkini untuk meningkatkan efisiensi layanan akademik adalah Universitas Nurul Jadid, sebuah perguruan tinggi berbasis pesantren yang berupaya mengikuti perkembangan zaman dengan tetap mempertahankan nilai-nilai pendidikannya [3]. Seiring dengan meningkatnya jumlah calon mahasiswa

yang tertarik untuk melanjutkan studi di Universitas Nurul Jadid setiap tahunnya, terjadi peningkatan volume pertanyaan dan kebutuhan informasi yang semakin besar, menuntut kecepatan serta keakuratan dalam memberikan respons yang memadai [4]. Kondisi ini menandakan adanya kebutuhan akan suatu sistem yang mampu memberikan jawaban cepat dan akurat terhadap pertanyaan masyarakat atau calon mahasiswa terkait informasi akademik maupun administrasi Universitas Nurul Jadid [5].

Selama ini, proses pelayanan informasi yang dilakukan secara manual oleh staf administrasi dalam menjawab pertanyaan calon mahasiswa dan masyarakat umum membutuhkan waktu serta tenaga yang cukup besar, sehingga efektivitasnya dirasa belum optimal [6]. Biasanya, calon mahasiswa atau masyarakat yang ingin memperoleh informasi menghubungi staf akademik dan panitia penerimaan mahasiswa baru melalui berbagai platform komunikasi pribadi seperti WhatsApp, Telegram, serta email [5]. Namun, keterbatasan jam kerja dan respons manual menyebabkan adanya penundaan dalam memperoleh jawaban atas pertanyaan yang diajukan, sehingga informasi yang diterima tidak selalu tersedia secara cepat dan real-time [7].

Chatbot merupakan salah satu implementasi dari teknologi Artificial Intelligence (AI) yang berfokus pada domain Natural Language Processing (NLP), di mana sistem ini dirancang untuk dapat berinteraksi dengan pengguna melalui percakapan berbasis teks atau suara secara otomatis tanpa campur tangan langsung dari manusia [8]. Meskipun chatbot dibuat untuk meniru percakapan manusia, seluruh proses komunikasi berlangsung secara otomatis dengan menggunakan mesin yang bekerja berdasarkan model AI, sehingga mampu menghemat waktu serta biaya operasional [9]. Berdasarkan survei dan penelitian yang telah dilakukan sebelumnya, penggunaan chatbot terbukti sangat efisien dalam

meningkatkan efektivitas layanan, menghemat anggaran, mempersingkat waktu respons, serta memberikan akses 24 jam bagi pengguna yang membutuhkan jawaban atas pertanyaan mereka kapan saja dan di mana saja [10].

Dalam pengembangannya, berbagai metode telah digunakan dalam penelitian terdahulu untuk membangun chatbot dengan tingkat akurasi yang cukup tinggi, berkisar antara 80 hingga 90% [11]. Salah satu metode yang sering digunakan dalam pembuatan chatbot adalah Recurrent Neural Network (RNN), yang dikenal memiliki kemampuan prediksi cukup baik dalam memproses informasi terbaru. Namun, meskipun RNN memiliki keunggulan dalam memprediksi kata berdasarkan data terkini, algoritma ini memiliki kelemahan yang cukup signifikan, yaitu kesulitan dalam mengingat informasi dalam jangka waktu yang panjang, sehingga akurasinya dapat menurun dalam beberapa kasus tertentu. Oleh karena itu, dikembangkanlah algoritma Long Short-Term Memory (LSTM) sebagai penyempurnaan dari RNN tanpa mengurangi keunggulan yang dimiliki RNN dalam memproses data secara akurat [12].

Berdasarkan berbagai referensi penelitian terdahulu, metode LSTM dipilih dalam penelitian ini sebagai pendekatan yang lebih optimal dalam membangun model chatbot. Algoritma LSTM sangat efektif dalam mengolah bahasa alami (Natural Language Processing) dengan tujuan utama untuk memudahkan pengguna dalam berkomunikasi dengan sistem komputer secara lebih interaktif dan natural [13].

Berdasarkan berbagai permasalahan yang telah diidentifikasi di atas, penelitian ini bertujuan untuk mengembangkan solusi berbasis teknologi dalam bentuk asisten virtual atau chatbot yang dapat memberikan layanan informasi akademik secara otomatis kepada calon mahasiswa dan masyarakat umum. Dengan menggunakan model berbasis LSTM, chatbot ini

diharapkan mampu memberikan respons secara real-time, cepat, serta akurat, sehingga dapat meningkatkan kualitas pelayanan informasi akademik di Universitas Nurul Jadid [14].

## II. TEORI DASAR

Dalam penelitian ini, terdapat berbagai teori dasar yang menjadi landasan dalam pengembangan teknologi dan sistem cerdas. Teori-teori ini memberikan pemahaman mendalam tentang cara kerja algoritma, pemrosesan data, serta interaksi antara manusia dan komputer. Pada penelitian ini, pengembangan chatbot berbasis kecerdasan buatan menggunakan algoritma Long Short-Term Memory (LSTM) mengacu pada beberapa teori utama dalam bidang informatika. Landasan teori berikut akan menjelaskan konsep-konsep yang mendukung implementasi chatbot sebagai asisten virtual dalam sistem informasi akademik Universitas Nurul Jadid.

1. Artificial Intelligence (AI) dan Natural Language Processing (NLP)  
Artificial Intelligence (AI) merupakan cabang ilmu komputer yang berfokus pada pengembangan sistem yang dapat meniru kecerdasan manusia. Salah satu bidang dalam AI yang berkembang pesat adalah Natural Language Processing (NLP), yaitu teknik yang memungkinkan komputer memahami, memproses, dan merespons bahasa manusia secara alami. Dalam penelitian ini, chatbot yang dikembangkan menggunakan AI berbasis NLP untuk menangani interaksi dengan pengguna melalui teks atau suara. Dengan adanya chatbot ini, Universitas Nurul Jadid dapat menyediakan layanan informasi akademik secara otomatis, efisien, dan responsif sepanjang waktu [15].

## 2. Deep Learning dan Long Short-Term Memory (LSTM)

Deep Learning adalah salah satu metode pembelajaran mesin yang menggunakan jaringan saraf tiruan untuk mengekstraksi pola dari data besar. Salah satu algoritma deep learning yang digunakan dalam penelitian ini adalah Long Short-Term Memory (LSTM), yang merupakan pengembangan dari Recurrent Neural Network (RNN). LSTM dirancang untuk mengatasi kelemahan RNN dalam mengingat informasi jangka panjang dengan menggunakan struktur khusus yang terdiri dari forget gate, input gate, dan output gate. Dengan model LSTM, chatbot mampu memahami konteks percakapan dan menghasilkan respons yang lebih relevan dan akurat [16].

## 3. Basis Data dan Data Preprocessing

Dalam sistem berbasis kecerdasan buatan, kualitas data menjadi faktor utama dalam membangun model yang efektif. Data yang digunakan dalam penelitian ini diperoleh secara manual dari website Universitas Nurul Jadid dan disimpan dalam format JSON, yang sesuai untuk pemrosesan data dalam machine learning. Sebelum digunakan untuk pelatihan model, data harus melalui tahap preprocessing, yaitu proses pembersihan dan transformasi data agar lebih mudah dipahami oleh model. Beberapa teknik preprocessing yang digunakan meliputi penghapusan tanda baca (remove punctuation), normalisasi kata (lemmatization), tokenisasi teks (tokenization), serta pengkodean data (encoding) agar dapat diolah dalam bentuk numerik [17].

## 4. Machine Learning dan Evaluasi Model

Machine Learning adalah metode pembelajaran otomatis yang memungkinkan sistem untuk mengenali pola dari data dan membuat prediksi berdasarkan pola tersebut. Dalam penelitian ini, chatbot dilatih menggunakan model LSTM untuk mengoptimalkan akurasi model. Evaluasi model dilakukan dengan mengukur akurasi model telah terlatih dengan baik tanpa mengalami overfitting atau underfitting. Dengan tingkat akurasi yang tinggi, chatbot mampu memberikan respons yang sesuai dengan pertanyaan pengguna, sehingga dapat digunakan dalam sistem informasi akademik Universitas Nurul Jadid [18].

## 5. Interaksi Manusia dan Komputer (HCI)

Interaksi Manusia dan Komputer (Human-Computer Interaction atau HCI) merupakan bidang ilmu yang berfokus pada bagaimana manusia berinteraksi dengan sistem berbasis komputer. Penelitian ini bertujuan untuk meningkatkan efisiensi pelayanan informasi akademik dengan mengembangkan chatbot yang dapat diakses kapan saja tanpa keterbatasan waktu. Dalam implementasinya, chatbot ini dirancang dengan antarmuka yang responsif, sehingga dapat digunakan melalui perangkat desktop maupun mobile. Selain itu, sistem ini juga dilengkapi dengan fitur speech recognition, yang memungkinkan pengguna memberikan input dalam bentuk suara maupun teks, sehingga meningkatkan pengalaman pengguna dalam berinteraksi dengan chatbot [19].

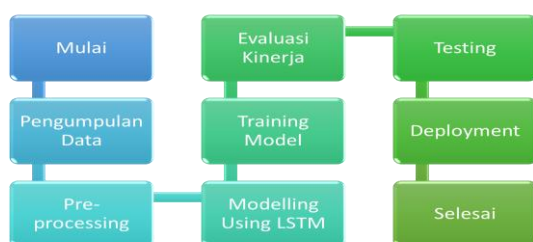
## 6. Pengembangan Perangkat Lunak dan Cloud Computing

Dalam implementasi chatbot, pengembangan perangkat lunak

berbasis web sangat penting agar sistem dapat digunakan secara luas. Model chatbot yang telah dilatih disimpan dalam format .h5 dan diterapkan menggunakan framework Flask berbasis Python. Flask digunakan untuk menghubungkan model chatbot dengan antarmuka pengguna di website Universitas Nurul Jadid. Selain itu, sistem ini dapat diintegrasikan dengan layanan Cloud Computing, yang memungkinkan chatbot untuk berjalan secara efisien dengan kapasitas penyimpanan dan komputasi yang lebih fleksibel [20].

### III. METODE

Penelitian ini dilakukan melalui beberapa tahap, dimulai dari pengambilan data melalui web, dilanjutkan dengan prapemrosesan data yang mencakup penghapusan tanda baca, lemmatisasi, tokenisasi, padding teks, dan encoding. Setelah itu, dilakukan pemodelan dan pelatihan model untuk meningkatkan kinerjanya. Model kemudian dievaluasi guna memastikan kesesuaian (goodfitting) dan diuji untuk menilai performanya. Tahap akhir adalah deployment, di mana model diterapkan pada aplikasi berbasis web [21]. Alur lengkap penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Kerangka Penelitian

#### 3.1 Pengumpulan Data

Data dalam penelitian ini diperoleh secara manual melalui referensi situs web <https://www.unuja.ac.id/> dan disimpan dalam format JSON, yang sangat sesuai untuk pengembangan model chatbot [22]. JSON merupakan

jenis data semi-struktural yang memuat elemen seperti tag, pola (pattern), konteks, dan respons [23]. Proses pengumpulan data berlangsung selama kurang lebih satu bulan dengan penyesuaian terhadap pertanyaan dan jawaban. Data yang dikumpulkan mencakup label tag, pertanyaan, dan jawaban, kemudian diproses untuk pelatihan model sebanyak 200 kali guna memperoleh tingkat akurasi yang optimal.

Tabel 1. Contoh dari Data yang Digunakan

| Tag      | Patterns                         | Responses                                                                                        |
|----------|----------------------------------|--------------------------------------------------------------------------------------------------|
| greeting | halo                             | Hai, aku bot siap membantumu.                                                                    |
| univ     | Apa itu Universitas Nurul Jadid? | Universitas Nurul Jadid adalah perguruan tinggi swasta favorit berbasis pesantren di Jawa Timur. |
| fakultas | Ada fakultas apa saja di sana?   | Ada fakultas teknik, fakultas agama Islam, fakultas sosial & humaniora, dan fakultas kesehatan.  |

#### 3.2 Preprocessing Data

Setelah data diperoleh dan diberikan label, langkah selanjutnya adalah melakukan preprocessing untuk membersihkan data sehingga lebih mudah diolah oleh mesin. Proses ini mencakup beberapa tahapan penting, yaitu remove punctuation, lemmatization, tokenization, padding text, dan encoding. Berikut adalah penjelasan masing-masing tahapan:

##### a. Remove Punctuation

Tahap ini bertujuan untuk menghapus tanda baca atau karakter khusus seperti '?', '!', '.', ',', ';', ':', dan simbol lainnya dalam teks. Penghapusan tanda baca dilakukan agar proses pengolahan data menjadi lebih efisien.

##### b. Lemmatization

Lemmatization merupakan proses normalisasi kata yang

menggunakan kosakata dan analisis morfologi untuk menghilangkan akhiran infleksi guna mengembalikan kata ke bentuk dasarnya. Contoh: kata menggunakan akan dikembalikan ke bentuk dasarnya, yaitu guna.

c. Tokenization

Proses ini membagi teks menjadi unit-unit yang lebih kecil yang disebut token, seperti kata atau frasa, dengan menghilangkan elemen tertentu seperti tanda baca. Contohnya, kalimat Aku pergi ke Makassar akan dipecah menjadi token 'Aku', 'pergi', 'ke', dan 'Makassar'.

d. Padding Text

Padding adalah proses menyesuaikan panjang teks agar seragam dengan menambahkan nilai nol (0) di awal atau akhir sesuai dengan panjang maksimum yang ditentukan. Dalam implementasinya, metode ini memanfaatkan pustaka `pad_sequence` dengan fungsi `pad_sequence()` yang menerima token sebagai parameter.

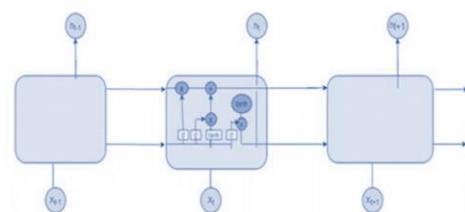
e. Encoding

Encoding adalah proses mengonversi data teks atau kategori menjadi bentuk numerik, seperti representasi biner (0 dan 1), agar dapat diolah oleh sistem komputasi secara lebih efektif.

### 3.3 Membuat Model

Setelah melalui tahap preprocessing, data yang telah disiapkan digunakan untuk membangun model machine learning. Teknik yang diterapkan dalam proses ini adalah Long Short-Term Memory (LSTM), yang merupakan salah satu metode pembelajaran mendalam yang banyak digunakan dan sangat sesuai untuk

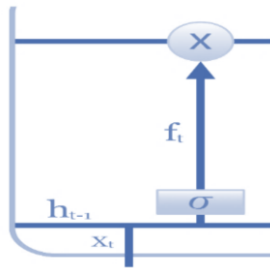
tugas klasifikasi serta prediksi data berbasis teks maupun deret waktu [23]. LSTM dikembangkan sebagai bentuk penyempurnaan dari algoritma Recurrent Neural Network (RNN) [12]. Meskipun RNN mampu menghasilkan prediksi yang akurat berdasarkan data yang tersedia, algoritma ini memiliki keterbatasan dalam menyimpan informasi dalam jangka waktu panjang. Untuk mengatasi keterbatasan tersebut sekaligus mempertahankan keunggulan RNN [11], dikembangkanlah algoritma LSTM yang memungkinkan penyimpanan informasi dalam durasi yang lebih panjang [24]. Selanjutnya, data deret waktu yang telah diproses digunakan dalam pembentukan model yang berfungsi untuk menganalisis, memprediksi, serta mengelompokkan data [25]. Struktur LSTM terdiri dari jaringan saraf dengan sejumlah memori blok yang disebut sel. Informasi dari cell state dan hidden state akan diteruskan ke sel berikutnya dalam jaringan, sebagaimana ditunjukkan pada Gambar 2, di mana elemen berbentuk persegi panjang berwarna biru merepresentasikan sel dalam arsitektur LSTM [26].



Gambar 2. Menunjukkan arsitektur dari model LSTM.

Data yang diproses melalui perhitungan LSTM disimpan dalam sel, sementara pengelolaan memorinya dikendalikan oleh komponen yang disebut pintu (gate). LSTM memiliki tiga jenis pintu, yaitu pintu penghapusan (forget gate), pintu

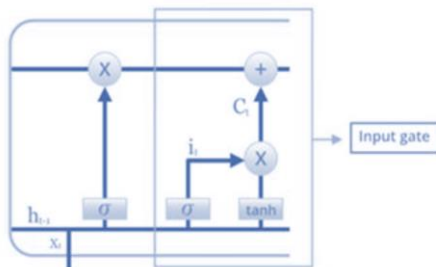
masuk (input gate), dan pintu keluaran (output gate) [27].



Gambar 3. Gerbang Lupa (Forget Gate)

$$F_1 = \sigma(W_f [h_{(t-1)}, x_t] + b_f)$$

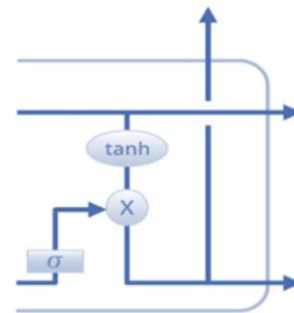
Persamaan (1) digunakan untuk menghitung forget gate, yang menentukan seberapa banyak informasi dari memori sebelumnya yang akan diabaikan berdasarkan masukan  $x(t)$  dan keluaran dari sel sebelumnya  $h(t-1)$ .  $F_1$  berperan sebagai output gate yang mengatur proses penghapusan informasi. Perhitungannya dilakukan dengan mengalikan vektor input dengan matriks bobot ( $W_f$ ), menambahkan bias ( $b_f$ ), lalu menerapkan fungsi aktivasi sigmoid ( $Q$ ). Hasil dari proses ini kemudian diproses melalui fungsi aktivasi untuk menghasilkan output biner. Jika output bernilai 0, informasi akan dihapus, sedangkan jika output bernilai 1, informasi akan dipertahankan [28].



Gambar 4. menunjukkan Gerbang Input.

$$i_t = \sigma(W_f [h_{t-1}, x_t] + b_i)$$

Persamaan (2) digunakan untuk menghitung input gate. Proses penambahan data yang berguna ke dalam sel dilakukan oleh information gate dengan memanfaatkan fungsi sigmoid untuk menyaring nilai yang relevan agar dapat disimpan. Siklus ini ditunjukkan pada Gambar 3, yang menggunakan input dari  $h(t-1)$  dan  $x(t)$ . Selanjutnya, sebuah vektor dibentuk menggunakan fungsi tanh, yang menghasilkan nilai dalam rentang -1 hingga +1 dan mencerminkan potensi manfaat dari  $h(t-1)$  dan  $x(t)$ . Terakhir, nilai tersebut direplikasi untuk memperoleh data yang bernilai [29].



Gambar 5. Output Gate

$$O_t = \sigma(W_o [h_{(t-1)}, x_t] + b_o)$$

Persamaan (3) digunakan untuk menghitung output gate dengan mengekstrak informasi yang relevan dari cell state sebagai nilai keluaran yang diproses oleh output gate. Proses ini diawali dengan menerapkan fungsi tanh pada cell state, kemudian hasilnya dikombinasikan dengan fungsi sigmoid untuk menyaring nilai berdasarkan input  $h_{t-1}$  dan  $x_t$ . Selanjutnya, nilai hasil penyaringan dikalikan dengan nilai keluaran sebelum akhirnya diteruskan ke sel berikutnya [25]. LSTM beroperasi melalui beberapa tahapan, dimulai dengan menentukan data yang akan dipertahankan dan data yang akan dihapus dari cell state. Lapisan sigmoid berfungsi untuk mengidentifikasi serta

mengeliminasi informasi yang tidak relevan, sementara fungsi tanh dan sigmoid bekerja sama dalam membedakan informasi yang penting dan tidak penting. Akhirnya, keluaran ditentukan melalui cell state yang telah difilter oleh kedua fungsi tersebut [29], [30].

### 3.4 Training

Pelatihan merupakan tahap untuk melatih model agar mesin dapat mengenali dan memahami pola dari model tersebut. Pemodelan yang dilakukan akan dilatih sebanyak 200 kali guna mengetahui tingkat akurasi, sehingga saat pengujian dilakukan, hasil yang diperoleh sesuai dengan yang diharapkan.

### 3.5 Evaluasi

Evaluasi dilakukan untuk menganalisis kinerja model dengan memplot hasil akurasi dan loss guna menilai performa model LSTM. Melalui analisis evaluasi berdasarkan tabel visualisasi, dapat ditentukan apakah model telah mencapai kinerja yang optimal serta mengidentifikasi kemungkinan terjadinya overfitting atau underfitting. Dengan demikian, kelayakan model untuk pengujian dapat diketahui dengan lebih mudah.

## IV. HASIL DAN PEMBAHASAN

Pengumpulan data secara manual dari halaman website unuja.ac.id menghasilkan 4.461 entri teks. Data tersebut kemudian diklasifikasikan ke dalam tiga kategori, yaitu "tag," "patterns," dan "responses." Proses pengambilan data dilakukan dalam rentang waktu November hingga Desember 2024.



Gambar 6. Website Universitas Nurul Jadid

Dari data yang telah diperoleh dan diklasifikasikan ke dalam label, selanjutnya akan dilakukan tahap prapemrosesan data guna menghasilkan data yang siap untuk dilatih.

|     | patterns                                     | tags                |
|-----|----------------------------------------------|---------------------|
| 0   | hallo                                        | greeting            |
| 1   | hai                                          | greeting            |
| 2   | halo                                         | greeting            |
| 3   | hei                                          | greeting            |
| 4   | hi                                           | greeting            |
| ... | ...                                          | ...                 |
| 584 | ada berapa program studi di s2 ?             | Prodi_pasca_sarjana |
| 585 | ada berapa prodi di pascasarjana ?           | Prodi_pasca_sarjana |
| 586 | apakah unuja ada program kuliah untuk s2 ?   | pascasarjana        |
| 587 | apakah menyediakan program kuliah untuk s2 ? | pascasarjana        |
| 588 | apa ada program s2 di unuja ?                | pascasarjana        |

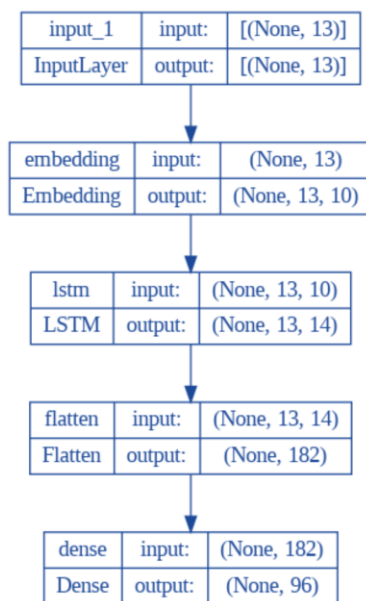
Gambar 7. Data Frame

Tabel 2. Proses Preprocessing Teks

| Sebelum                                     | Metode                    | Sesudah                                                              |
|---------------------------------------------|---------------------------|----------------------------------------------------------------------|
| "Jelaskan tentang universitas nurul jadid!" | <i>Remove Punctuation</i> | ["Jelaskan tentang universitas nurul jadid"]                         |
|                                             | <i>Lemmatization</i>      | ["jelas", "tentang", "universitas", "universitas", "nurul", "jadid"] |
|                                             | <i>Tokenization</i>       | ["jelaskan", "tentang", "universitas", "nurul", "jadid"]             |
| "Fasilitas apa yang di sediakan?"           | <i>Remove Punctuation</i> | ["Fasilitas apa yang disediakan"]                                    |
|                                             | <i>Lemmatization</i>      | ["fasilitas", "apa", "sedia"]                                        |
|                                             | <i>Tokenization</i>       | ["fasilitas", "apa", "yang", "di", "sediakan"]                       |

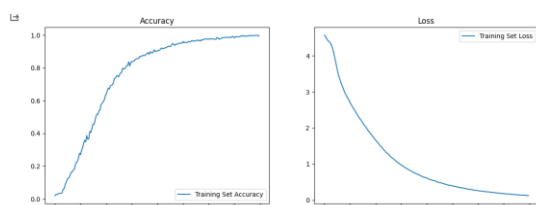
Pada Gambar 7, hanya data dengan label "tag" dan "patterns" yang diproses. Setelah melalui tahap *preprocessing*, diperoleh 4.461 data dengan 227 kata unik dan 96 kelas. Proses *preprocessing* yang

dilakukan, sebagaimana ditunjukkan pada Tabel 2, hanya mencakup data berlabel “tag” dan “patterns,” yang kemudian akan digunakan untuk pelatihan model *Long Short-Term Memory* (LSTM). Sebelum memasuki tahap pelatihan, model LSTM harus terlebih dahulu dibangun menggunakan arsitektur dengan 14 layer, yang dibuat menggunakan pustaka Keras/TensorFlow. Visualisasi arsitektur model LSTM dapat dilihat pada Gambar 8 di bawah ini.



Gambar 8. Arsitektur Model

Setelah model selesai dibangun, langkah selanjutnya adalah mengevaluasi kinerjanya dengan menggunakan plot model untuk menganalisis tingkat akurasi dan loss, seperti yang ditunjukkan pada Gambar 9.



Gambar 9. Evaluasi Model

Berdasarkan visualisasi hasil model LSTM, pelatihan model sebanyak 200 kali menghasilkan akurasi sebesar 99,32% dan

loss sebesar 12,57%. Hal ini menunjukkan bahwa model telah mencapai performa yang baik tanpa mengalami overfitting atau underfitting, sehingga layak untuk diuji lebih lanjut dan diterapkan dalam deployment.



Gambar 10. Pengujian

Pada Gambar 10, model chatbot LSTM memberikan jawaban yang akurat dan sesuai dengan yang diharapkan. Oleh karena itu, model ini dapat disimpan untuk di-deploy menggunakan framework web Python, yaitu Flask. Untuk tampilan aplikasi, dapat digunakan antarmuka berbasis front-end dasar, serta model berformat .h5 yang telah disimpan dapat diintegrasikan menggunakan bahasa Python.



Gambar 11. Demo Chatbot

Gambar 11 menampilkan demo chatbot berbasis website yang mengintegrasikan model LSTM untuk memproses input dan menghasilkan output sesuai dengan pertanyaan pengguna. Chatbot ini bersifat responsif dan dapat diakses melalui ponsel maupun laptop. Selain tampilan yang intuitif, chatbot ini juga dilengkapi dengan fitur *Speech Recognition* atau pengenalan suara, sehingga pengguna dapat memberikan input baik melalui suara maupun dengan mengetik teks.

## V. KESIMPULAN

Penelitian ini menawarkan solusi berupa aplikasi chatbot berbasis kecerdasan buatan yang mampu merespons pertanyaan secara cepat dan akurat selama 24 jam sehari. Data yang digunakan diperoleh secara manual dari situs web Universitas Nurul Jadid, kemudian diberi pelabelan untuk keperluan pemrosesan lebih lanjut. Setelah melalui tahap pra-pemrosesan, data digunakan dalam pelatihan model menggunakan algoritma Long Short-Term Memory (LSTM), yang terbukti efektif dalam klasifikasi teks. Hasil pelatihan sebanyak 200 kali menunjukkan akurasi sebesar 99,32% dan nilai loss sebesar 12,57%, tanpa indikasi overfitting maupun underfitting, sehingga model dinilai layak untuk diuji lebih lanjut. Pengujian menunjukkan bahwa model mampu memahami pertanyaan pengguna dengan baik dan menghasilkan output yang sesuai dengan input yang diberikan. Model ini kemudian disimpan dalam format .h5 untuk diintegrasikan ke dalam chatbot berbasis website menggunakan framework Flask pada Python. Chatbot ini bersifat responsif dan dilengkapi dengan fitur speech recognition atau pengenalan suara, sehingga memungkinkan pengguna memberikan input dalam bentuk suara maupun teks. Dengan demikian, penelitian ini menghasilkan chatbot berbasis kecerdasan buatan yang mampu memberikan jawaban secara cepat dan akurat serta dapat dimanfaatkan sebagai asisten virtual dalam membantu staf atau

panitia penerimaan mahasiswa baru (PMB) merespons pertanyaan secara lebih efisien.

## REFERENSI

- [1] M. Faid, K. Malik, and A. Supriadi, "Evaluasi Model Jaringan Saraf Tiruan Berbasis LSTM dalam Memprediksi Fluktuasi Harga Bitcoin," *Journal of Advanced Research in Informatics*, vol. 2, no. 3025–1796, pp. 15–22, Jun. 2024, doi: <https://doi.org/10.24929/jars.v2i2.3398>.
- [2] M. Ali Hafid and M. Ade Kurniawan, "Deteksi Akun Kaggle Bot Menggunakan Linear Regression," *Journal homepage: Journal of Electrical Engineering and Computer (JEECOM)*, vol. 6, no. 2, 2024, doi: 10.33650/jeeecom.v4i2.
- [3] N. Ningsih, N. A. S. Iskandar, S. Rizqiyah, and S. Sudriyanto, "Prediksi Churn Pelanggan Industri Telekomunikasi Menggunakan Metode Artificial Neural Network Berbasis Streamlit," *JUSTIFY: Jurnal Sistem Informasi Ibrahimy*, vol. 3, no. 2, pp. 105–114, Dec. 2024, doi: 10.35316/justify.v3i2.5544.
- [4] S. Sudriyanto, "Optimizing Neural Networks Using Particle Swarm Optimization (PSO) Algorithm for Hypertension Disease Prediction," *JEECOM Journal of Electrical Engineering and Computer*, vol. 5, no. 2, 2023, doi: 10.33650/jeeecom.v5i2.6759.
- [5] R. Hidayat, R. A. Ronaldo, R. A. Prasetyo, and S. A. Edho Wicaksono, "Optimasi Parameter Support Vector Machine Menggunakan Algoritma Genetika untuk Meningkatkan Prediksi Pergerakan Harga Saham," *COREAI: Jurnal Kecerdasan Buatan, Komputasi dan Teknologi Informasi*, vol. 3, no. 1, 2022, doi: 10.33650/coreai.v3i1.3859.
- [6] A. Ghazvini, N. Mohd Sharef, and F. B. Sidi, "Prediction of Course Grades in Computer Science Higher Education Program via a Combination of Loss Functions in LSTM Model," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3351186.
- [7] Y. Huo, M. Jin, and S. You, "LSTM-Based Framework for the Synthesis of Original Soundtracks," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3372581.
- [8] Y. Pande, "Project EngiBot: Engineering Insights through NLP- driven Chatbot," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 11, 2023, doi: 10.22214/ijraset.2023.56590.
- [9] Dr. S. M. Patil, "STES Chatbot Using Flutter: STES Dialogic AI," *Int J Res Appl*

- Sci Eng Technol*, vol. 11, no. 11, 2023, doi: 10.22214/ijraset.2023.56507.
- [10] Q. D. L. Tran and A. C. Le, "A deep reinforcement learning model using long contexts for chatbots," in *Proceedings of 2021 International Conference on System Science and Engineering, ICSSE 2021*, 2021. doi: 10.1109/ICSSE52999.2021.9538427.
- [11] M. Dhyani and R. Kumar, "An intelligent Chatbot using deep learning with Bidirectional RNN and attention model," in *Materials Today: Proceedings*, 2019. doi: 10.1016/j.matpr.2020.05.450.
- [12] P. Anki, A. Bustamam, H. S. Al-Ash, and D. Sarwinda, "Intelligent Chatbot Adapted from Question and Answer System Using RNN-LSTM Model," in *Journal of Physics: Conference Series*, 2021. doi: 10.1088/1742-6596/1844/1/012001.
- [13] D. Udayan, D. B. N. Krishna, T. Reddy, and L. Dinesh, "Conversational Chatbot for College Management Using LSTM," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4027039.
- [14] N. Cannavaro, "Aplikasi Chatbot untuk Layanan Akademik Menggunakan Platform RASA Open Source dengan Fitur Two Stage Fallback," *Jurnal Ilmu Komputer dan Informatika*, vol. 3, no. 1, 2023, doi: 10.54082/jiki.73.
- [15] A. A. Chandra, V. Nathaniel, F. R. Satura, and F. D. Adhinata, "Pengembangan Chatbot Informasi Mahasiswa Berbasis Telegram dengan Metode Natural Language Processing," *Journal ICTEE*, vol. 3, no. 1, 2022, doi: 10.33365/jictee.v3i1.1886.
- [16] R. E. Alden, H. Gong, E. S. Jones, C. Ababei, and D. M. Ionel, "Artificial Intelligence Method for the Forecast and Separation of Total and HVAC Loads with Application to Energy Management of Smart and NZE Homes," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3129172.
- [17] R. Fitri, *Pemrograman Basis Data Menggunakan MySQL*. 2020.
- [18] Z. H. Zhou, *Machine Learning*. 2021. doi: 10.1007/978-981-15-1967-3.
- [19] Lolanda Hamim Annisa and Y. H. C. Pratama, "Implementasi Paradigma Interaksi Manusia & Komputer Pada di Era Society 5.0: Systematic Literature Review," *Technology and Informatics Insight Journal*, vol. 1, no. 2, 2022, doi: 10.32639/tij.v1i2.183.
- [20] T. H. Binh, D. B. Son, H. Vo, B. M. Nguyen, and H. T. T. Binh, "Reinforcement Learning for Optimizing Delay-Sensitive Task Offloading in Vehicular Edge-Cloud Computing," *IEEE Internet Things J*, vol. 11, no. 2, 2024, doi: 10.1109/JIOT.2023.3292591.
- [21] E. Dwi Pratama, "Implementasi Model Long-Short Term Memory (LSTM) pada Klasifikasi Teks Data SMS Spam Berbahasa Indonesia," *The Journal on Machine Learning and Computational Intelligence (JMLCI)*, vol. 1, no. 2, 2022.
- [22] A. M. A. Sai, O. Balamurali, M. Karthikeya, and S. Anand, "A Web-Based Chatbot for Indian Cities: A Comparison of CNN, ANN, and LSTM Models," in *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*, 2023. doi: 10.1109/ICCCNT56998.2023.10307912.
- [23] F. A. Al Farisi, R. S. Perdana, and P. P. Adikara, "Klasifikasi Intensi dengan Metode Ling Short-Term Memory pada Chatbot Bahasa Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 7, 2023, doi: 10.25126/jtiik.1078000.
- [24] A. W. Sugiyarto and A. M. Abadi, "Prediction of Indonesian palm oil production using long short-term memory recurrent neural network (LSTM-RNN)," in *Proceedings - 2019 1st International Conference on Artificial Intelligence and Data Sciences, AiDAS 2019*, 2019. doi: 10.1109/AiDAS47888.2019.8970735.
- [25] Y. L. Hsueh and T. L. Chou, "A Task-oriented Chatbot Based on LSTM and Reinforcement Learning," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, no. 1, 2022, doi: 10.1145/3529649.
- [26] D. Apriliani, S. F. Handayani, and I. T. Saputra, "Implementasi Natural Language Processing (NLP) Dalam Pengembangan Aplikasi Chatbot Pada SMK YPE Nusantara Slawi," *Techno.Com*, vol. 22, no. 4, 2023, doi: 10.33633/tc.v22i4.9155.
- [27] J. Y. Lee, "Can an artificial intelligence chatbot be the author of a scholarly article?," 2023. doi: 10.6087/kcse.292.
- [28] N. Lhasiw, N. Sanglerdsinlapachai, and T. Tanantong, "A Bidirectional LSTM Model for Classifying Chatbot Messages," in *16th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2021*, 2021. doi: 10.1109/iSAI-NLP54397.2021.9678173.
- [29] M. Ilyas Tri Khaqiqi, N. H. Harani, and C. Prianto, "Performance Analysis and Development of QnA Chatbot Model Using LSTM in Answering Questions," *The Indonesian Journal of Computer Science*,

- vol. 12, no. 3, 2023, doi:  
10.33022/ijcs.v12i3.3249.
- [30] M. A. Nugroho, A. Damayanti, M. F. Rifai,  
and S. Windarti, "Pengembangan Aplikasi  
Qna Untuk Pendaftaran Mahasiswa Baru

Stmik Akakom," *Journal of Information  
System Management (JOISM)*, vol. 3, no. 1,  
2021, doi: 10.24076/joism.2021v3i1.408.