

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Seiring dengan perkembangan teknologi, *Artificial Intelligence* (AI) telah menjadi sistem komputer yang mampu melakukan tugas-tugas yang biasanya membutuhkan kecerdasan manusia [1]. Salah satu kemampuan AI yang semakin berkembang adalah berkomunikasi seperti manusia melalui *Natural Language Processing* (NLP), cabang AI yang berkaitan dengan interaksi antara mesin dan manusia menggunakan bahasa alami. NLP memiliki berbagai penerapan seperti penerjemah bahasa (*Google Translate*), asisten personal (Siri), pemeriksa keakuratan bahasa (*Grammarly*). Chatbot, sebagai salah satu penerapan NLP, dirancang untuk berkomunikasi dengan manusia melalui percakapan daring, menjadikannya solusi menjanjikan di berbagai konteks, termasuk sistem informasi akademis.

Dalam sistem informasi akademik, kemampuan untuk memberikan respons cepat dan akurat terhadap pertanyaan mahasiswa sangat penting. Chatbot berpotensi besar untuk meningkatkan efisiensi dan efektivitas dalam penyampaian informasi akademik, mengatasi keterbatasan metode konvensional yang seringkali kurang responsif. Namun, tantangan utama dalam pengembangan chatbot ini adalah memastikan bahwa chatbot dapat memahami dan merespons pertanyaan mahasiswa dengan tepat, terutama mengingat keragaman dan kompleksitas pertanyaan yang diajukan.

Penelitian mengenai chatbot telah berkembang pesat. Contohnya, penelitian yang dilakukan oleh Tjut Awaliyah Zuraiyah, Dian Kartika Utami, dan Degi Herlambang mengenai chatbot untuk pendaftaran mahasiswa baru menggunakan *Recurrent Neural Network*, total data yang ia gunakan berjumlah 300 data lalu dibagi menjadi data latih sebanyak 250 data dan 50 data digunakan sebagai data uji menunjukkan hasil akurasi 88%, presisi 95%, dan recall 92% [1].

Pada tahun 2021, Vincentius Riandaru Prasetyo, Njoto Benarkah, dan Vioni Jannet Chrisintha dari Universitas Surabaya mengembangkan chatbot berbasis NLP untuk membantu pengguna mendapatkan informasi tentang program teknologi

informasi, mencapai akurasi di atas 75% [2]. Penelitian Muhammad Rahaji Jhaerol dan Sudianto Sudianto menunjukkan bahwa algoritma LSTM yang digunakan pada chatbot Program Merdeka Belajar Kampus Merdeka memperoleh akurasi 100% dan F1-score 100% [3]. Studi lain oleh M.-H. Su et al. dalam "A Chatbot Using LSTM-based Multi-Layer Embedding for Elderly Care" menunjukkan akurasi 79,96% untuk top-1 dan 94,85% untuk top-10 pasangan pesan[4] .

Keunggulan kinerja model RNN dengan arsitektur LSTM dalam pengenalan aktivitas juga telah dibuktikan dalam penelitian Singh dkk [5] Dalam penelitian tersebut mengusulkan model Convolutional Neural Network (CNN) dan model RNN untuk memodelkan pengenalan aktivitas. Model CNN digunakan karena mampu menangkap (*capture*) beberapa pergerakan aktivitas melalui ekstraksi fitur. Berdasarkan hasil eksperimen, kedua model *Deep Learning* CNN dan RNN menghasilkan kinerja yang baik. Secara keseluruhan eksperimen yang dilakukan, arsitektur LSTM menunjukkan kinerja yang lebih baik dibandingkan CNN, dengan perbandingan akurasi 89,8% (LSTM) dan 88,2% (CNN).

Algoritma *Long Short-Term Memory* (LSTM) dipilih dalam penelitian ini karena kemampuannya yang unggul dalam menangkap dependensi jangka panjang dalam data teks, yang esensial untuk memahami pertanyaan dengan konteks yang kompleks. Berdasarkan hal tersebut, penelitian ini bertujuan untuk menguji performa algoritma LSTM dalam konteks chatbot yang dirancang khusus untuk sistem informasi akademik. Chatbot ini berfungsi sebagai media untuk mengevaluasi seberapa baik algoritma LSTM dapat beroperasi dalam memberikan respons yang akurat dan efektif terhadap pertanyaan mahasiswa.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang masalah, maka rumusan masalahnya adalah bagaimana performa algoritma *Long Short-Term Memory* (LSTM) ketika diuji dalam konteks chatbot untuk sistem informasi akademik.

1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk mengimplementasikan algoritma *Long Short-Term Memory* (LSTM) pada chatbot informasi akademik. Dengan tujuan yang akan dicapai adalah untuk mengukur seberapa baik performa algoritma tersebut berdasarkan indikator seperti akurasi, serta ketepatan klasifikasi yang memprediksi kategori pertanyaan oleh model terhadap pertanyaan yang diajukan oleh pengguna.

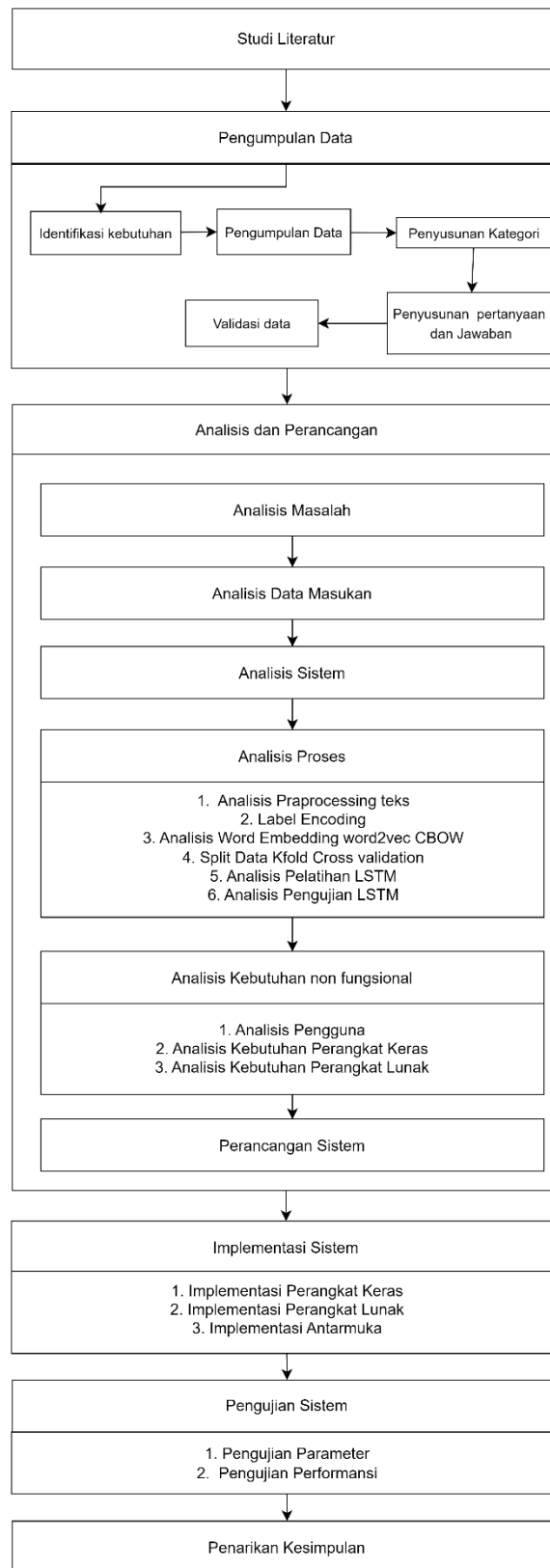
1.4 Batasan Masalah

Adapun Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Sistem chatbot yang digunakan berbasis teks yang berbahasa indonesia
2. Bahasa Pemograman yang digunakan ialah python.
3. Penelitian ini berfokus pada informasi yang relevan dengan Program Studi Teknik Informatika Universitas Komputer Indonesia (UNIKOM), seperti data jadwal kuliah, identitas dosen, kalender akademik, berkas skripsi, berkas Kerja Praktik, berkas akademik, informasi lokasi ruangan dan informasi mengenai fasilitas yang berlaku pada tahun akademik 2023–2024.
4. Model yang dibangun menggunakan arsitektur LSTM.
5. Metode pembobotan data terfokus pada penggunaan Word2Vec (CBOW).
6. Pembagian data menggunakan metode *K-fold Cross validation*.

1.5 Metode Penelitian

Metode penelitian digunakan sebagai panduan atau referensi dalam pelaksanaan penelitian agar hasil yang diperoleh sesuai dengan tujuan yang ingin dicapai. Metode penelitian yang akan diterapkan dalam penelitian ini dapat dilihat pada Gambar 1.1.



Gambar 1. 1 Metode Penelitian

1.5.1. Studi Literatur

Studi literatur ini merupakan teknik pengumpulan data dengan meninjau beberapa e-book, makalah, tesis, skripsi, dan jurnal terdahulu yang berkaitan dengan evaluasi performa algoritma LSTM dan penerapan chatbot serupa dalam konteks sistem informasi akademik.

1.5.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari berbagai sumber sekunder yang dapat diakses publik dan berkaitan dengan program studi Teknik Informatika Universitas Komputer Indonesia (UNIKOM). Proses pengumpulan data mencakup beberapa langkah penting:

1. Identifikasi Kebutuhan

Langkah awal adalah mengidentifikasi kebutuhan informasi mahasiswa melalui diskusi dan refleksi pribadi. Dari proses ini, ditemukan beberapa poin utama yang sering ditanyakan oleh mahasiswa, seperti Identitas dan Kontak Dosen, Lokasi Ruangan Kuliah, Jadwal Kuliah, Persyaratan dan Prosedur Akademik, Kalender Akademik, dan Daftar Mata Kuliah.

2. Pengumpulan informasi

Informasi data dikumpulkan dari berbagai sumber eksternal yang dapat diakses publik, termasuk website resmi UNIKOM dan Teknik Informatika, media sosial resmi, serta publikasi terkait. Data yang dikumpulkan mencakup informasi seperti nama dosen, jadwal kuliah, lokasi ruangan, dan persyaratan akademik lainnya.

3. Penyusunan Label Kategori / TAG

Setelah informasi data terkumpul, kemudian dipecah menjadi kategori yang lebih spesifik, yang nantinya akan digunakan dalam penyusunan pertanyaan dan jawaban chatbot.

4. Penyusunan Pertanyaan & Jawaban

Kategori yang telah dibentuk kemudian digunakan untuk menyusun pertanyaan dan jawaban yang relevan, memastikan keakuratan dan relevansi informasi.

5. Validasi Data

Validasi data ini dilakukan secara manual untuk memastikan akurasi dan kesesuaian data dengan sumber aslinya.

1.5.3. Analisis

Tahapan analisis dimulai dengan identifikasi masalah, yaitu kurangnya efisiensi dan responsivitas dalam penyediaan informasi akademik yang masih dilakukan secara konvensional. Analisis data dilakukan untuk memahami jenis informasi yang paling sering dicari oleh mahasiswa. Selain itu, dilakukan pemetaan struktur data dan kategorisasi pertanyaan untuk memastikan model LSTM yang dibangun dapat memahami konteks pertanyaan dengan baik dan memberikan jawaban yang relevan. Analisis ini juga mencakup evaluasi metode yang digunakan dalam penelitian, termasuk praproses data, teknik pembobotan menggunakan Word2Vec (CBOW), dan penilaian performa model LSTM.

1.5.4. Implementasi

Pada tahap ini, model LSTM dikembangkan dan dilatih menggunakan data yang telah dianalisis. Implementasi mencakup pemrograman model LSTM, pelatihan model dengan data yang tersedia, serta integrasi model dengan sistem chatbot yang akan digunakan. Hasil dari tahap implementasi ini adalah terbentuknya simulator chatbot yang akan digunakan pada tahap pengujian dan evaluasi selanjutnya.

1.5.5. Pengujian/Evaluasi Model

Pengujian dan evaluasi model dilakukan dalam dua langkah utama. Pertama, pengujian parameter dilakukan untuk menentukan kombinasi parameter terbaik dari model LSTM yang menghasilkan nilai akurasi validasi tertinggi. Parameter yang diuji mencakup dimensi *hidden*, jumlah *epoch*, *batch size*, dan *learning rate*. Kedua, pengujian performansi dilakukan dengan cara menguji model yang telah dilatih menggunakan parameter terbaik tadi pada pengguna (*user*). Dalam pengujian ini, model diberikan pertanyaan oleh pengguna, dan hasil jawaban dari model dievaluasi untuk melihat apakah sesuai dengan pertanyaan yang diajukan atau tidak.

1.5.6. Penarikan Kesimpulan

Tahap terakhir adalah penarikan kesimpulan dari penelitian yang telah dilakukan. Pada tahap ini, akan dirangkum hasil dari semua tahapan sebelumnya, membahas temuan-temuan utama, serta memberikan rekomendasi untuk penelitian selanjutnya. Kesimpulan ini juga mencakup evaluasi apakah tujuan penelitian telah tercapai dan kontribusi penelitian terhadap bidang ilmu terkait.

1.6 Sistematika Penulisan

BAB 1 PENDAHULUAN

Bab 1 menguraikan tentang latar belakang yang terdiri dari permasalahan yang berkaitan dengan penelitian, rumusan masalah, maksud dan tujuan penelitian, batasan masalah, metode penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab 2 membahas dan menjelaskan mengenai landasan teori yang mendukung penulisan skripsi, mulai dari *literature review*, kecerdasan buatan, chatbot, Word2Vec, LSTM, *K-Fold Cross Validation*.

BAB 3 ANALISIS

Bab 3 berisikan tentang analisis terhadap model yang akan dibangun meliputi analisis masalah, analisis data masukan, analisis pada pemrosesan, dan analisis metode.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisikan hasil implementasi dari analisis terhadap model yang sudah dibangun. Pengujian juga dilakukan untuk mengetahui tingkat performansi dari metode yang digunakan

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan mengenai hasil penelitian dan saran yang diharapkan akan berguna untuk penelitian lebih lanjut pada kasus yang sama