

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat telah merambah berbagai sektor, termasuk bidang Kesehatan (Prasanti & Indriani, 2018). Transformasi digital dalam layanan kesehatan menjadi semakin penting dengan meningkatnya jumlah pengguna internet di Indonesia. Menurut Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), jumlah pengguna internet di Indonesia pada tahun 2024 diperkirakan mencapai 221.563.479 jiwa, dari total populasi 278.696.200 jiwa pada tahun 2023. Dengan demikian, lebih dari 79% masyarakat Indonesia memiliki kemampuan dalam menggunakan internet di era modern ini. Akses terhadap informasi kesehatan yang akurat dan layanan medis yang cepat menjadi faktor krusial dalam peningkatan kualitas kesehatan masyarakat (Strekalova, 2014).

Namun, meskipun infrastruktur digital semakin berkembang, tantangan dalam layanan kesehatan konvensional seperti ketersediaan tenaga medis dan biaya pelayanan tetap menjadi hambatan utama. Salah satu permasalahan yang dihadapi adalah keterbatasan jumlah tenaga medis, terutama dokter spesialis, yang menyebabkan waktu tunggu konsultasi menjadi lama dan menyulitkan masyarakat dalam mengakses layanan kesehatan, baik karena sulit memperoleh informasi medis yang akurat, jauhnya fasilitas kesehatan, maupun tingginya biaya konsultasi (Anandan dkk, 2022). Selain itu, tingginya biaya konsultasi dokter dan perawatan kesehatan juga menjadi kendala utama bagi banyak masyarakat dalam memperoleh layanan yang mereka butuhkan (Yunianto, 2023).

Bidang kesehatan gigi dan mulut merupakan salah satu sektor yang terdampak, karena sebagian besar tenaga medisnya terkonsentrasi di wilayah perkotaan, sehingga masyarakat di daerah dengan fasilitas kesehatan yang tidak memiliki layanan tersebut seringkali kesulitan mendapatkan pemeriksaan atau konsultasi (Sinaga, 2022). Selain itu, rendahnya kesadaran masyarakat Indonesia terhadap pentingnya kesehatan gigi dan

mulut masih menjadi persoalan yang signifikan di berbagai wilayah. Survei Kesehatan Indonesia mencatat bahwa meskipun 56,9% masyarakat mengaku mengalami masalah kesehatan gigi dalam satu tahun terakhir, hanya 11,2% dari mereka yang memutuskan untuk memeriksakan diri ke tenaga medis (Media Indonesia, 2023). Kondisi ini berdampak pada rendahnya praktik perawatan gigi yang benar, di mana hanya 2,8% masyarakat menyikat gigi dengan benar, dan sebanyak 88% mengalami karies gigi (Ekaptiningrum, 2023).

Misalnya, di Dusun Pendul, Kabupaten Bantul, mayoritas masyarakat memiliki tingkat pengetahuan kesehatan gigi pada kategori sedang (65%) hingga rendah (22%), menandakan kurangnya edukasi dan perhatian terhadap kesehatan mulut (Ramadhan dkk, 2023). Temuan serupa juga terjadi di wilayah kerja Puskesmas Leuwigajah, Cimahi, di mana tingginya kasus penyakit pulpa dan periapikal disebabkan oleh minimnya kesadaran akan pentingnya perawatan gigi (Himawati, 2023). Tidak hanya itu, di Kota Tegal, masyarakat masih mengalami kebingungan dalam mengenali gejala penyakit gigi serta kurang memahami hak dan kewajiban sebagai pasien, yang turut menghambat tindakan preventif sejak dini (Sabrina Putri & Sofiani, 2023).

Berbagai kondisi tersebut mencerminkan rendahnya kesadaran masyarakat tentang pentingnya kesehatan gigi di berbagai daerah di Indonesia, yang berkontribusi pada pola perilaku masyarakat dalam mengakses layanan kesehatan gigi. Tidak hanya itu, banyak masyarakat Indonesia juga menunda konsultasi ke dokter gigi hingga kondisi sudah parah, dengan alasan seperti rasa takut, biaya yang dianggap mahal, atau ketidaktahuan mengenai tempat konsultasi yang tersedia (Santoso dkk, 2020). Penundaan ini umumnya disebabkan oleh minimnya edukasi mengenai gejala awal penyakit gigi, serta kurangnya informasi tentang tindakan awal yang tepat saat menghadapi kondisi darurat (Maramis dkk, 2023).

Banyak masyarakat yang bingung apakah mereka harus menemui dokter umum atau spesialis berdasarkan gejala yang dirasakan (Matthew dkk, 2022). Akibatnya, sebagian besar masyarakat cenderung menunda pemeriksaan karena minimnya

kesadaran terhadap gejala penyakit serta sulitnya mengakses layanan Kesehatan (Anjum dkk, 2023). Penundaan ini dapat menyebabkan keterlambatan diagnosis dan penanganan penyakit, yang berpotensi memperburuk kondisi pasien. Selain itu, informasi yang diperoleh dari internet sering kali tidak kredibel dan dapat menyesatkan pasien dalam mengambil keputusan (Bharti dkk, 2020).

Dengan tingkat kesadaran masyarakat yang masih rendah terhadap pentingnya kesehatan gigi dan mulut, banyak individu cenderung menganggap perawatan gigi sebagai prioritas yang tidak mendesak (Maybury dkk, 2019). Hal ini menyebabkan masyarakat sering menunda kunjungan ke dokter gigi hingga gejala yang dirasakan menjadi parah, serta jarang melakukan pemeriksaan secara rutin. Kondisi ini diperburuk oleh kurangnya pemahaman mengenai gejala awal penyakit gigi, kebingungan dalam menentukan jenis tenaga medis yang harus dikunjungi, serta keterbatasan akses terhadap informasi yang kredibel dan layanan kesehatan yang terjangkau.

Untuk mengatasi tantangan tersebut, pemanfaatan teknologi digital seperti chatbot dinilai potensial dalam menyediakan informasi seputar kesehatan gigi secara cepat, mudah diakses, dan hemat biaya. Pendekatan chatbot modern banyak mengandalkan model bahasa besar (*Large Language Model*) seperti ChatGPT untuk menghasilkan percakapan yang alami dan fleksibel. ChatGPT dikembangkan menggunakan arsitektur *causal language modeling* dan disempurnakan melalui *Reinforcement Learning from Human Feedback* (RLHF), sehingga mampu memahami serta merespons berbagai topik dalam banyak bahasa, termasuk Bahasa Indonesia. Penelitian oleh (Brown dkk, 2020) menunjukkan bahwa model ini mampu menghasilkan respons yang koheren dan kontekstual dalam percakapan yang bervariasi, menjadikannya populer dalam berbagai aplikasi berbasis teks.

Namun demikian, dalam konteks kesehatan gigi, penggunaan ChatGPT memiliki sejumlah keterbatasan. Sebagai model generatif, ChatGPT berpotensi menghasilkan informasi yang keliru atau menyesatkan (*hallucination*), yakni ketika

model menciptakan jawaban yang tampak benar padahal tidak didukung fakta yang valid (Nisar dkk, 2023; Brown dkk, 2024). Selain itu, ChatGPT belum menjamin perlindungan privasi data pengguna, yang menjadi isu penting dalam pengembangan aplikasi berbasis kesehatan.

Metode MLM memungkinkan model memahami konteks kata secara menyeluruh dengan cara menyembunyikan (*masking*) sebagian token dalam kalimat dan memprediksi token yang hilang tersebut, sehingga model belajar secara bidirectional dan memiliki pemahaman konteks yang lebih dalam (Devlin dkk, 2018). Sementara itu, NSP berperan dalam melatih model agar dapat mengenali hubungan antar-kalimat dengan memprediksi apakah sebuah kalimat logisnya mengikuti kalimat sebelumnya (Yu dkk, 2024). Hal ini penting untuk menjaga koherensi dalam percakapan dan meningkatkan ketepatan respons. Menurut (Kumar, 2024), kombinasi MLM dan NSP dapat menghasilkan chatbot yang lebih fokus, relevan, dan minim *hallucination*, serta lebih efisien untuk dijalankan pada perangkat dengan sumber daya terbatas.

Namun, ketersediaan dataset publik untuk pelatihan dan *fine-tuning* model bahasa Indonesia khusus pada domain kesehatan gigi sangat terbatas. Tidak terdapat dataset yang memuat percakapan seputar gejala, tindakan awal, maupun edukasi kesehatan gigi dalam format yang sesuai untuk *intent classification* maupun *fine-tuning* menggunakan pendekatan *Masked Language Model* (MLM) dan *Next Sentence Prediction* (NSP). Oleh karena itu, dalam penelitian ini penyusunan dataset dilakukan secara khusus dengan melibatkan seorang dokter gigi sebagai ahli untuk memastikan validitas konten dan kesesuaian dengan kebutuhan domain.

Untuk mengatasi masalah tersebut, dikembangkan chatbot berbasis IndoBERT yang di-*fine-tune* menggunakan pendekatan *Masked Language Model* (MLM) dan *Next Sentence Prediction* (NSP). Metode ini memungkinkan chatbot memahami struktur dan hubungan antar-kalimat dalam Bahasa Indonesia secara lebih baik, sehingga dapat memberikan jawaban yang lebih akurat, kontekstual, dan relevan terhadap keluhan

medis. Selain itu, model ini juga dilengkapi dengan modul klasifikasi *intent* untuk mendeteksi maksud pertanyaan pengguna, sehingga respons yang diberikan semakin sesuai dengan kebutuhan informasi. Selain lebih aman dari sisi privasi, model ini juga ringan dan dapat dijalankan di perangkat dengan sumber daya terbatas.

Oleh karena itu, pengembangan chatbot kesehatan gigi dapat menjadi salah satu solusi untuk membantu masyarakat dalam memperoleh informasi dan konsultasi awal secara cepat, mudah diakses, dan hemat biaya. Chatbot ini ditujukan bagi masyarakat yang cenderung menunda perawatan gigi karena rendahnya kesadaran akan pentingnya kesehatan gigi, keterbatasan waktu, kekhawatiran terhadap biaya, atau anggapan bahwa perawatan gigi bukanlah prioritas. Selain itu, chatbot ini juga bermanfaat bagi individu yang tinggal di wilayah yang tidak memiliki akses langsung ke fasilitas kesehatan gigi. Dengan memberikan informasi awal yang relevan dan edukatif, chatbot diharapkan mampu mendorong masyarakat untuk mengambil langkah preventif lebih dini sebelum gejala memburuk.

Dengan menyediakan layanan gratis yang dapat diakses 24 jam, chatbot ini diharapkan mampu memberikan kemudahan bagi masyarakat dalam mengenali gejala sejak dini dan mengambil keputusan yang tepat. Penelitian ini mengembangkan chatbot berbasis web menggunakan model bahasa alami IndoBERT dengan pendekatan *Masked Language Model* (MLM), *Next Sentence Prediction* (NSP) serta *Intent Classification* untuk menghasilkan jawaban yang sesuai dengan konteks pertanyaan pengguna.

1.2 Rumusan Masalah

Berdasarkan dengan latar belakang yang telah dipaparkan, maka dapat dituliskan rumusan masalah untuk penelitian ini adalah sebagai berikut:

1. Bagaimana cara menyusun dan membangun *Knowledge Based* serta dataset khusus domain kesehatan gigi berbahasa Indonesia yang mencakup *Masked Language Model* (MLM), *Next Sentence Prediction* (NSP), *intent classification* dengan melibatkan ahli dokter gigi?

Nadhief Athallah Isya, 2025

IMPLEMENTASI MODEL IndoBERT PADA CHATBOT KESEHATAN GIGI DENGAN ALGORITMA MASKED LANGUAGE MODEL DAN NEXT SENTENCE PREDICTION

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

2. Bagaimana mengembangkan model IndoBERT pada domain kesehatan gigi melalui *fine-tuning* sehingga dapat dimanfaatkan sebagai basis chatbot berbasis *rule-based* dan *knowledge-based*?
3. Bagaimana mengevaluasi performa model IndoBERT yang telah di-*fine-tune* pada domain kesehatan gigi menggunakan metrik klasifikasi?
4. Bagaimana merancang dan membangun sistem chatbot kesehatan gigi berbasis *website* yang mengintegrasikan model IndoBERT yang telah di-*fine-tune*?
5. Bagaimana menguji kemampuan sistem chatbot kesehatan gigi dalam memberikan respons yang cepat, dapat diakses kapanpun, dan relevan terhadap pertanyaan pengguna?

1.3 Batasan Masalah

Karena begitu luasnya objek penelitian yang dihadapi penulis dan untuk memperjelas arah permasalahan yang dibahas maka penulis membatasi masalah-masalah sebagai berikut:

1. Masalah yang diangkat terbatas pada domain kesehatan gigi, khususnya kesulitan masyarakat dalam mengenali gejala awal, menentukan tenaga medis yang tepat, keterbatasan waktu dan biaya, serta rendahnya kesadaran dan edukasi mengenai pentingnya kesehatan gigi.
2. Chatbot yang dikembangkan hanya berfokus pada enam jenis penyakit gigi umum, yaitu karies gigi, gigi sensitif, gigi retak, gingivitis, pulpitis, dan abses, dengan ruang lingkup informasi terbatas pada sembilan fitur utama, yaitu pengertian, gejala, penyebab, tindakan awal, obat, pencegahan, komplikasi, kebutuhan untuk menemui dokter, durasi pemulihan, serta hal-hal yang harus dihindari.
3. Data yang digunakan untuk pengembangan chatbot terbatas pada dataset berbahasa Indonesia yang relevan dengan topik kesehatan gigi. Penyusunan dataset *Next Sentence Prediction (NSP)*, *intent classification*, dan *knowledge*

base dilakukan dengan melibatkan satu orang ahli dokter gigi sebagai sumber utama untuk memastikan validitas dan kesesuaian konten.

4. Sebagian variasi data diperluas dengan bantuan teknologi AI generatif, namun tetap melalui proses validasi manual oleh penulis dan ahli untuk menjamin akurasi serta kesesuaian dengan konteks penelitian.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Membangun dataset khusus domain kesehatan gigi berbahasa Indonesia yang mencakup *Masked Language Model (MLM)*, *Next Sentence Prediction (NSP)*, *intent classification*, serta *knowledge base* dengan melibatkan ahli dokter gigi.
2. Mengembangkan model IndoBERT pada domain kesehatan gigi melalui proses *fine-tuning* sehingga dapat dimanfaatkan sebagai basis chatbot berbasis *rule-based* dan *knowledge-based*.
3. Mengevaluasi performa model IndoBERT yang telah di-*fine-tune* pada domain kesehatan gigi dengan menggunakan metrik klasifikasi guna memastikan akurasi.
4. Membangun sistem chatbot kesehatan gigi berbasis website yang mengintegrasikan model IndoBERT hasil *fine-tuning*.
5. Menguji kemampuan sistem chatbot kesehatan gigi melalui *blackbox testing* untuk memastikan fungsionalitas, *non-performance testing* untuk menilai aspek keandalan dan kemudahan akses, serta validasi pakar untuk menjamin relevansi dan ketepatan konten medis yang diberikan.

1.5 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penelitian yang dilakukan adalah sebagai berikut:

1. Bagi Masyarakat

Nadhief Athallah Isya, 2025

IMPLEMENTASI MODEL IndoBERT PADA CHATBOT KESEHATAN GIGI DENGAN ALGORITMA MASKED LANGUAGE MODEL DAN NEXT SENTENCE PREDICTION

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Penelitian ini diharapkan dapat memberikan manfaat langsung kepada masyarakat dengan menyediakan akses layanan kesehatan gigi virtual yang cepat, mudah, dan bebas biaya. Sistem ini memungkinkan masyarakat, terutama yang tinggal di daerah dengan keterbatasan fasilitas kesehatan, untuk mendapatkan informasi kesehatan gigi yang relevan dan akurat tanpa harus bergantung pada fasilitas kesehatan konvensional yang sering sulit dijangkau. Dengan adanya sistem ini, masyarakat dapat memperoleh edukasi dan rekomendasi terkait kesehatan gigi kapan saja dan di mana saja, yang berpotensi meningkatkan kesadaran dan kebiasaan perawatan gigi yang baik.

2. Bagi Peneliti

Penelitian ini memberikan kontribusi dalam pengembangan teknologi chatbot berbasis NLP yang lebih spesifik pada konteks kesehatan gigi. Dengan menggunakan model IndoBERT dan teknik MLM serta NSP, penelitian ini memberikan wawasan baru dalam pengolahan bahasa alami untuk aplikasi kesehatan, yang dapat digunakan sebagai dasar bagi penelitian selanjutnya. Selain itu, penelitian ini juga membuka peluang untuk eksplorasi lebih lanjut terkait penerapan AI dalam layanan kesehatan di Indonesia.

3. Bagi Peneliti lain

Penelitian ini memberikan referensi yang bermanfaat bagi peneliti lain yang tertarik mengembangkan aplikasi serupa dalam bidang kesehatan atau bidang lain yang memanfaatkan NLP dan teknik pembelajaran mesin. Hasil penelitian ini juga dapat dijadikan acuan dalam mengembangkan sistem serupa untuk topik-topik kesehatan lain, serta sebagai langkah awal untuk memperkenalkan solusi digital yang dapat meningkatkan akses layanan kesehatan secara lebih merata, terutama di daerah yang kurang memiliki sumber daya.

Semoga penelitian ini tidak hanya meningkatkan pemahaman masyarakat tentang kesehatan, tetapi juga memperluas jangkauan layanan kesehatan yang lebih efisien dan terjangkau melalui teknologi.