

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Di era modern saat ini, pembelajaran mandiri bahasa Jepang tingkat dasar, khususnya untuk standar *Japanese-Language Proficiency Test* (JLPT), mengalami perkembangan pesat seiring meluasnya akses terhadap materi edukasi berbasis teknologi (*What*). Secara ideal, penguasaan bahasa kedua bagi pembelajar pemula memerlukan pemahaman terstruktur yang mengintegrasikan tiga elemen utama secara hierarkis, yaitu karakter Kanji, kosakata (*vocabulary*), dan tata bahasa (*grammar*). Dalam psikologi pendidikan, pengorganisasian informasi secara sistematis di dalam kognisi pembelajar sangat krusial untuk membangun fondasi pemahaman yang bermakna. Pembelajaran mandiri yang efektif menuntut adanya panduan atau alur instruksional yang ketat agar pelajar dapat mengaitkan relasi antar-elemen bahasa tersebut dengan tepat, tanpa mengalami kebingungan arah atau disorientasi kurikulum.

Namun, kenyataan di lapangan menunjukkan bahwa para pembelajar mandiri sering kali menghadapi hambatan besar akibat kompleksitas struktural silabus bahasa Jepang yang memiliki dependensi sangat tinggi (*Why*). Pelajar tidak akan bisa memahami suatu kosakata tanpa menguasai Kanji penyusunnya, dan tidak dapat menerapkan suatu struktur tata bahasa tanpa memahami kategori kata kerja atau kata sifat yang membentuknya. Masalah ini diperparah oleh keterbatasan media belajar konvensional seperti buku teks statis atau aplikasi

hafalan satu arah yang membuat proses belajar terasa membosankan dan menurunkan motivasi. Di sisi lain, pelacakan kurikulum berbasis basis data relasional biasa (RDBMS) juga tidak efisien karena membutuhkan operasi komputasi yang sangat kompleks untuk memetakan hubungan hierarkis tersebut, adapun teknologi Artifisial berupa model *Large Language Model* yang ternyata memiliki banyak sekali celah dalam merangkai materi bahasa Jepang yang ada di database digital. Akibatnya, terjadi fragmentasi informasi di mana pelajar terjebak dalam pemahaman yang dangkal tanpa mampu melacak rantai pengetahuannya sendiri.

Untuk mengatasi hambatan belajar mandiri tersebut, banyak pembelajar mencoba beralih menggunakan *AI generatif* sebagai tutor virtual pengganti (*Who & Where*). Namun, pemanfaatan Large Language Models (LLM) murni justru menimbulkan masalah baru berupa halusinasi semantik akibat ketiadaan kontrol kurikulum eksternal. Guna memvalidasi fenomena ini, penulis melakukan studi kasus dan wawancara dengan Anna-sensei, pengajar dari Jasmine Studi Budaya Bahasa Jepang di Surabaya. Menurut Anna-sensei, penggunaan AI generatif murni bagi pembelajar otodidak sangat berisiko: ini dikarenakan kosakata yang diberikan sering salah, struktur kalimat kerap tidak tepat, bahkan arti dan cara baca Kanji keliru. Hal ini terjadi akibat bentrok data pelatihan AI global tanpa kurasi kurikulum yang jelas, dan juga sumber antar materi yang tersebar di seluruh database digital yang saling kontras satu sama lain. Temuan empiris ini memperkuat urgensi perlunya tutor AI yang mampu mengeliminasi inkonsistensi tersebut demi menjamin keandalan materi.

Secara pedagogis, kelemahan LLM murni tersebut berdampak fatal karena sifat stokastik pada arsitektur dasarnya rentan mencampurkan materi tingkat menengah kepada pembelajar N5 (*When*). Selain itu, sistem bimbingan yang masih bersifat pasif berbasis teks kehilangan elemen interaksi multimodal untuk melatih kemampuan mendengar serta berbicara, sehingga gagal membangun keterlibatan aktif dan daya kritis dari sisi pembelajar. Oleh karena itu, diperlukan sebuah terobosan hibrida yang menggabungkan keluwesan komunikasi LLM dengan kepastian struktur pengetahuan untuk mengunci batasan silabus secara ketat.

Secara teknis, solusi ini diwujudkan ke dalam sebuah arsitektur agen edukasi cerdas (*intelligence tutor system*) lokal yang mengintegrasikan berbagai pipa teknologi mutakhir secara asinkron (*How*). Sistem backend berbasis FastAPI akan mengolah masukan pengguna melalui proses tokenisasi, lalu melakukan pencarian graf semantik pada Neo4j untuk disuntikkan ke dalam instruksi model bahasa Qwen via metode GraphRAG, model akan mematuhi segala instruksi berupa prompt, dan merangkai kata atau rangkaian materi berdasarkan aturan pedagogis yang sudah ditentukan oleh sistem. Guna mendongkrak motivasi dan interaksi oral, sistem dilengkapi mesin Speech-to-Text, Text-to-Speech, animasi *lip-sync* avatar 3D, serta metode gamifikasi (XP dan lencana) di basis data Supabase. Melalui ekosistem *Neuro-Symbolic AI* ini, tutor virtual mampu mempertahankan urutan instruksional yang logis, adaptif, dan aman secara akademis. Berdasarkan seluruh urgensi tersebut, penelitian ini

diberi judul "*Sistem Tutor Pembelajaran Bahasa Jepang dengan GraphRAG Berbasis Neuro-Symbolic AI*".

B. Pembatasan Masalah

Demi menghindari ruang lingkup penelitian agar tidak meluas, penelitian ini dibatasi pada hal-hal berikut :

1. Penelitian ini hanya difokuskan pada pengembangan dan pengujian tutor virtual bahasa Jepang untuk materi dasar tingkat JLPT N5 dengan acuan silabus atau *Nihongo Sou Matome* N5.
2. Penelitian ini membatasi representasi pengetahuan pada interaksi struktural Kanji, Kosakata, dan Tata Bahasa dalam *Knowledge Graph*.
3. Penelitian ini tidak mengevaluasi seluruh aspek pedagogis pembelajaran bahasa Jepang, tetapi hanya pada fitur tutor virtual yang dirancang dalam ruang lingkup sistem.
4. Penelitian ini tidak mencakup proses pelatihan ulang (*fine-tuning*) pada arsitektur model bahasa dasar yang digunakan, melainkan berfokus pada metode rekayasa *prompt* (*prompt engineering*).
5. Penelitian ini fokus pada optimasi sistem pada device spesifikasi Vram Terbatas, dengan hanya menggunakan Model LLM berukuran kecil.

C. Perumusan Masalah

Berdasarkan latar belakang masalah diatas dapat diambil rumusan masalah sebagai berikut:

1. Bagaimana merancang Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*?

2. Bagaimana mengimplementasikan Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*?
3. Bagaimana mengevaluasi Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*?

D. Tujuan Penelitian

Berdasarkan uraian latar belakang, pembatasan masalah dan perumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Merancang Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*.
2. Mengimplementasikan Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*.
3. Menguji Sistem Tutor Pembelajaran Bahasa Jepang Berbasis *Neuro-Symbolic AI*.

E. Kegunaan Penelitian

1. Kegunaan Teoritis

- a. Bagi pengembangan ilmu keinformatikaan

Penelitian ini memberikan kontribusi terhadap pengembangan ilmu keinformatikaan, khususnya dalam sub-bidang Kecerdasan Buatan (*Artificial Intelligence*) dan Sistem Pembelajaran Cerdas (*Intelligent Tutoring System*). Riset ini menyumbangkan pemikiran baru mengenai metode integrasi *Neuro-symbolic* hibrida yang menggabungkan fleksibilitas pemrosesan bahasa alami dari *Large Language Models*

(LLM) dengan kepastian data dari struktur *Knowledge Graph* (KG) untuk mitigasi halusinasi informasi.

b. Bagi universitas dan program studi

Hasil penelitian ini dapat dijadikan sebagai referensi ilmiah dan praktis bagi universitas serta program studi Teknik Informatika dalam mengimplementasikan serta mengembangkan teknologi *Generative AI* tingkat lanjut. Luaran riset ini juga dapat memperkaya literatur riset lokal mengenai pemanfaatan arsitektur *Knowledge Graph-Enhanced Retrieval-Augmented Generation* (KG-RAG) dalam domain EdTech (*Educational Technology*).

c. Bagi peneliti selanjutnya

Penelitian ini dapat menjadi dasar atau acuan untuk penelitian lanjutan di bidang tutor virtual edukasi cerdas. Peneliti berikutnya dapat mengembangkan sistem ini agar lebih interaktif lagi, seperti integrasi dengan teknologi *Text-to-Speech* (TTS) untuk latihan percakapan verbal, perluasan cakupan kurikulum ke level JLPT yang lebih tinggi (N4 hingga N1), ataupun penerapan umpan balik adaptif berbasis deteksi emosi pembelajar.

2. Kegunaan Praktis

a. Bagi objek

Penelitian tentang sistem tutor virtual ini dapat membantu pembelajar mandiri bahasa Jepang level dasar (JLPT N5) sebagai objek utama untuk mendapatkan lingkungan belajar yang aman, terarah, dan terstruktur.

Melalui visualisasi relasi konsep dan metode perancah (scaffolding) Sokratik, pengguna dapat memahami korelasi antara Kanji, kosakata, dan tata bahasa secara runut tanpa risiko terpapar materi yang melompati prasyarat kompetensinya.

b. Bagi Masyarakat

Khususnya lingkungan institusi pendidikan formal maupun lembaga kursus bahasa asing, penelitian ini memberikan contoh konkret tentang pentingnya pemanfaatan kecerdasan buatan yang tervalidasi kurikulum untuk membantu guru. Sistem ini dapat digunakan sebagai asisten virtual yang meringankan beban pengajar dalam memberikan pengayaan materi, pembuatan latihan soal adaptif, serta pemberian umpan balik personal kepada banyak siswa sekaligus secara konsisten.