

**CHATBOT MENGGUNAKAN MODEL ALGORITMA BERT PADA FITUR CHAT
LAYANAN PENDIDIKAN**



Disusun Oleh:

Dio Rakasiwi 1197050032

JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS ISLAM NEGERI SUNAN
GUNUNG DJATI
BANDUNG 2023

LEMBAR PENGESAHAN
CHAT BOT MENGGUNAKAN MODEL BERT PADA FITUR CHAT LAYANAN
PENDIDIKAN

Disusun Oleh:

Dio Rakasiwi

1197050032

Telah diseminarkan dan disahkan pada tanggal :

Mengesahkan,

Ketua Jurusan Teknik Informatika

Dosen Penguji

Cepy Slamet, Ph.D.

NIP. 198002252011011007

Mohamad Irfan, ST., M.Kom. Ph.D

NIP. 198310232009121005

DAFTAR ISI

LEMBAR PENGESAHAN	i
DAFTAR ISI.....	ii
PENDAHULUAN	3
1.1 Latar Belakang.....	3
1.2 Rumusan Masalah.....	4
1.3 Batasan Penelitian.....	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 <i>The State of The Art</i>	5
1.7 Kerangka Pemikiran	7
1.8 Metodologi Penelitian.....	8
1.9 Jadwal dan Lokasi Penelitian.....	12
DAFTAR PUSTAKA	Error! Bookmark not defined.
LAMPIRAN.....	Error! Bookmark not defined.

PENDAHULUAN

1.1 Latar Belakang

Secara umum, komunikasi adalah suatu proses penyampaian dan penerimaan pesan antar dua orang atau lebih. Bentuk pesan dapat berupa komunikasi tulisan atau teks, komunikasi verbal, dan komunikasi non-verbal. Pada saat ini layanan pendidikan menjadi salah satu kebutuhan yang penting bagi para pelajar [1]. Komunikasi dalam bentuk teks merupakan suatu proses penyampaian pesan komunikasi dengan menggunakan kata-kata yang memiliki makna tertentu. Sistem komunikasi merupakan salah satu solusi untuk membangun sistem layanan yang baik pada Pendidikan karena mampu memberikan media virtual interaksi [2].

Chatbot adalah salah satu aplikasi atau layanan yang berinteraksi dengan pengguna melalui percakapan teks. Chatbot bekerja untuk menggantikan manusia dalam melayani pembicaraan melalui aplikasi pesan [3]. Salah satu yang diterapkan chatbot dengan adanya Question and Answer pada layanan pendidikan.

Question and Answering adalah salah satu cara yang bisa diterapkan pada sistem komunikasi dan penyampaian informasi. Adanya sistem ini dapat memudahkan pihak administrasi untuk menyampaikan informasi melalui komputer. Komputer dapat memberi informasi dan menjawab pertanyaan secara online yang diberikan oleh penggunanya [4].

Secara Diskusi yang dilakukan pada siswa di Yayasan Thoriqul Jannah sering mengalami kesulitan untuk menanyakan atau menerima jawaban dari pertanyaan perihal layanan pendidikan karena tidak ada ruang forum yang mampu mewadahi hal tersebut. Berdasarkan diskusi para guru Yayasan Thariqul Jannah, banyaknya pertanyaan umum yang masuk dari para siswa menanyakan pertanyaan yang sering ditanya oleh siswa lainnya. Hal tersebut pada layanan bidang pendidikan siswa sering memberikan jawaban yang sama berulang kali kepada siswa yang menyebabkan informasi kurang efektif dan tidak beraturan.

Untuk menjawab pertanyaan pada permasalahan yang sering dikeluhkan para siswa maka diperlukan adanya sistem yang mampu menerima pertanyaan dan mampu memberikan jawaban dari pertanyaan tersebut. Dengan dibangunnya sistem otomatis agar jika sistem menerima pertanyaan yang ditanyakan siswa dapat langsung diberikan jawabannya dengan cepat.

Sistem ini ditargetkan untuk pengguna siswa yang ingin bertanya tentang layanan pendidikan. Dengan adanya sistem ini dapat membantu menjawab pertanyaan dari siswa dan mampu memberikan jawaban secara otomatis. Dalam penelitian ini akan membangun sistem yang mampu membantu kebutuhan tersebut dengan penerapan pada fitur chat layanan pendidikan.

Model yang digunakan dalam membangun sistem chatbot ialah model BERT. BERT berguna untuk mengolah representasi dua arah yang berada dalam teks tanpa nama dengan menggabungkan sisi kanan dan kiri pada sebuah konteks dalam segala bagian.

Berdasarkan permasalahan tersebut maka penelitian ini akan membuat sistem Question and Answering menggunakan model BERT memberikan jawaban atas pertanyaan yang diberikan dan menyimpulkan sebuah perintah/bahasa Indonesia pada fitur chat both layanan pendidikan siswa. Dengan adanya penelitian ini dapat membantu siswa menyiapkan wadah untuk membantu menjawab pertanyaan mengenai layanan pendidikan [5].

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas, maka dapat di rumuskan permasalahan pada penelitian adalah sebagai berikut :

1. Bagaimana mengimplementasikan algoritma BERT untuk melakukan fitur chatbot menggunakan fitur chat layanan Pendidikan?
2. Bagaimana kinerja algoritma BERT dalam menjalankan fitur chatbot pada layanan Pendidikan?

1.3 Batasan Penelitian

Pada penelitian ini ditetapkan beberapa batasan masalah, diantaranya :

1. Sistem ini hanya menguji jawaban teks, tidak menguji jawaban berbentuk file gambar dan suara.
2. Sistem both chat hanya dilakukan pada teks yang menggunakan bahasa Indonesia.

1.4 Tujuan Penelitian

Tujuan yang hendak dicapai dalam penelitian ini adalah sebagai berikut :

1. Untuk mengimplemasikan algoritma BERT dalam melakukan chat both pada layanan chat Pendidikan.
2. Untuk mengetahui kinerja algoritma BERT dalam melakukan chatbot pada layanan chat Pendidikan.

1.5 Manfaat Penelitian

Adapun manfaat yang didapatkan dari sistem ini, diantaranya :

1. Membantu Pembina atau admin dalam melakukan komunikasi secara otomatis melalui aplikasi telegram.
2. Dapat memudahkan admin dalam menerima respon dari lawan bicaranya dan memberi pesan balasan merespon chat yang masuk secara otomatis.

1.6 *The State of The Art*

The State of The Art digunakan untuk menelaah hasil penelitian sebelumnya, yang memiliki korelasi atau kesamaan dengan penelitian yang dilakukan saat ini, yang berguna untuk memudahkan admin dalam melihat chat yang masuk secara otomatis. Untuk mendukung penelitian mengenai perancangan sistem tersebut dibutuhkan berbagai kajian literatur diantaranya yaitu akan dijelaskan dalam beberapa paragraph di bawah ini :

1. Pada jurnalnya yang berjudul “Question answering system dalam bentuk chatboth pada platfrom line untuk mata pelajaran sejarah SMA/MA dengan menggunakan algoritma Levenshteing distance” penelitian ini menyatakan bahwa chatbot dapat membantu siswa dalam memberikan jawaban secara benar mengenai mata pelajaran sejarah. Untuk membantu dan memudahkan siswa SMA/MA dalam mencari informasi mengenai mata pelajaran sejarah maka dirancang dan dibangun question answering sytem dalam bentuk chatbot pada platform LINE. Hasil yang diperoleh pada penelitian ini akurasi sebesar 100% dengan mengguna menggunakan pertanyaan yang berasal dari knowledge base dan 72% dengan menggunakan pertanyaan yang ditanyakan langsung oleh siswa SMA/MA [6].
2. Pada jurnalnya yang berjudul “Real Life Application of a Question Answering System Using BERT Language Model” penelitian ini menerapkan salah satu novasi terbaru dalam Deep Learning untuk tugas teks klasifikasi. Tujuannya untuk membuat pertanyaan sistem penjawab dalam bahasa Italia yang memberikan informasi tentang subjek tertentu, e-faktur

dan penagihan digital. Dari hasil penelitian ini memiliki kemampuan untuk menjawab pertanyaan berdasarkan frase dari kata yang ditanyakan memiliki akurasi sebesar 93,6 % yang ditingkatkan dalam dataset pemodelan bahasanya [7].

3. Pada jurnalnya yang berjudul “Question Answering Systems: A Systematic Literature Review” penelitian ini menerapkan untuk membuat Web lebih cocok untuk digunakan manusia dengan menghilangkan kebutuhan untuk menyaring melalui banyak hasil pencarian secara manual untuk menentukan yang benar menjawab pertanyaan, Untuk menjawab permasalahan yang akan diselesaikan oleh question and answering system ini maka akan dilakukan pengujian untuk mengetahui akurasi dari model perancangan. Dari hasil penelitian ini mengidentifikasi delapan puluh studi dimasukkan dalam tinjauan sistematis: 15 adalah kualitatif sedangkan 65 bersifat kuantitatif [8].
4. Pada jurnalnya yang berjudul “Question Answering using Ontology for Sumedang Larang History with Support Vector Machine Based on Telegram Bot” penelitian ini sistem yang dapat menerima pertanyaan dalam bahasa alami, pencarian informasi tentang pertanyaan di basis pengetahuan dan memberikan informasi yang akurat. Selain itu, teknologi dapat dijadikan sebagai daya tarik pembelajaran sejarah dengan sistem tanya jawab (QA). Penerapan algoritma support vector machine (SVM) untuk memproses teks klasifikasi dapat mempermudah pencarian kategori teks. Hasil pengujian menunjukkan performansi metode SVM dengan parameter ukuran pengujian 0,5 yaitu 74% dan 78% [9].
5. Pada jurnal yang berjudul “Mobile-Based Question Answering System Development With Ontology Based Knowledge” penelitian ini, dikembangkan sebuah question answering system (QAS) terkait penerimaan mahasiswa baru agar pengguna mendapatkan informasi yang sesuai dengan kebutuhannya, selalu bernilai benar, dan dapat diakses kapan saja. Hasil pengujian terhadap sistem dengan beberapa cara, yaitu: usability testing, dan pengujian akurasi jawaban. Pengujian menunjukkan QAS yang dibangun dapat diimplementasikan dengan baik sesuai dengan kebutuhan dengan akurasi jawaban sebesar 82.14% [10].
6. Pada jurnal yang berjudul “COBERT: Covid-19 Question Answering System Using BERT” oleh Jafar A. Alzubi, pemodelan ini memiliki akurasi besar 87,3%. “Pendekatan metode transformers untuk deteksi bahasa kasar dalam komentar berita online Indonesia” oleh

adriansyah dwi rendragraha. Pemodelan ini mampu memberikan deteksi dan klasifikasi dari permasalahan tersebut” [4].

Dari beberapa penelitian diatas terdahulu telah disebutkan, penulis mencoba melakukan penelitian dengan mengambil objek tentang Question and Answer menggunakan algoritma BERT. Dengan menggunakan penelitian model BERT mampu memberikan perubahan yang baik dan dapat dilakukan dengan sistem layanan pendidikan yang sudah ada. Manfaat keuntungan peneleitian ini untuk pengembangan aplikasi layanan chat pendidikan agar mampu memiliki fitur chat pada layanan siswa.

1.7 Kerangka Pemikiran

Kerangka pemikiran dari penelitian tugas akhir ini dapat diuraikan sebagai berikut :

Tabel 1.1 Kerangka Pemikiran Sistem

Opportunities
<ul style="list-style-type: none"> - Adanya teknologi untuk melakukan chatbot pada sistem layanan Pendidikan sehingga menjadi wadah pertanyaan pada layanan Pendidikan dan mampu menjawab pertanyaan yang umum ditanyakan. - Adanya kemajuan teknologi informasi untuk mengola teks. - Membantu menjawab tentang layanan Pendidikan secara online
Problem
Chatbot merupakan salah satu solusi yang bisa diterapkan pada sistem komunikasi dan penyampaian informasi. Dan biasanya pengguna chattboot mendapatkan beberapa masalah seperti ada nya kendala seperti tidak dapatnya menjawab semua keluhan dan keinginan konsumen tidak bisa langsung terjawab oleh chatbot dan nantinya konsumen tidak akan merasa puas bahkan bisa menambah masalah baru.
Approach
Algoritma BERT

Software Developments
<i>CRIPS-DM (Cross-Industry Standard Process for Data Mining)</i>
Software Implementation
Dibangun dengan bahasa pemograman Python, PHP, dan database MySQL
Result
Aplikasi dapat menjawab beberapa pertanyaan yang masuk secara otomatis.

1.8 Metodologi Penelitian

1. Tahapan penelitian

Dalam penelitian ini akan meliputi beberapa langkah seperti :

1. Mencari Referensi

Mencari Referensi, yaitu proses melakukan pencarian referensi atau tinjauan literatur mengenai masalah serupa. Pada tahap ini dilakukan juga analisis masalah dan melakukan perumusan untuk solusi permasalahan berdasarkan tinjauan literatur yang ada.

2. Perencanaan Program

Perencanaan program, yaitu hasil rumusan dari analisis sebelumnya dilakukan pembuatan aplikasi sebagai solusi dari permasalahan yang ada. Perencanaan program disini harus mampu memenuhi kebutuhan dari penelitian yang dilakukan

3. Pengujian

Pengujian, yaitu tahapan melakukan testing terhadap aplikasi yang sudah di bangun sebelumnya. Pengujian ini dilakukan untuk memastikan apakah aplikasi sudah mampu berjalan baik untuk menyelesaikan permasalahan atau tidak.

3. Penyusunan Laporan

Penyusunan Laporan, yaitu tahap pendokumentasian dari keseluruhan penelitian. Dalam penyusunan laporan ini membahas tentang keseluruhan proses yang dilakukan selama penelitian.

2. Teknik Penyelesaian Masalah

Untuk penyelesaian masalah ini diterapkan model BERT. Model BERT memanfaatkan dataset sbagai pengetahuan yang dimiliki dari sistem. Adapun dalam rancangan sistem menggunakan model BERT adalah sebagai berikut :

1. Analisis dataset

Tahap ini mengelola dataset yang telah di kumpulkan. Pengelolaan ini bertujuan untuk mengetahui apakah data yang dikumpulkan sudah baik atau terdapat data yang bersifat kosong. Analisis ini dapat dikelompokan apakah data sudah memenuhi kebutuhan yang diperlukan atau tidak.

2. Text Preprocessing

Tahap ini merupakan pembersihan data dari unsur yang tidak digunakan pada proses perancangan model. Tahap ini meliputi case folding, yaitu mengubah huruf dalam data menjadi huruf kecil, cleaning, yaitu menghapus kata dari data yang tidak memiliki hubungan dengan informasi yang digunakan seperti emoji, normalization yaitu perubahan kata yang digunakan untuk mengganti kata yang salah ketik atau tidak baku pada dataset, stopword removal merupakan penghapusan kata kata yang sering muncul namun dianggap kurang penting, dan wordpiece tokenizing yaitu untuk memisahkan kata menjadi kalimat dasar.

3. Analisis BERT

Pada tahap ini dilakukan analisis berdasarkan dataset pretrain BERT yang akan digunakan pada penelitian ini. Model pretrain yang digunakan merupakan model indoBERT yang didapatkan dari penelitian. Model ini dipilih berdasarkan model dataset indoBERT base-Uncased fine-tuned untuk Question and Answering downstream task.

4. Model BERT

Pada tahap ini dilakukan untuk menerapkan, proses BERT dapat melakukan analisis secara otomatis melalui quens and answerling. dan pengguna bisa secara cepat mendapatkan jawaban yang di inginkan. Dan BERT dapat digunakan, seperti sentimen analysis (analisis sentimen), question answerling (penjawab pertanyaan).

3. Pengumpulan Data

Pengumpulan data merupakan salah satu kebutuhan awal untuk melakukan penelitian ini. Data yang dikumpulkan akan di gunakan dalam perancangan sistem, maka pengumpulan data harus dilakukan secara sistematis dan terarah.

3.1 Jenis Data

Sumber data yang diperoleh dalam tugas akhir ini dibagi menjadi dua jenis, yaitu :

a. Data Primer

Data Primer adalah data yang diperoleh langsung. Data berasal dari observasi dan wawancara dengan layanan Pendidikan yang ada di Yayasan Thariqul Jannah.

b. Data Sekunder

Data Sekunder adalah data yang diperoleh secara tidak langsung dari narasumber. Data ini digunakan sebagai pendukung data primer seperti melalui media studi Pustaka dan media informasi lain. Untuk mendukung keabsahannya data sekunder ini diambil dari buku, catatan, serta arsip yang dapat dipercaya kebenarannya

3.2 Metode Pengumpulan Data

Untuk mendapatkan data yang diperlukan dalam tugas akhir ini dibagi menjadi tiga jenis, yaitu :

a. Wawancara

Pengumpulan melalui wawancara diarahkan untuk mendapatkan informasi yang relevan mengenai informasi yang dibutuhkan. Narasumber yang akan diwawancara ialah bagian admin layanan Pendidikan Yayasan Thariqul Jannah, Informasi yang akan menjadi pokok ialah mengenai kebutuhan aplikasi, informasi yayasan yang berada didalam ruang lingkup.

b. Studi Literatur

Pengumpulan melalui studi Pustaka ialah dengan mencari referensi atau pustaka mengenai kajian dengan permasalahan serupa. Referensi ini diambil dari karya tulis ilmiah, jurnal, dan prosiding, serta buku. Untuk meningkatkan kerelevanan informasi yang perlukan maka dilakukan pembatasan pencarian Pustaka yaitu maksimal penelitian yang dilakukan dalam ruang lingkup 5 tahun terakhir.

c. Observasi

Pengumpulan data melalui observasi ialah dengan melakukan pengamatan langsung mengenai keadaan perusahaan yang akan di amati.. Salah satu objek yang menjadi pokok perhatian ialah keadaan di ruang lingkup Yayasan. Sebagai contoh ialah memperhatikan kondisi atau isu yang terjadi diruang lingkup yayasan sebagai pendukung dari fitur layanan Pendidikan.

4. Teknik Analisis

Pendekatan analisis yang dilakukan pada penelitian kali ini ialah teknik kuantitatif. Teknik kuantitatif ini melakukan pendekatan penyelesaian masalah dengan melakukan pengujian terhadap variabel. Untuk menjawab permasalahan yang akan diselesaikan oleh question and answering system ini maka akan dilakukan pengujian untuk mengetahui akurasi dari model perancangan tersebut.

1.9 Jadwal dan Lokasi Penelitian

Untuk melakukan penelitian ini tentunya diperlukan data. Data ini akan di peroleh berdasarkan keadaan di ruang lingkup Yayasan Thoriqul Jannah. Adapun rincian lokasi penelitian ini ialah sebagai berikut :

Lokasi : Yayasan Thoriqul Jannah.

Alamat : Dusun Babakan RT.02 RW.08 Desa Pemulihan Kec. Pamulihan Sumedang.

Adapun jadwal penelitian ini akan dilaksanakan pada gambaran selama Pendidikan berlangsung adalah sebagai berikut :

No	Penelitian	Desember				Januari				Februari				Maret			
		1	2	3	4	1	4	3	4	4	1	2	3	4	1	4	3
1	Mencari referensi dan pengumpulan data																
2	Melakukan analisis dari kebutuhan yang dilakukan																
3	Perancangan aplikasi dan pengujian berdasarkan pemodelan																
4	Penyusunan Laporan Penelitian																

(Tabel 1.1 Jadwal Penelitian)

Dalam tabel 1.1 dapat diketahui bahwa penelitian ini dilakukan berjenjang sesuai dengan sistematika penelitian yang dilakukan. Dalam penelitian ini terdiri dari Langkah untuk mencari referensi dan survey permasalahan, kemudian melakukan analisis dari kebutuhan data yang dilakukan, selanjutnya dilakukan perancangan pemodelan dan pengujian sistem yang akan dibangun, serta penyusunan laporan penelitian.

DAFTAR PUSTAKA

- [1] F. Rayyan, *Pengembangan chatbot untuk aplikasi online chat telegram dengan pendekatan klasifikasi emosi pada teks menggunakan metode indobert-lite*. 2022.
[Online]. Available:
https://repository.uinjkt.ac.id/dspace/handle/123456789/65718%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/65718/1/FAISHAL_RAYYAN-FST.pdf
- [2] A. Siahaan, R. Akmalia, Y. Amelia, T. Wulandari, S. A. F. Hrp, and K. Pasaribu, "Manajemen Peningkatan Mutu Pendidikan," *J. Educ.*, vol. 5, no. 2, pp. 3840–3848, 2023, doi: 10.31004/joe.v5i2.1068.
- [3] E. Larasati Amalia and D. Wahyu Wibowo, "Rancang Bangun Chatbot Untuk Meningkatkan Performa Bisnis," *J. Ilm. Teknol. Inf. Asia*, vol. 13, no. 2, pp. 137–142, 2019.
- [4] J. A. Alzubi, R. Jain, A. Singh, P. Parwekar, and M. Gupta, "COBERT: COVID-19 Question Answering System Using BERT," *Arab. J. Sci. Eng.*, vol. 19, 2021, doi: 10.1007/s13369-021-05810-5.
- [5] Ridho, "Bab II Landasan Teori," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2018.
- [6] A. Setiawan, O. N. Pratiwi, and ..., "Question Answering System Dalam Bentuk Chatbot Pada Platform Line Untuk Mata Pelajaran Sejarah Sma/ma Dengan Menggunakan Algoritma Levenshtein ...," *eProceedings ...*, vol. 8, no. 5, pp. 9794–9802, 2021, [Online]. Available:
<https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15756%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15756/15469>
- [7] F. Alloatti, L. Di Caro, and G. Sportelli, "Real life application of a question answering system using bert language model," *SIGDIAL 2019 - 20th Annu. Meet. Spec. Interes. Gr. Discourse Dialogue - Proc. Conf.*, no. January, pp. 250–253, 2019, doi: 10.18653/v1/w19-5930.

- [8] S. S. A. N. Elfadil, M. Jarajreh, and S. Algarni, "Question Answering Systems: A Systematic Literature Review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 3, pp. 495–502, 2021, doi: 10.14569/IJACSA.2021.0120359.
- [9] E. S. B. Perangin-Angin, Z. K. A. Baizal, and D. Richasdy, "Question Answering using Ontology for Sumedang Larang History with Support Vector Machine Based on Telegram Bot," *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 2438, 2022, doi: 10.30865/mib.v6i4.4574.
- [10] R. A. C and I. W. W. Wisesa, "Pengembangan Mobile based Question Answering System dengan Basis Pengetahuan Ontologi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 4, p. 693, 2020, doi: 10.25126/jtiik.2020742255.