

**SISTEM *INTENT DETECTION* PADA KELUHAN PASIEN UNTUK
REKOMENDASI DOKTER MENGGUNAKAN METODE *BIOBERT***

PROPOSAL TUGAS AKHIR

Proposal ini disusun guna memenuhi salah satu syarat untuk menyelesaikan program studi Teknik Informatika S-1 pada Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang



Disusun Oleh :

Nama : Reysita Nazela Fitrah

Nim : 32602100110

Program Studi : Teknik Informatika

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM SULTAN AGUNG
SEMARANG**

2025

LEMBAR PENGESAHAN
TUGAS AKHIR

SISTEM *INTENT DETECTION* PADA KELUHAN PASIEN UNTUK
REKOMENDASI DOKTER MENGGUNAKAN METODE
BIOBERT

REYSITA NAZELA FITRAH
NIM 32602100110

Telah dipertahankan di depan tim penguji ujian sarjana tugas akhir
Program Studi Teknik Informatika
Universitas Islam Sultan Agung
Pada tanggal : ...19 Februari 2025

TIM PENGUJI UJIAN SARJANA :

Imam Much Ibnu S. ST.MSc, Ph.D

NIDN.0613037301(

(Ketua Penguji)

3-3-2025

Andi Riansyah, ST, M.Kom

NIK. 060908802

(Anggota Penguji)

24-02-2025

Badie'ah, ST, M.Kom

NIK. 210615044

(Pembimbing)

3-3-2025

Semarang,

Mengetahui,

Kaprodi Teknik Informatika
Universitas Islam Sultan Agung



Moch Tarfik, ST, MIT

NIK/ 210604034

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Reysita Nazela Fitrah
NIM : 32602100110
Judul Tugas Akhir : SISTEM INTENT DETECTION PADA KELUHAN
MEDIS UNTUK REKOMENDASI DOKTER
MENGUNAKAN METODE BIOBERT

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, 06 Maret 2025

Yang Menyatakan,



Reysita Nazela Fitrah

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Reysita Nazela Fitrah

NIM : 32602100110

Program Studi : Teknik Informatika

Fakultas : Teknologi industri

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul : Sistem Intent Detection Pada Keluhan Medis untuk Rekomendasi Dokter menggunakan Metode BioBert

Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan diinternet dan media lain untuk kepentingan akademis selama tetap menyantumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiarisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan agung.

Semarang, 07 Maret 2025 .

Yang menyatakan,



Reysita Nazela Fitrah

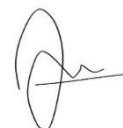
KATA PENGANTAR

Dengan mengucapkan syukur alhamdulillah atas kehadiran Allah SWT yang telah memberikan rahmat dan karunianya kepada penulis, sehingga dapat menyelesaikan Tugas Akhir dengan judul “Sistem *Intent Detection* Pada Keluhan Medis Untuk Rekomendasi Dokter Menggunakan Metode *BioBert*” ini untuk memenuhi salah satu syarat menyelesaikan studi serta dalam rangka memperoleh gelar sarjana (S-1) pada Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang. Tugas Akhir ini disusun dan dibuat dengan adanya bantuan dari berbagai pihak, materi maupun teknis, oleh karena itu saya selaku penulis mengucapkan terima kasih kepada:

1. Rektor UNISSULA Bapak Prof. Dr. H. Gunarto, S.H., M.H yang mengizinkan penulis menimba ilmu di kampus ini.
2. Dekan Fakultas Teknologi Industri Ibu Dr. Novi Marlyana, S.T., M.T., IPU.Eng
3. Dosen pembimbing penulis Ibu Badie'ah, S.T., M.Kom yang telah meluangkan waktu dan memberi ilmu.
4. Orang tua penulis, Ayah Widi Wicaksono dan Ibu Kristiana yang telah memberi suport dan mengizinkan untuk menyelesaikan laporan ini.
5. Dan kepada semua pihak yang tidak dapat saya sebutkan satu persatu.

Dengan rendah hati, penulis menyadari bahwa laporan masih memiliki banyak kekurangan dalam hal kuantitas, kualitas, dan ilmu pengetahuan. Oleh karena itu, penulis mengharapkan kritikan dan saran yang membangun untuk membantu laporan ini menjadi lebih baik di masa depan.

Semarang, 7 Maret 2025



Reysita Nazela Fitrah

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	Error! Bookmark not defined.
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	v
KATA PENGANTAR.....	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	viii
Abstrak.....	ix
BAB 1	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Pembatasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat	4
1.6 Sistematika Penulisan.....	4
BAB II	6
TINJAUAN PUSTAKA DAN DASAR TEORI.....	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori.....	8
2.2.1 <i>Intent</i>	8
2.2.2 <i>Intent Detection</i>	8
2.2.3 <i>Deep Learning</i>	8
2.2.4 <i>Natural Language Processing (NLP)</i>	9
2.2.5 <i>Transformer</i>	9
2.2.6 <i>Bidirectional Encoder Representations (BERT)</i>	12
2.2.7 <i>BioBert</i>	14
2.2.8 Identifikasi Keluhan Pasien	15
BAB III.....	16
METODOLOGI PENELITIAN	16

3.1 Tahapan Penelitian	16
3.1.1 Pengumpulan Data	17
3.1.2 Pra-Pemrosesan Data	17
3.1.3 Memasukan Teks dalam Model <i>BioBert</i>	17
3.1.4 Evaluasi Model	18
3.2 Analisis kebutuhan	18
3.3 Analsis sistem	20
3.4 Perancangan User Interface	21
3.4.1 Halaman awal sistem rekomendasi dokter	21
3.4.2 Halaman Hasil Sistem Rekomendasi Dokter	22
BAB IV	23
HASIL ANALISIS DAN PENELITIAN	23
4.1 Hasil Penelitian	23
4.1.1 Pengumpulan Data	23
4.1.2 Pra-Pemrosesan Data	23
4.1.3 Implementasi Model <i>BioBert</i>	23
4.2 Hasil Evaluasi	24
4.2.1 Model <i>BioBert</i>	24
4.3 Hasil Implementasi Antarmuka	27
4.2.1 Halaman Utama Sistem Rekomendasi Dokter	27
4.2.2 Halaman Hasil Sistem Rekomendasi	28
BAB V	30
KESIMPULAN DAN SARAN	30
5.1 Kesimpulan	30
5.2 Saran	30
DAFTAR PUSTAKA	31

DAFTAR GAMBAR

Gambar 2. 2 Architecture Transformer(Mohiuddin dkk., 2023)	10
Gambar 2. 3 Scaled Dot-Product Attention dan multi-head attention(Mohiuddin dkk., 2023)	11
Gambar 2. 4 Pre-training dan fine-tuning BERT(Translation, 2021)	14
Gambar 3.1 Flowchart sistem	16
Gambar 3. 2 Tahapan dalam Model BioBert	18
Gambar 3. 3 Flowchar Sistem Rekomendasi Dokter	20
Gambar 3. 4 Halaman Awal Sistem Rekomendasi Dokter	21
Gambar 3. 5 Halaman Hasil Sistem Rekomendasi Dokter	22
Gambar 4. 1 train loss	25
Gambar 4. 2 train accuracy dan validation accuracy	26
Gambar 4. 3 Hasil Akurasi.....	27
Gambar 4. 4 Halaman Utama Sistem Rekomendasi Dokter.....	28
Gambar 4. 5 menginputkan teks.....	29
Gambar 4. 6 Hasil Rekomendasi dokter	29



Abstrak

Dalam dunia kesehatan, proses identifikasi keluhan pasien merupakan langkah awal yang sangat penting untuk menentukan perawatan yang sesuai. Keluhan pasien sangat perlu dipahami dengan cepat dan akurat untuk merekomendasikan dokter yang tepat. Sistem berbasis *Intent Detection* dapat membantu pengguna dalam mengidentifikasi gejala penyakit yang sedang dialami dan mendapatkan rekomendasi dokter yang tepat. Penelitian ini menerapkan model *BioBert* (*Bidirectional Encoder Representations from Transformers for Biomedical Text Mining*), sebuah model berbasis NLP (*Natural Language Processing*) yang dirancang untuk teks biomedis, guna meningkatkan akurasi terhadap intent keluhan pasien dalam memberikan rekomendasi dokter. Metode yang digunakan dalam penelitian ini mencakup pengumpulan data keluhan medis, pra-pemrosesan teks, implementasi model *BioBert* untuk deteksi *intent*, serta evaluasi model berdasarkan metrik akurasi dan loss. Data yang digunakan berjumlah 1004 yang diperoleh dari platform digital. Pada proses awal pra-pemrosesan dilakukan pembersihan data yang meliputi tokenisasi, *stopword removal*, dan *stemming* guna menghilangkan kata-kata yang tidak diperlukan. Lalu, data kemudian dipetakan menjadi beberapa bagian untuk menentukan dokter yang sesuai. Sistem yang dibangun berupa Sistem rekomendasi dokter yang dapat mendeteksi *intent* dari keluhan penyakit yang dirasakan pengguna.

Hasil penelitian menunjukkan bahwa model *BioBert* mampu meningkatkan performa sistem dalam mengklasifikasikan keluhan medis. Model ini dapat memahami konteks bahasa medis dengan lebih baik, sehingga dapat mendeteksi dan mengelompokkan keluhan pasien dengan lebih akurat. Penggunaan model berbasis NLP ini juga memungkinkan pengolahan data dalam jumlah besar dengan efisiensi yang tinggi, sehingga sistem dapat terus diperbarui dan diperbaiki berdasarkan data baru yang masuk. Evaluasi terhadap sistem menunjukkan tingkat akurasi yang cukup tinggi dalam mendeteksi intent keluhan pasien, dengan hasil yang lebih baik dibandingkan model NLP standar lainnya. Hal ini menunjukkan

bahwa penerapan *BioBert* dalam sistem deteksi *intent* berbasis teks memiliki potensi besar dalam dunia medis.

Selain itu, penelitian ini juga menyoroti pentingnya kualitas data dalam meningkatkan performa model NLP. Proses pra-pemrosesan yang mencakup pembersihan data, normalisasi teks, dan penghapusan kata-kata yang tidak relevan terbukti berperan penting dalam meningkatkan akurasi deteksi *intent*. Dengan demikian, optimalisasi data input menjadi faktor kunci dalam keberhasilan implementasi sistem ini. Kedepannya, penelitian ini dapat diperluas dengan menambahkan lebih banyak data dari berbagai sumber, termasuk data keluhan pasien dari rumah sakit dan klinik, guna meningkatkan cakupan dan akurasi sistem. Selain itu, pengembangan fitur tambahan seperti dukungan multibahasa dan integrasi dengan rekam medis elektronik dapat lebih meningkatkan kegunaan sistem dalam lingkungan medis yang lebih luas.

Implikasi dari penelitian ini sangat luas, terutama dalam konteks layanan kesehatan berbasis teknologi. Sistem yang dikembangkan dapat digunakan sebagai alat bantu bagi tenaga medis dalam menyaring keluhan pasien sebelum konsultasi langsung dilakukan. Dengan adanya sistem ini, pasien dapat lebih cepat mendapatkan rekomendasi dokter yang sesuai dengan kondisi mereka tanpa harus menunggu proses administrasi yang lama. Selain itu, sistem ini juga dapat membantu dalam mengurangi beban kerja tenaga medis dengan menyaring dan mengelompokkan keluhan secara otomatis sebelum dilakukan diagnosis lebih lanjut oleh dokter. Dengan demikian, efisiensi layanan kesehatan dapat meningkat, sementara waktu tunggu pasien dapat dikurangi secara signifikan.

Secara keseluruhan, penelitian ini membuktikan bahwa penerapan teknologi NLP dalam bidang kesehatan memiliki dampak yang signifikan dalam meningkatkan akurasi deteksi *intent* keluhan pasien dengan nilai akurasi yang tinggi yaitu 94%. Dengan implementasi model *BioBert*, sistem dapat menjadi solusi inovatif dalam memberikan rekomendasi dokter secara cepat dan akurat. Pengembangan lebih lanjut terhadap sistem ini, baik dari segi data maupun fitur tambahan, dapat semakin memperkaya manfaatnya dalam dunia medis. Oleh karena itu, penelitian ini

menjadi langkah awal yang menjanjikan dalam penerapan kecerdasan buatan untuk meningkatkan kualitas layanan kesehatan di masa depan.

Kata kunci : *Intent Detection, BioBert , Natural Language Processing, Rekomendasi Dokter.*

Abstract

In the world of health, the process of identifying patient complaints is a very important initial step in determining the appropriate treatment. Patient complaints need to be understood quickly and accurately to recommend the right doctor. An Intent Detection-based system can help users identify symptoms of the disease they are experiencing and get the right doctor recommendations. This study applies the BioBert (Bidirectional Encoder Representations from Transformers for Biomedical Text Mining) model, an NLP (Natural Language Processing)-based model designed for biomedical texts, to improve the accuracy of patient complaint intent in providing doctor recommendations. The methods used in this study include collecting medical complaint data, text pre-processing, implementing the BioBert model for intent detection, and evaluating the model based on accuracy and loss metrics. The data used amounted to 1004 obtained from the digital platform. In the initial pre-processing process, data cleaning was carried out which included tokenization, stopword removal, and stemming to remove unnecessary words. Then, the data was mapped into several parts to determine the appropriate doctor. The system built is a doctor recommendation system that can detect the intent of the disease complaints felt by the user.

The results of the study show that the BioBert model is able to improve the performance of the system in classifying medical complaints. This model can understand the context of medical language better, so it can detect and group patient complaints more accurately. The use of this NLP-based model also allows processing large amounts of data with high efficiency, so that the system can be continuously updated and improved based on new incoming data. Evaluation of the

system shows a fairly high level of accuracy in detecting patient complaint intent, with better results than other standard NLP models. This shows that the application of BioBert in a text-based intent detection system has great potential in the medical world.

In addition, this study also highlights the importance of data quality in improving the performance of the NLP model. The pre-processing process including data cleaning, text normalization, and removal of irrelevant words has been shown to play an important role in improving the accuracy of intent detection. Thus, optimizing input data is a key factor in the successful implementation of this system. In the future, this study can be expanded by adding more data from various sources, including patient complaint data from hospitals and clinics, to increase the coverage and accuracy of the system. In addition, the development of additional features such as multilingual support and integration with electronic medical records can further enhance the usability of the system in a broader medical environment.

The implications of this study are very broad, especially in the context of technology-based healthcare services. The developed system can be used as an aid for medical personnel in filtering patient complaints before direct consultations are carried out. With this system, patients can more quickly get doctor recommendations that suit their conditions without having to wait for a long administrative process. In addition, this system can also help reduce the workload of medical personnel by automatically filtering and grouping complaints before further diagnosis is carried out by doctors. Thus, the efficiency of healthcare services can be increased, while patient waiting times can be significantly reduced. Overall, this study proves that the application of NLP technology in the health sector has a significant impact on increasing the accuracy of detecting patient complaint intent with a high accuracy value of 95%. With the implementation of the BioBert model, the system can be an innovative solution in providing doctor recommendations quickly and accurately. Further development of this system, both in terms of data and additional features, can further enrich its benefits in the

medical world. Therefore, this study is a promising first step in the application of artificial intelligence to improve the quality of healthcare services in the future.

Keywords: Intent Detection, BioBert, Natural Language Processing, Doctor Recommendation.



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dalam dunia kesehatan, proses identifikasi keluhan pasien merupakan langkah awal yang sangat penting untuk menentukan perawatan yang sesuai. Keluhan pasien sangat perlu dipahami dengan cepat dan akurat untuk merekomendasikan dokter yang tepat. Proses manual bisa memerlukan waktu yang sangat lama sehingga tidak efisien dan rentan terhadap kesalahan, terutama pada saat jumlah pasien sedang banyak. Hal ini tidak efektif dan tidak cukup baik untuk menanggapi keluhan pasien.

Dalam era digital saat ini, akses cepat dan tepat ke layanan kesehatan menjadi kebutuhan yang penting. Masyarakat modern mengharapkan layanan kesehatan yang responsif, praktis, dan mudah diakses melalui perangkat digital. Peningkatan beban kerja tenaga medis akibat jumlah pasien yang meningkat juga memerlukan solusi yang dapat membantu meringankan pekerjaan tersebut, terutama untuk keluhan awal. (Nurul Huda dkk., 2021). Dalam situasi di mana tenaga medis dan staf administrasi terbatas, diperlukan layanan kesehatan yang bisa membantu mengurangi beban mereka dalam menjawab pertanyaan umum dan keluhan. Sistem pengolahan keluhan medis untuk memberikan rekomendasi yang tepat kepada pasien.

Sistem berbasis yang dilengkapi dengan kemampuan *Natural Language Processing* (NLP) dapat menjadi solusi alat yang efektif untuk membantu pasien dalam mengidentifikasi keluhan dan mendapatkan rekomendasi dokter yang tepat. Dalam proses ini, deteksi intens keluhan pasien menjadi aspek penting untuk memberikan solusi yang sesuai dengan gejala yang dialami. *Natural Language Processing* merupakan salah satu cabang dari kecerdasan buatan yang berfokus

pada interaksi antara komputer dan bahasa manusia. NLP memuat berbagai teknik dan metode untuk mengolah dan menganalisis data teks. (Nur Oktavia dkk., 2024)

Intent detection merupakan suatu proses mengklasifikasi *intent* pengguna berdasarkan tulisan yang digunakan (Fatharani dkk., 2022). Salah satu pemanfaatan *intent detection* adalah pada sistem rekomendasi dokter. Dengan menerapkan *intent detection* dalam sistem untuk mengidentifikasi keluhan medis dengan tujuan memberikan rekomendasi dokter yang tepat, Sehingga metode ini dapat membantu pengguna terkait keluhan medis yang dirasakan dan dapat memberikan informasi rekomendasi dokter yang tepat, Metode ini menggunakan algoritma pembelajaran mesin untuk menemukan pola dari data keluhan pasien dan mengidentifikasi penyakit dari keluhan pasien.

Salah satu metode yang sangat efektif dalam *detection* teks adalah *BioBert* yang merupakan representasi dari model *BERT*. *BERT* merupakan model dua arah yang sudah dilatih sebelumnya menggunakan sejumlah teks besar dengan model bahasa yang ditujukan untuk memprediksi kata-kata yang disamarkan secara acak dari konteksnya (Asiva Noor Rachmayani, 2015). *BERT* telah dikembangkan secara luas untuk klasifikasi *intent* karena kemampuannya untuk memahami konteks dengan memproses teks dua arah, menawarkan akurasi yang tinggi dalam tugas klasifikasi *intent*. Dalam kasus keluhan medis, pendekatan pada model ini penting karena memungkinkan model memahami konteks setiap kata dalam kalimat pasien, baik dari sisi kiri maupun kanan. Misalnya, pada kalimat seperti “sakit kepala parah dan mual,” model akan memahami bahwa konteks kata “sakit” dan “mual” saling berhubungan dan keduanya mungkin mengindikasikan perlunya rekomendasi dokter spesialis tertentu, seperti spesialis saraf. Arsitektur *BERT* merupakan salah satu arsitektur yang berbasis transformasi dengan menggunakan mekanisme *self-attention*, *BERT* bisa memberikan perhatian lebih pada kata-kata tertentu sehingga dapat memahami konteks atau maksud suatu kata dengan baik. *BioBert* sendiri adalah salah satu representasi dari model *BERT* yang ditujukan untuk pemahaman kontek dalam bidang biomedis atau kesehatan. Kemampuannya dalam memahami setiap kata medis sangat cocok digunakan dalam sistem ini,

dimana *BioBert* dapat membantu memahami maksud dari keluhan pasien untuk menghasilkan intent dokter yang cocok.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, penulis mengidentifikasi permasalahan yang akan dibahas.

1. Bagaimana mengembangkan sistem deteksi intent berbasis NLP yang dapat memahami keluhan pasien secara otomatis?
2. Bagaimana cara mendeteksi *intent* dari keluhan medis pengguna menggunakan metode *BioBert* untuk menghasilkan rekomendasi dokter yang sesuai?
3. Seberapa efektif model klasifikasi intens berbasis *BioBert* dalam mengelompokkan jenis keluhan medis dan merekomendasikan dokter yang tepat dan akurat?

1.3 Pembatasan Masalah

Batasan masalah dari penulisan tugas akhir ini adalah sebagai berikut:

1. Sistem ini hanya akan menangani keluhan atau gejala medis umum yang dapat di tangani oleh dokter yang tersedia, keluhan yang rumit, spesifik atau jarang mungkin tidak dapat terdeteksi dengan baik.
2. Data keluhan pasien yang digunakan untuk pelatihan dan pengujian model hanya mencakup dataset terbatas.
3. Sistem deteksi ini berbasis *BioBert* representasi dari *BERT* dan hanya mampu memberikan rekomendasi dokter secara umum berdasarkan jenis gejala yang ada di dataset.
4. Sistem ini hanya akan meberikan rekomendasi dokter yang tepat kepada calon pasien berdasarkan deskripsi keluhan.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem deteksi intens keluhan pasien menggunakan *BioBert* untuk proses *embedding* dan *BERT* sebagai *classifier* untuk proses klasifikasi.
2. Mengukur performa model klasifikasi *BERT* dalam mengelompokkan keluhan pasien serta memberikan solusi rekomendasi dokter spesialis yang tepat.

1.5 Manfaat

Manfaat dari penelitian ini adalah :

1. Memberikan rekomendasi yang tepat kepada pengguna dari keluhan yang di alami.
2. Memberikan kemudahan pengguna dalam mengetahui dokter mana yang harus dituju.
3. Mengurangi beban kerja kepada tenaga medis.

1.6 Sistematika Penulisan

Berikut adalah sistematika penulisan yang digunakan dalam proses pembuatan laporan Tugas Akhir

BAB I : PENDAHULUAN

Pada Bab 1, penulis menjelaskan latar belakang dari penelitian yang dilakukan, permasalahan yang ingin dipecahkan, tujuan penelitian, batasan-batasan masalah yang dihadapi, serta manfaat dari penelitian ini. Pada akhir bab, diberikan gambaran umum tentang sistematika penulisan skripsi ini.

BAB II : DASAR TEORI DAN TINJAUAN PUSTAKA

Pada bab 2 ini penulis membahas teori-teori dan penelitian terdahulu yang mendukung penelitian yang dilakukan. Beberapa topik yang dibahas meliputi, *Natural Language Processing (NLP)*, *intent detection*, metode *BERT*, dan konsep sistem rekomendasu dokter.

BAB III : METODOLOGI PENELITIAN

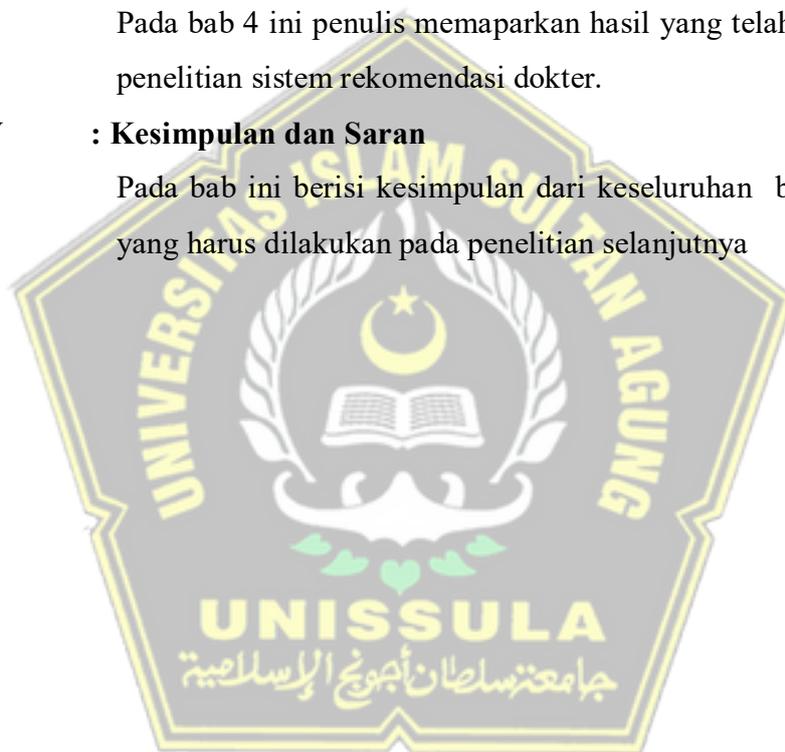
Bab 3 ini memaparkan metode penelitian yang digunakan, mulai dari pengumpulan data, pemrosesan data menggunakan teknik NLP, hingga implementasi metode BERT untuk klasifikasi intent. Selain itu, bab ini juga menjelaskan arsitektur sistem tahapan implementasi, serta metode evaluasi yang diterapkan untuk mengukur performa sistem yang dikembangkan.

BAB IV : Hasil dan Analisis Penelitian

Pada bab 4 ini penulis memaparkan hasil yang telah diproses dari penelitian sistem rekomendasi dokter.

BAB V : Kesimpulan dan Saran

Pada bab ini berisi kesimpulan dari keseluruhan bab dan saran yang harus dilakukan pada penelitian selanjutnya



BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Implementasi teknologi pada merupakan salah satu bentuk aplikasi NLP (*Natural Language Processing*). NLP sendiri merupakan salah satu bidang ilmu Kecerdasan buatan (*Artificial Intelligence*) yang mempelajari komunikasi antara manusia dengan komputer melalui bahasa alami. Model komputasi seperti ini berguna untuk memudahkan komunikasi antara manusia dengan komputer dalam hal pencarian informasi, sehingga dapat terjadi suatu interaksi antara keduanya dengan menggunakan bahasa alami. (Eldi & Syaputra, 2020). Dalam hal ini sistem ini dirancang untuk memudahkan pasien mendapatkan saran yang bermanfaat atas keluhan yang dirasakan dengan cepat, dan pasien mendapatkan rekomendasi dokter yang sesuai.

Pada penelitian sebelumnya (Kumar, 2021) mengembangkan sistem berbasis machine learning untuk merekomendasikan dokter spesialis berdasarkan analisis teks keluhan pasien. Penelitian ini mencakup integrasi model pembelajaran mesin seperti Random Forest dan Naïve Bayes untuk memprediksi kemungkinan penyakit berdasarkan gejala yang dimasukkan oleh pasien. Berdasarkan hasil prediksi penyakit, sistem ini dapat merekomendasikan dokter spesialis yang sesuai, seperti kardiolog jika penyakit yang terdeteksi berhubungan dengan jantung. Proses ini melibatkan pengumpulan data gejala dari pengguna, pemrosesan data menggunakan algoritma pembelajaran mesin, dan menghasilkan rekomendasi berbasis penyakit yang diprediksi. Sistem ini menggunakan data yang diambil dari berbagai dataset medis yang valid, seperti dataset Cleveland untuk penyakit jantung dan dataset Pima Indian untuk diabetes, guna melatih model prediksi mereka.

Pada penelitian yang dilakukan (Bian dkk., 2023) Mendemonstrasikan kemampuan *BioBert* , sebuah model berbasis pemrosesan bahasa alami yang telah dilatih sebelumnya secara khusus pada data biomedis, untuk menyelesaikan tugas-tugas seperti *Named Entity Recognition* (NER) dan klasifikasi teks dalam domain

biomedis dengan akurasi yang signifikan. Model ini dioptimalkan untuk teks biomedis dengan memanfaatkan korpus besar seperti PubMed dan PMC, yang memungkinkan identifikasi entitas medis seperti penyakit, obat-obatan, dan protein dengan lebih baik dibandingkan model NLP umum lainnya. *BioBert* menunjukkan kinerja luar biasa dalam tugas NER, misalnya dalam ekstraksi informasi yang terstruktur dari teks medis yang tidak terstruktur, seperti mencocokkan gejala dengan kondisi klinis atau klasifikasi hubungan antar-entitas dalam teks. Peningkatan akurasi signifikan juga dicapai dalam tugas-tugas seperti relasi entitas dan penjawaban pertanyaan berbasis data biomedis, menunjukkan kemampuan BioBert untuk memahami konteks domain secara mendalam

Pada penelitian yang dilakukan oleh (Berge dkk., 2023) mengembangkan model berbasis *Natural Language Processing* (NLP) untuk mendeteksi keluhan pasien dari teks tidak terstruktur dan mencocokkannya dengan spesialisasi dokter yang relevan. Model ini bertujuan untuk mengotomatisasi analisis teks keluhan pasien, sehingga dapat memberikan rekomendasi spesialisasi dokter dengan lebih cepat dan akurat. Algoritma yang digunakan memanfaatkan teknik NLP modern untuk memahami hubungan semantik antara gejala pasien dan bidang spesialisasi medis. Sistem ini bekerja dengan mengekstraksi informasi penting dari data pasien menggunakan pendekatan berbasis konsep dan mengklasifikasikan keluhan secara otomatis. Dalam penerapannya, pendekatan ini memperbaiki efisiensi pencarian dan pencocokan data medis melalui algoritma pembelajaran mesin yang mengintegrasikan konsep-konsep medis dengan istilah sinonim atau frasa terkait. Hal ini memungkinkan deteksi keluhan dan pencocokan yang lebih presisi meskipun data pasien memiliki variasi format atau ekspresi linguistik. Teknologi seperti ini sangat relevan dalam konteks peningkatan layanan kesehatan berbasis teknologi, di mana efisiensi dalam menganalisis keluhan pasien secara langsung berdampak pada kepuasan dan hasil perawatan pasien. Pendekatan ini juga mengurangi beban kerja administratif dokter, sehingga lebih banyak waktu dapat dialokasikan untuk perawatan klinis langsung

Pada penelitian ini akan, akan digunakan model *BioBert* (*Bidirectional Encoder Representations from Transformers for Biomedical Text Mining*) untuk

mengembangkan sistem *intent detection* pada keluhan medis guna memberikan rekomendasi dokter spesialis. Pendekatan ini mengintegrasikan teknologi *Natural Language Processing* (NLP) dengan pemrosesan teks biomedis *BioBert* digunakan untuk mendeteksi input keluhan medis pengguna yang akan memberikan hasil rekomendasi dokter spesialis dari keluhan medis pasien. Menggunakan pendekatan *BioBert* untuk melakukan proses klasifikasi dan deteksi intent pada teks medis.

2.2 Dasar Teori

2.2.1 Intent

Intent adalah maksud atau tujuan yang dapat memahami keinginan pengguna. ditujukan untuk menyelesaikan suatu tugas tertentu, seperti mencari informasi secara spesifik (Joey Ferelestian dkk., 2023). *Intent* disini digunakan untuk membantu sistem dalam mengklasifikasi jenis keluhan medis yang dapat menentukan rekomendasi dokter spesialis. Setiap input teks diklasifikasikan berdasarkan jenis penyakit sehingga dapat di proses untuk memberikan rekomendasi dokter spesialis yang tepat sesuai dari keluhan medis pengguna.

2.2.2 Intent Detection

Intent detection merupakan proses mengklasifikasi *intent* pengguna berdasarkan tulisan atau lisan yang digunakan. Dalam konteks sistem kesehatan, *intent detection* bertujuan untuk mengenali maksud dari pengguna berdasarkan keluhan yang dialami. *Intent* yang diklasifikasikan akan menentukan rekomendasi dokter spesialis apa yang akan diberikan kepada pasien atau pengguna. (Fatharani dkk., 2022) sehingga sistem dapat dengan mudah menentukan rekomendasi yang sesuai.

2.2.3 Deep Learning

Deep Learning (DL) adalah salah satu cabang dari *Machine Learning* yang terdiri dari algoritma pemodelan abstraksi tingkat tinggi pada data menggunakan sekumpulan fungsi transformasi non-linear yang ditata berlapis-lapis dan mendalam. DL sangat baik untuk diterapkan pada *supervised learning*, *unsupervised learning* dan *semi-supervised learning* maupun untuk *reinforcement*

learning dalam berbagai aplikasi seperti pengenalan citra, suara, klasifikasi teks, dan sebagainya.(Homepage dkk., 2024)

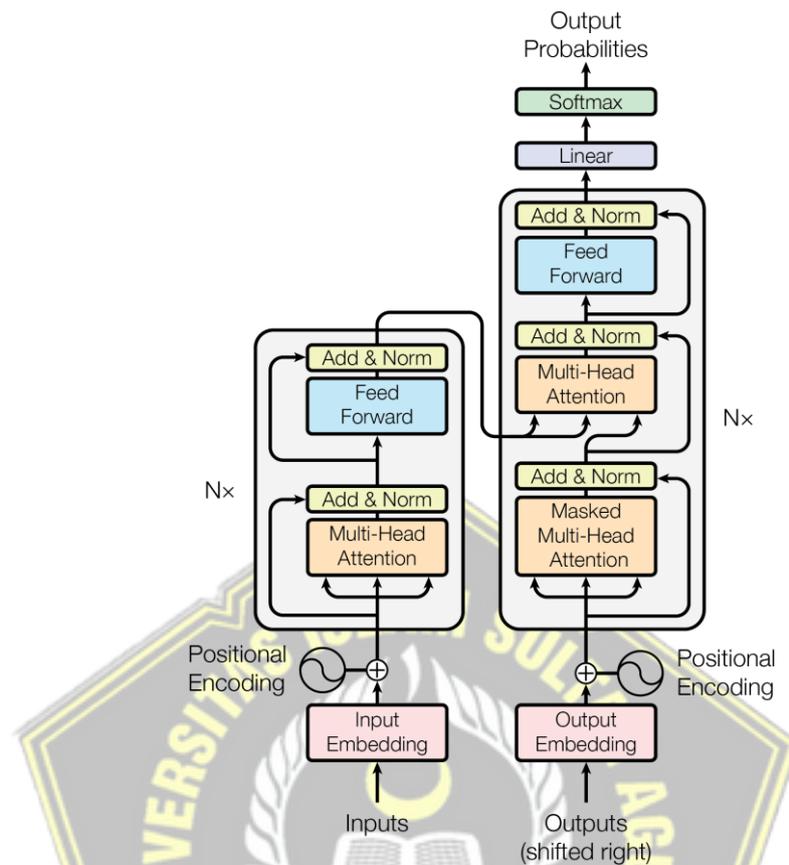
Penerapan *deep learning* dalam sistem rekomendasi dokter ini mampu mendeteksi *intent* pengguna, khususnya dalam konteks keluhan medis, *Deep Learning* adalah jenis pembelajaran mesin yang mengajarkan komputer untuk melakukan tugas dengan menggunakan contoh manusia. (Al Fadil Syahputra dkk., 2024)

2.2.4 Natural Language Processing (NLP)

Natural Language Processing adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP mencakup berbagai teknik dan metode untuk mengolah dan menganalisis data teks yang bertujuan untuk memahami arti atau maksud dari teks atau ucapan yang diberikan oleh manusia yang mencakup *intent classification*. Proses di mana komputer menghasilkan teks atau ucapan dalam bahasa alami. Contohnya adalah merespons pertanyaan pengguna dalam sistem *intent detection* rekomendasi dokter. (Nur Oktavia dkk., 2024)

2.2.5 Transformer

Transformer adalah jenis arsitektur neural baru yang mengodekan data input sebagai fitur-fitur canggih melalui mekanisme perhatian. Pada dasarnya, *transformer* visual pertama-tama membagi gambar input menjadi beberapa petak lokal, lalu menghitung representasi dan hubungannya. Karena gambar alami memiliki kompleksitas tinggi dengan detail dan informasi warna yang melimpah, granularitas pembagian petak tidak cukup baik untuk menggali fitur objek dalam skala dan lokasi yang berbeda.(Han dkk., 2021) Memiliki 2 struktur utama yaitu, Encoder Proses mengubah input menjadi representasi internal. Terdiri dari beberapa lapisan layer yang memiliki bloks self-attention dan lapisan feed forward. Arsitektur pada transformer bisa dilihat pada Gambar 2.1 Dibawah ini.



Gambar 2. 1 Architecture Transformer (Mohiuddin dkk., 2023)

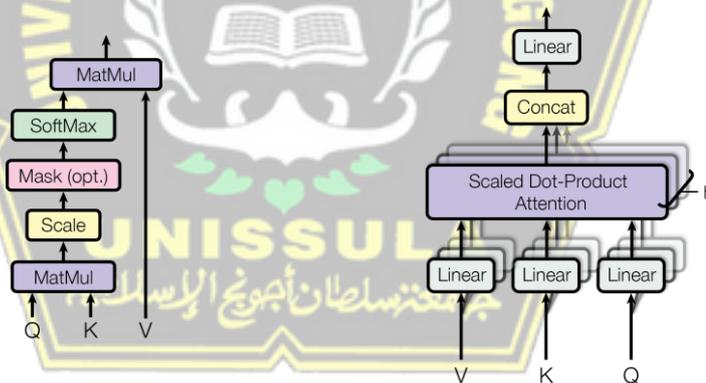
a. *Encoder dan Decoder*

Encoder dalam implementasinya digunakan untuk membaca semua teks yang diinput secara bersamaan. *Encoder* terdiri dari beberapa tumpukan lapisan yang identik. Pada Setiap lapisannya memiliki dua sub-lapisan, yaitu lapisan pertama *self-attention* dan lapisan kedua jaringan saraf *feed-forward*. dengan *self-attention* (Han dkk., 2021). *Decoder* merupakan bagian yang mengambil representasi dari *encoder* dan dapat menghasilkan output. Memiliki beberapa lapisan yang terdiri dari *self-attention*, *cross-attention*, dan lapisan *feed forward*, *Dekoder* digunakan untuk menghasilkan urutan keluaran yang prediktif. *Dekoder* juga terdiri dari tumpukan lapisan yang dapat diidentifikasi. Setiap lapisan terdiri dari dua sublapisan seperti lapisan *encoder*, dengan lapisan perhatian tambahan di antara kedua lapisan tersebut untuk membantu *node* saat ini mendapatkan konten utama yang perlu diperhatikan dengan melakukan perhatian *multi-head* pada keluaran *encoder*. *Self-attention* mekanisme yang memungkinkan model

memperhatikan bagian-bagian lain dari input saat memproses setiap kata. *Transformer* dapat memahami setiap kata atau kalimat saling berkaitan serta memahami konteks dari kata tersebut. (Mohiuddin dkk., 2023) Misal kata ‘bisa’ dapat diartikan beda jika konteksnya beda “bisa ular” dan “bisa gila” dua kata yang sama namun dengan konteks yang berbeda maka maknanya pun juga ikut berbeda. Dengan *self-attention* ini membantu untuk memahami makna dalam setiap kata.

b. Scaled Dot-Product Attention

Input terdiri dari kueri dan kunci berdimensi d_k , dan nilai berdimensi d_v . Kami menghitung produk titik dari kueri dengan semua kunci, bagi masing-masing dengan $\frac{1}{\sqrt{d_k}}$, dan terapkan fungsi softmax untuk memperoleh bobot pada nilai-nilai tersebut. Dua fungsi perhatian yang paling umum digunakan adalah perhatian aditif, dan perhatian produk titik (perkalian). Perhatian produk titik identik dengan algoritma kami, kecuali untuk faktor skala. Gambar 2.2 dibawah ini menunjukkan diagram alur dari mekanisme *Attention* dalam Transformer, khususnya *Multi-Head Attention*. Pada Gambar 2.2 dibawah ini merupakan alur model *transformer*.



Gambar 2. 2 Scaled Dot-Product Attention dan multi-head attention (Mohiuddin dkk., 2023)

c. Multi-head Attention

Multi-head attention layer pada ViT merupakan salah satu layer yang dapat digunakan untuk memproses informasi dari gambar. *Multi-head attention* terdiri dari beberapa layer *attention* yang dijalankan secara paralel. Multi-head attention memungkinkan model untuk menangani informasi secara bersamaan dari representasi dan posisi yang berbeda. Sederhananya, multi-head attention berfungsi

agar patch gambar yang ada dapat berkomunikasi dengan mengirimkan informasi ganda dengan menggunakan proses looping yang memanfaatkan *scale dot-product attention*, *query*, *keys*, dan *value* yang dapat dieksekusi secara parallel.(Halim, 2024).

d. Position-wise Feed-Forward Networks

Selain sublapisan perhatian, masing-masing lapisan dalam encoder dan decoder kami berisi jaringan umpan maju yang terhubung penuh, yang diterapkan ke setiap posisi secara terpisah dan identik. Ini terdiri dari dua transformasi linier dengan aktivasi ReLU. Meskipun transformasi linier sama di berbagai posisi, transformasi tersebut menggunakan parameter yang berbeda dari satu lapisan ke lapisan lainnya. Cara lain untuk mendeskripsikannya adalah sebagai dua konvolusi dengan ukuran kernel 1 Dimensi masukan dan keluaran adalah $d_{model} = 512$, dan lapisan dalam memiliki dimensi $d_{ff} = 2048$. (Mohiuddin dkk., 2023)

e. Embeddings and Softmax

Embedding yang dipelajari untuk mengubah token *input* dan token *output* menjadi vektor berdimensi model. Menggunakan transformasi linear yang dipelajari dan fungsi *softmax* untuk mengubah output *decoder* menjadi probabilitas token berikutnya yang diprediksi. Dalam model kami, kami berbagi matriks bobot yang sama antara dua lapisan embedding dan transformasi linear pra-softmax, mirip dengan. Di lapisan *embedding*, kami mengalikan bobot tersebut dengan $\frac{1}{d_{model}}$.(Mohiuddin dkk., 2023)

f. Positional Encoding

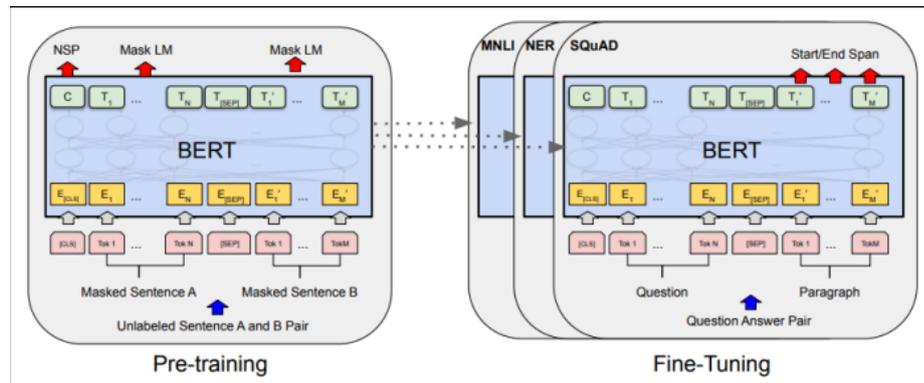
Positional Embedding adalah sebuah mekanisme pada model transformer (seperti BERT dan BioBert) yang bertujuan menyampaikan informasi tentang urutan atau posisi kata dalam sebuah kalimat kepada model. Hal ini penting karena model transformer, yang menggunakan arsitektur self-attention, tidak secara alami mampu mengenali urutan kata, berbeda dengan model seperti RNN yang memproses kata secara berurutan.(Mohiuddin dkk., 2023)

2.2.6 Bidirectional Encoder Representations (BERT)

Bidirectional Encoder Representations from Transformers (BERT) merupakan model dua arah yang telah dilatih sebelumnya menggunakan sejumlah

besar teks dengan tujuan model bahasa bertopeng yang tujuannya adalah untuk memprediksi kata-kata yang disamarkan secara acak dari konteksnya. (Asiva Noor Rachmayani, 2015) dengan menggunakan Metode pendekatan *BioBert* dari *BERT*, sistem dapat memahami konteks keluhan medis dengan lebih akurat dan memberikan rekomendasi spesialis yang sesuai. Metode *BERT* melatih representasi dua arah yaitu kanan dan kiri sehingga dapat memahami konteks bahasa lebih baik dan tepat. *BERT* melakukan 2 tugas utama, yaitu *Pre-Training* dan *Fine-Tuning*. Pada tahap *Pre-Training BERT* melakukan dua tugas utama diantaranya MLM dan NSP. MLM merupakan sebuah pemodelan bahasa masal yang dilakukan untuk memahami konteks dari suatu kata atau bahasa, sedangkan NSP adalah proses pengambilan 2 kalimat untuk memahami hubungan antar kalimat dan memprediksi apakah 2 kalimat tersebut memiliki hubungan berurutan atau tidak. Lalu pada Tahap yang kedua yaitu *Fine-Tuning*, *BERT* melakukan proses klasifikasi teks, dan deteksi entitas yang nantinya setelah proses ini dapat dilakukan *Mapping Detection*. Pada dasarnya *BERT* juga dapat dipahami sebagai suatu arsitektur stack transformer. *BERT* memiliki dua buah model, yaitu *BERT Base* yang memiliki 12 layer *stack encoder* layer dan *BERT Large* yang memiliki 24 *stack encoder* layer. *BERT Base* memiliki sekitar 110M parameter sedangkan *BERT Large* memiliki 340M parameter (Bei dan Sudin, 2021).

Dalam *BERT* terdapat dua langkah yang dilalui yaitu yang disebut pre-training dan fine-tuning. *Pre-training* merupakan tahap dimana model dilatih dengan menggunakan data yang tidak berlabel (*unsupervised task*) melalui pre-training task, sedangkan fine-tuning merupakan model yang diinisialisasi dengan parameter hasil training sebelumnya, dan *fine-tuning* menggunakan data yang sudah berlabel. (Awalina, n.d.). arsitektur kerja dari *BERT* dapat dilihat pada Gambar 2.3 dibawah ini:



Gambar 2. 3 *Pre-training* dan *fine-tuning* BERT(Translation, 2021)

a *Pre-Training*

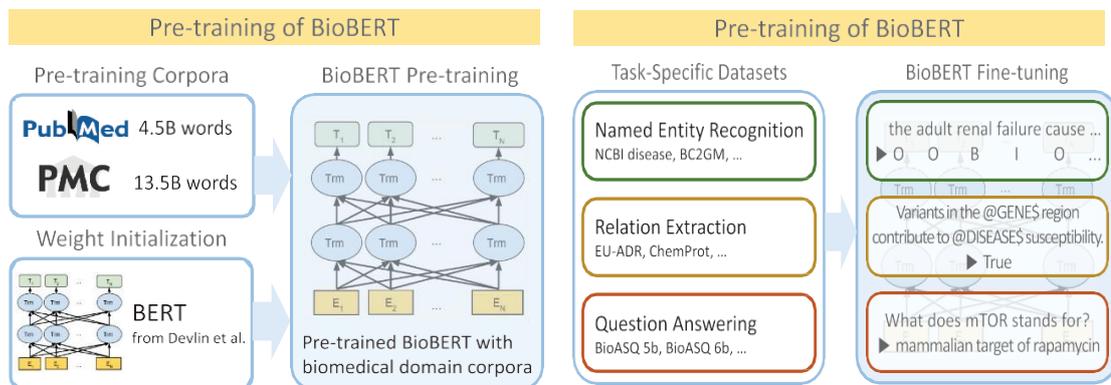
Pre-Training merupakan sebuah proses suatu model yang diekspos ke sejumlah teks besar dari berbagai sumber yang ada di internet, yang memungkinkannya untuk mempelajari struktur dan pola bahasa yang kompleks.(Setiawan dkk., 2023) *Pre-Training* dilakukan untuk membantu sistem dalam memahami konteks kalimat dan memahami hubungan antar kalimat pengguna.

b *Fine-Tuning*

Fine-Tuning merupakan proses yang dilakukan setelah *pre-training* menggunakan dataset yang telah diberi label secara khusus untuk tugas tertentu, seperti menerjemahkan bahasa, menjawab pertanyaan, atau berinteraksi dalam bentuk percakapan.(Setiawan dkk., 2023)

2.2.7 BioBert

BioBert merupakan representasi dari model *BERT* yang telah dilatih sebelumnya untuk domain biomedis. *BioBert* memiliki struktur yang sama dengan model *BERT* hanya saja dalam prosesnya *BioBert* difokuskan pada satu tugas seperti NER,RE, dan QA. *BioBert* memanfaatkan arsitektur *BERT*, yang mengandalkan *self-attention* dan *transformer* untuk memahami konteks kata dalam kedua arah yang dirancang secara khusus menangani pengenalan entitas biomedis. Gambar 3.1 dibawah ini adalah arsitektur dari model *BioBert*.



Gambar 3.1 Arsitektur *BioBert* (Lee dkk., 2020)

BioBert secara khusus dirancang untuk tugas seperti:

- Named Entity Recognition (NER)*
Proses ekstraksi relasi untuk mengklasifikasikan relasi relasi yang telah diberi nama entitas biomedis. Mepresentasikan [CLS]
- Relation Extraction (ER)*
Proses mengidentifikasi hubungan antara entitas keluhan medis.
- Question Answering (QA)*
Output yang diperoleh sehingga dapat menjawab pertanyaan teks biomedis.
(Lee dkk., 2020)

2.2.8 Identifikasi Keluhan Pasien

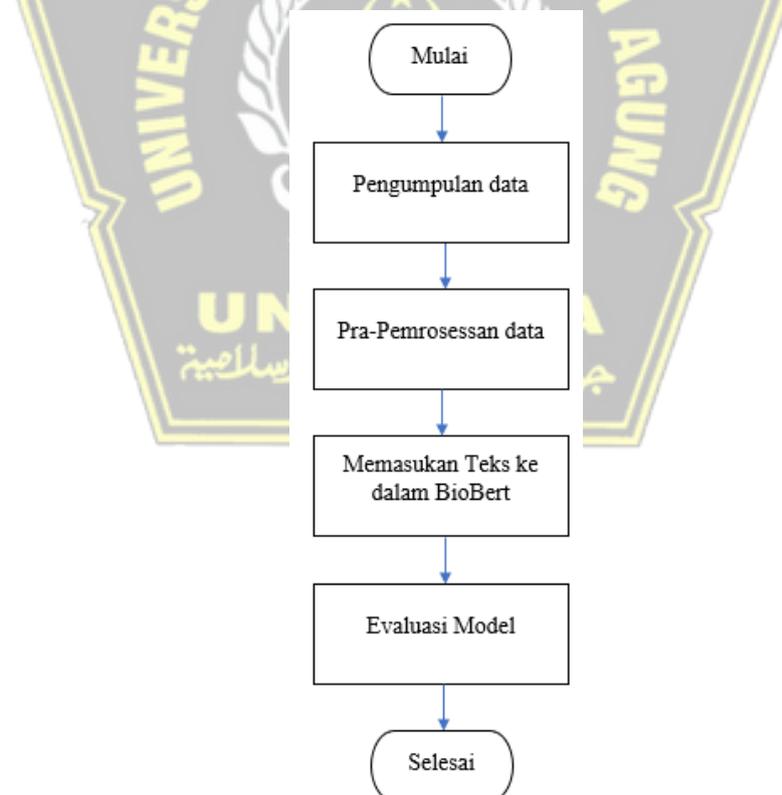
Keluhan pasien merupakan tahap awal pemeriksaan yang dilakukan oleh dokter untuk mengumpulkan informasi terkait kondisi yang sedang dirasakan oleh pasien agar dokter dapat menyimpulkan hasil diagnosis dari pasien tersebut. (Hendrawati, 2017) maka dibutuhkan sistem yang dapat membantu pengguna dalam mengenali keluhan yang dirasakan untuk selanjutnya dapat dianalisis keluhan atau gejala yang disampaikan pasien dan menentukan spesialisasi medis yang tepat.

BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Penelitian ini berfokus pada pengembangan sistem yang mampu mendeteksi intent dari keluhan medis pengguna dan merekomendasikan dokter yang sesuai. Pendekatan yang digunakan adalah dengan menerapkan *Natural Language Processing* (NLP), khususnya menggunakan model *BioBert* (*Bidirectional Encoder Representations from Transformers*) untuk deteksi intent yang mampu memahami maksud pengguna. Metode pada penelitian ini memuat langkah-langkah yang diantaranya adalah pengumpulan data, pra-pemrosesan data, memasukan teks pada model *BioBert* dan evaluasi model. Alur metode penelitian ini dapat dilihat pada Gambar 3.1



Gambar 3.1 *Flowchart sistem*

3.1.1 Pengumpulan Data

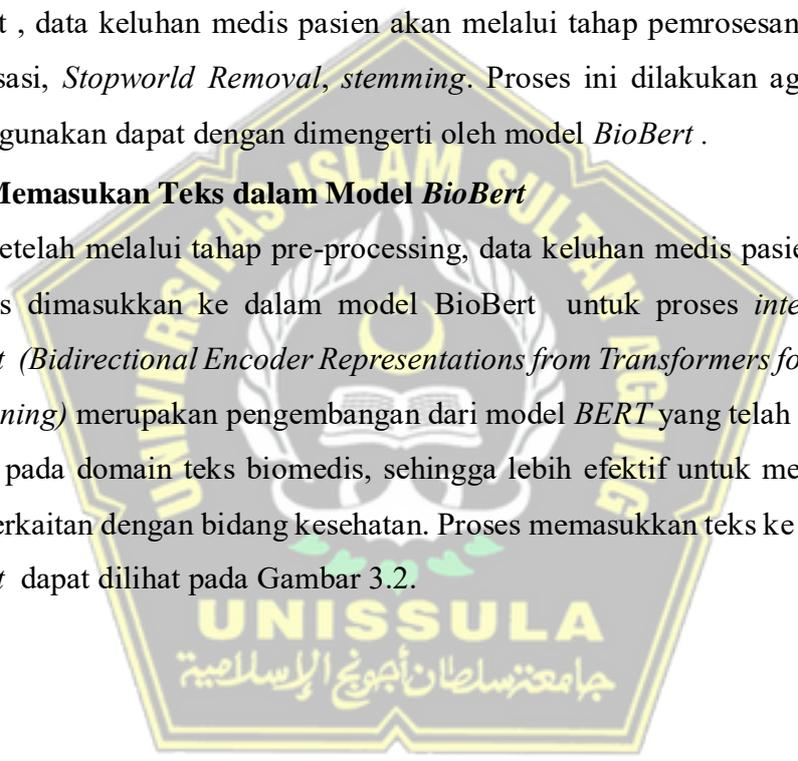
Data yang akan digunakan dalam penelitian ini akan berupa Dataset keluhan medis dan spesialis dokter yang diambil dari dataset terbuka yang diperoleh dari kaggle. Data yang telah di peroleh diberi label sesuai dengan keluhan dan dokter spesialisnya.

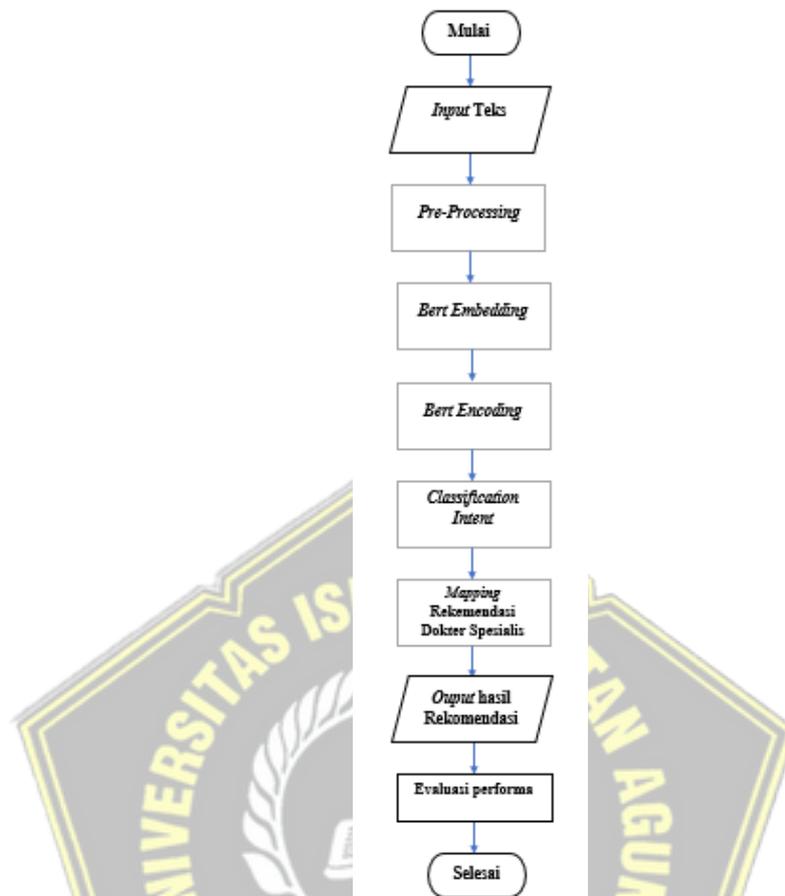
3.1.2 Pra-Pemrosesan Data

Data yang telah dikumpulkan akan melalui tahap pra-pemrosesan untuk membersihkan, merapikan, dan mempersiapkan data Sebelum masuk ke model BioBert , data keluhan medis pasien akan melalui tahap pemrosesan awal seperti Tokenisasi, *Stopworld Removal*, *stemming*. Proses ini dilakukan agar data yang akan digunakan dapat dengan dimengerti oleh model *BioBert* .

3.1.3 Memasukan Teks dalam Model *BioBert*

Setelah melalui tahap pre-processing, data keluhan medis pasien yang telah diproses dimasukkan ke dalam model BioBert untuk proses *intent detection*. *BioBert (Bidirectional Encoder Representations from Transformers for Biomedical Text Mining)* merupakan pengembangan dari model *BERT* yang telah dilatih secara khusus pada domain teks biomedis, sehingga lebih efektif untuk menangani data yang berkaitan dengan bidang kesehatan. Proses memasukkan teks ke dalam model *BioBert* dapat dilihat pada Gambar 3.2.





Gambar 3. 2 *Flowchart Tahapan dalam Model BioBert*

3.1.4 Evaluasi Model

Pada penelitian ini, evaluasi dilakukan untuk mengukur sejauh mana model dapat mendeteksi intent dari keluhan medis pasien secara akurat dan memberikan rekomendasi dokter yang sesuai. Proses evaluasi dilakukan melalui beberapa metrik kinerja yang relevan klasifikasi.

3.2 Analisis kebutuhan

Pada tahap ini, penulis menganalisis software apa saja yang diperlukan dalam proses pengerjaan sistem rekomendasi dokter ini, sehingga proses penelitian dapat berjalan dengan baik dan sesuai. Berikut, merupakan beberapa software yang dibutuhkan dalam proses penelitian :

1. *Python*

Python merupakan bahasa pemrograman tingkat tinggi yang memiliki sintaks sederhana yang mudah dipahami dan bersifat open-source. Pada penelitian ini

digunakan python versi 3.11.5 yang memiliki versi stabil dan lebih baik sehingga dalam penggunaannya lebih mudah dan efisien

2. *Library Transformer*

Transformers merupakan pustaka bahasa pemrograman *python* yang dikembangkan oleh *Huggingface*. Pada penelitian ini *transformer* digunakan untuk mempermudah penggunaan model *BioBert*.

3. *Library Random*

Pada penelitian ini *Library Random* digunakan untuk melakukan proses simulasi dan pembangkitan data dummy sehingga dapat menghasilkan probabilitas yang tinggi

4. *Library Numpy*

Numpy (Numerical Python) adalah sebuah library python yang digunakan untuk komputasi array. *Library* ini menyediakan struktur data *array* multidimensi yang sangat efisien untuk perhitungan kemiripan antara gejala yang di inputkan pasien dan rekomendasi dokter yang akan diberikan.

5. *Library NLTK*

Library NLTK (Natural Language Toolkit) merupakan *library python* yang digunakan dalam pemrosesan bahasa alami atau NLP (*Natural Language Process*). Pada penelitian ini *Library* ini digunakan untuk proses tokenisasi, *lemmatization*, dan *Stopword removal*.

6. *Google Colaboratory*

Google Colaboratory adalah sebuah layanan *cloud* yang disediakan oleh *Google* untuk menjalankan kode *python* dalam *notebook jupyter*. Dalam pengembangan aplikasi ini *google colab* digunakan untuk menjalankan kode program seperti training.

7. *Visual Studio Code*

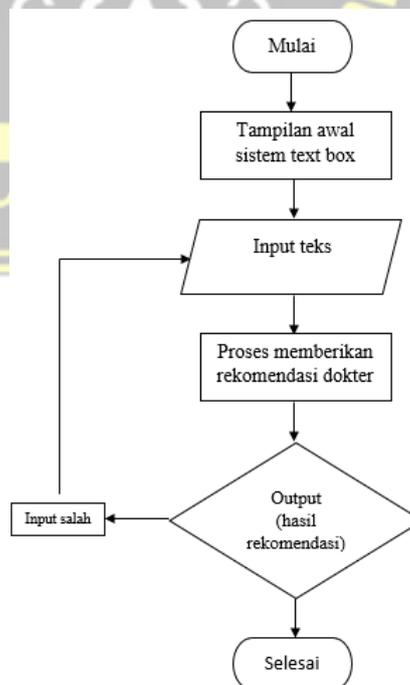
Visual Studio Code adalah sebuah aplikasi teks editor yang dikembangkan oleh *microsoft*. Dalam pengembangan pembuatan sistem pada penelitian ini *VSC* digunakan karena ringan dan mendukung berbagai bahasa pemrograman termasuk *python* sehingga pengelolaan kode dan pengembangan aplikasi sistem ini lebih efisien.

8. *Framework Flask*

Flask adalah sebuah Framework web berbasis python yang dirancang untuk memulai dengan cepat dan mudah dalam proses membangun sebuah *web* atau API (Santoso & Saian, 2023). Pada penelitian ini *Flask* digunakan sebagai pengembangan aplikasi *website* karena flask ringan dan mudah dikembangkan sehingga implementasinya mudah dijalankan dengan model *BioBert*.

3.3 Analisis sistem

Di tahap analisa sistem ini, penulis melakukan analisis terhadap sistem yang akan di buat menggunakan *flowchart*. Dimana pada tahap pertama user akan masuk pada tampilan rekomendasi dokter, lalu user menginputkan teks yang berisi gejala penyakit yang sedang dialami. Setelah itu sistem akan mulai memproses teks yg diinputkan. Lalu setelahnya akan muncul output hasil rekomendasi dokter. Jika user salah memasukan atau kurang lengkap memasukan teks, user dapat kembali menginputkan teks di laman rekomendasi dokter, dan sistem akan memproses kembali teks yang telah diinputkan.



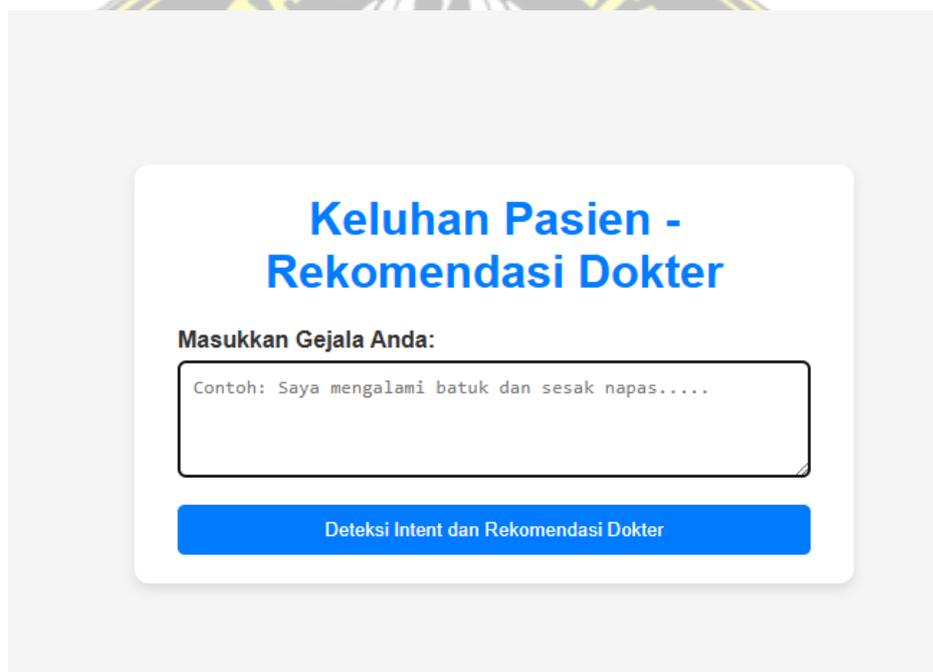
Gambar 3. 3 *Flowchart* Sistem Rekomendasi Dokter

- a. Mulai : Proses dimulai
- b. *Input* teks gejala penyakit : User akan masuk pada laman sistem rekomendasi dokter. Pada tampilan sistem, user memasukan teks gejala penyakit.
- c. Proses sistem : Setelah memasukan teks, sistem akan mulai memproses gejala penyakit tersebut untuk dicocokkan dokter mana yang sesuai dengan teks yang telah diinputkan.
- d. Output hasil : Setelah teks berhasil diproses, sistem akan memberikan hasil rekomendasi dokter yang sesuai dengan gejala penyakit yang sedang dialami.

3.4 Perancangan User Interface

3.4.1 Halaman awal sistem rekomendasi dokter

Halaman ini menampilkan halaman awal sebelum memasukan teks keluhan dari pasien. Dapat di lihat pada gambar 3.4



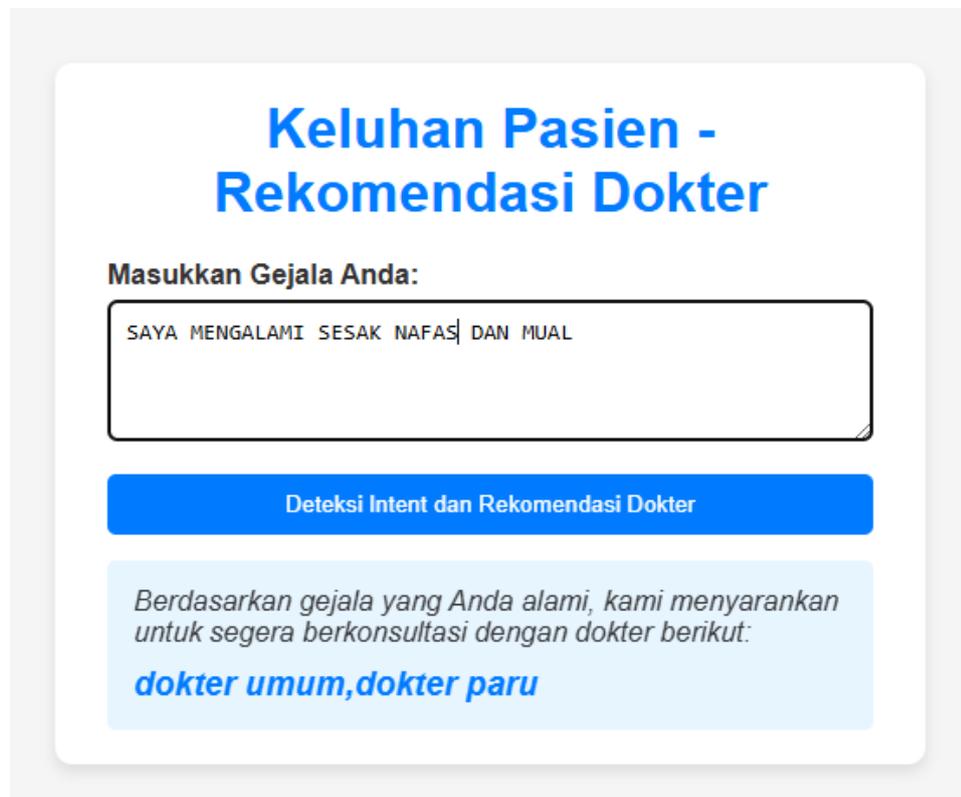
The image shows a user interface for a doctor recommendation system. At the top, there is a blue header with the text "Keluhan Pasien - Rekomendasi Dokter". Below the header, there is a text input field with the label "Masukkan Gejala Anda:". Inside the input field, there is a placeholder text: "Contoh: Saya mengalami batuk dan sesak napas.....". Below the input field, there is a blue button with the text "Deteksi Intent dan Rekomendasi Dokter".

Gambar 3. 4 Halaman Awal Sistem Rekomendasi Dokter

Pada gambar 3.4 Menampilkan halaman input teks yang akan digunakan user dalam memasukan teks gejala penyakit yang dialami oleh user lalu dibawahnya terdapat tombol pencarian rekomendasi dokter.

3.4.2 Halaman Hasil Sistem Rekomendasi Dokter

Pada halaman ini menampilkan hasil dari inputan teks yang dimasukkan oleh user. Hasilnya dapat di lihat pada gambar 3. Dibawah ini



**Keluhan Pasien -
Rekomendasi Dokter**

Masukkan Gejala Anda:

SAYA MENGALAMI SESAK NAFAS DAN MUAL

Deteksi Intent dan Rekomendasi Dokter

Berdasarkan gejala yang Anda alami, kami menyarankan untuk segera berkonsultasi dengan dokter berikut:

dokter umum, dokter paru

Gambar 3. 5 Halaman Hasil Sistem Rekomendasi Dokter

Pada gambar 3.5 Menampilkan hasil dari inputan teks user yang memberikan rekomendasi dokter yang sesuai. Dalam proses inputan teks sistem akan memberikan rekomendasi dokter yang harus dikunjungi atau di konsultasikan oleh user.

BAB IV

HASIL ANALISIS DAN PENELITIAN

4.1 Hasil Penelitian

4.1.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset keluhan medis dan spesialis dokter yang tersedia di sumber dataset terbuka Kaggle. Data tersebut berisi kumpulan jenis penyakit, gejala penyakit dan Dokter spesialis yang berjumlah 1004 data dengan masing-masing kategori keluhan telah diberi label yang sesuai.

4.1.2 Pra-Pemrosesan Data

Data yang telah dikumpulkan sebelumnya kemudian dilakukan pra-pemrosesan data untuk memastikan bahwa data telah siap digunakan dalam model BioBert agar hasil yang diperoleh sesuai dan akurat. Langkah-langkah yang dilakukan dalam pra-pemrosesan meliputi :

- a *Tokenisasi*
Memisahkan teks menjadi kata-kata atau token yang lebih kecil sehingga dapat dimnegerti sistem dengan mudah.
- b *Stopworld Removal*
Menghapus kata-kata umum yang tidak memiliki kaitan yang signifikan dalam proses analisis intent.
- c *Stemming*
Mengubah kata-kata tersebut kedalam bentuk dasarnya untuk mempermudah dalam pemrosesan teks.

4.1.3 Implementasi Model BioBert

Setelah data diproses, selanjutnya adalah memasukan teks ke dalam model *BioBert* untuk melakukan deteksi intent. *BioBert* (*Bidirectional Encoder Representations from Transformers for Biomedical Text Mining*) merupakan pengembangan dari model *BERT* yang telah dilatih secara khusus pada domain teks biomedis, sehingga lebih efektif untuk menangani data yang berkaitan

dengan bidang kesehatan. Proses memasukkan teks ke dalam model *BioBert* meliputi :

a Embedding

Pada tahap ini, teks input diubah menjadi representasi numerik agar dapat diproses oleh model BioBert . Prosesnya meliputi Tokenisasi yang menggunakan tokenizer dari BERT (WordPiece Tokenizer) untuk membagi teks menjadi sub-kata (tokens). Hasil dari tahap embedding ini adalah representasi vektor dari teks yang dapat digunakan oleh model BioBert untuk pemrosesan lebih lanjut.

b Encoding

Tahap encoding dilakukan menggunakan arsitektur transformer dari BioBert , yang terdiri dari beberapa *layer self-attention* untuk memahami konteks kata dalam teks medis. BioBert menganalisis hubungan antara setiap kata dengan kata lain dalam satu kalimat atau antar kalimat.

c Klasifikasi intent

Pada tahap ini, output dari encoder diteruskan ke lapisan klasifikasi untuk menentukan *intent* dari teks input. Setelah intent diklasifikasikan, sistem dapat melanjutkan ke tahap selanjutnya, yaitu mapping rekomendasi dokter.

d Mapping Rekomendasi Dokter

Setelah intent dikenali, sistem melakukan pemetaan ke dokter yang sesuai dengan gejala atau kebutuhan pasien. sistem akan mencocokkan gejala dengan daftar spesialisasi dokter di database

e Output Rekomendasi Dokter

Sistem akan memberikan hasil rekomendasi dokter yang relevan dari input teks yang telah di masukan ke sistem .

4.2 Hasil Evaluasi

4.2.1 Model *BioBert*

Proses penelitian ini menggunakan beberapa library yang mendukung bahasa pemrograman python yang dapat diimplemetasikan pada model BioBert .

```

original_text \
0      lelah muntah nyeri dada
1 demam sakit kepala leher kaku mual muntah peka...
2      mati rasa mual pusing nyeri dada demam
3      sakit perut pusing mati rasa batuk
4 nyeri bengkak merah hangat nyeri bagi kena ses...
5 hasil tes pap tidak normal darah vagina tidak ...
6      ruam kulit merah rasa bakar gatal
7 lelah sesak napas pale skin lightheadedness di...
8      demam diarrhea abdominal cramps mual muntah
9      batuk mati rasa pusing muntah gatal
10     lelah pusing batuk

tokenized_text \
0 ['le', '##lah', 'm', '##unta', '##h', 'n', '##...
1 ['dem', '##am', 'sa', '##ki', '##t', 'k', '##e...
2 ['mat', '##i', 'r', '##asa', 'm', '##ual', 'pu...
3 ['sa', '##ki', '##t', 'per', '##ut', 'pu', '##...
4 ['n', '##yer', '##i', 'ben', '##g', '##ka', '#...
5 ['has', '##il', 'te', '##s', 'p', '##ap', 't',...
6 ['r', '##ua', '##m', 'k', '##uli', '##t', 'me'...
7 ['le', '##lah', 'se', '##sa', '##k', 'nap', '#...
8 ['dem', '##am', 'di', '##ar', '##r', '##hea', ...
9 ['bat', '##uk', 'mat', '##i', 'r', '##asa', 'p...
10  ['le', '##lah', 'pu', '##sing', 'bat', '##uk']

tokenized_ids
0 [101, 5837, 10358, 182, 19529, 1324, 183, 7904...
1 [101, 26707, 2312, 21718, 2293, 1204, 180, 804...
2 [101, 22591, 1182, 187, 18384, 182, 4746, 2360...
3 [101, 21718, 2293, 1204, 1679, 3818, 23609, 42...
4 [101, 183, 7904, 1182, 26181, 1403, 1968, 1377...
5 [101, 1144, 2723, 21359, 1116, 185, 11478, 189...
6 [101, 187, 6718, 1306, 180, 15818, 1204, 1143,...
7 [101, 5837, 10358, 14516, 3202, 1377, 22956, 2...
8 [101, 26707, 2312, 4267, 1813, 1197, 13836, 24...
9 [101, 7693, 7563, 22591, 1182, 187, 18384, 236...
10 [101, 5837, 10358, 23609, 4253, 7693, 7563, 10...

```

Gambar 4. 1 Tokenisasi

Terlihat pada Gambar 4.1 merupakan hasil dari proses tokenisasi yang dilakukan pada text dengan mengubah text menjadi vektor atau angka agar mudah dipahami oleh sistem.

```

5) before_stopwords \
0 nyeri,bengkak,kemerahan,hangat,nyeri di bagian...
1 mual,mati rasa,sakit perut,demam,nyeri dada
2 nyeri, kemerahan, radang pada mata
3 demam,mati rasa,sakit perut,gatal-gatal,pusing
4 demam,diarrhea,abdominal cramps,mual,muntah
5 demam,kelelahan,pusing
6 sakit perut,muntah,pusing,gatal-gatal
7 mual,nyeri dada,batuk,kelelahan
8 Ruam, kulit memerah, rasa terbakar, gatal
9 kesulitan berinteraksi sosial, perilaku repeti...
10 demam,sakit perut,kelelahan,nyeri dada

after_stopwords
0 nyeri,bengkak,kemerahan,hangat,nyeri bagian te...
1 mual,mati rasa,sakit perut,demam,nyeri dada
2 nyeri, kemerahan, radang mata
3 demam,mati rasa,sakit perut,gatal-gatal,pusing
4 demam,diarrhea,abdominal cramps,mual,muntah
5 demam,kelelahan,pusing
6 sakit perut,muntah,pusing,gatal-gatal
7 mual,nyeri dada,batuk,kelelahan
8 Ruam, kulit memerah, rasa terbakar, gatal
9 kesulitan berinteraksi sosial, perilaku repeti...
10 demam,sakit perut,kelelahan,nyeri dada

```

Gambar 4. 2 proses *stopworld removal*

Pada Gambar 4.2 merupakan proses *stopworld removal* dengan menghapus kata-kata yang tidak penting.

```

5) before_lemmatization \
0 demam,gatal-gatal,nyeri dada,muntah,kelelahan
1 sakit perut,mual,pusing,nyeri dada,batuk
2 Kesulitan mendapatkan, mempertahankan ereksi
3 nyeri, kemerahan, radang mata
4 mati rasa,sakit perut,mual,muntah
5 nyeri dada,batuk,demam
6 nyeri,bengkak,kemerahan,hangat,nyeri bagian te...
7 kehilangan ingatan, kebingungan, kesulitan ber...
8 kelelahan,pusing,muntah
9 sakit perut,batuk,pusing,demam
10 mual,pusing,sakit perut

after_lemmatization
0 demam , gatal - gatal , nyeri dada , muntah , ...
1 sakit perut , mual , pusing , nyeri dada , batuk
2 sulit dapat , tahan ereksi
3 nyeri , merah , radang mata
4 mati rasa , sakit perut , mual , muntah
5 nyeri dada , batuk , demam
6 nyeri , bengkak , merah , hangat , nyeri bagi ...
7 hilang ingat , bingung , sulit pikir , ubah pr...
8 lelah , pusing , muntah
9 sakit perut , batuk , pusing , demam
10 mual , pusing , sakit perut

```

Gambar 4. 3 proses lemmatisasi

Pada gambar 4.3 merupakan proses dari lemmatisasi. Dengan Mengubah kata-kata tersebut kedalam bentuk dasarnya untuk mempermudah dalam pemrosesan teks.

```

Epoch 1, Loss: 3.3863, Accuracy: 0.3740
Validation Accuracy: 0.5822
Epoch 2, Loss: 2.2649, Accuracy: 0.6358
Validation Accuracy: 0.7455
Epoch 3, Loss: 1.6763, Accuracy: 0.8018
Validation Accuracy: 0.8574
Epoch 4, Loss: 1.2666, Accuracy: 0.8860
Validation Accuracy: 0.9034
Epoch 5, Loss: 1.0764, Accuracy: 0.9242
Validation Accuracy: 0.9238
Epoch 6, Loss: 0.9904, Accuracy: 0.9416
Validation Accuracy: 0.9319
Epoch 7, Loss: 0.9452, Accuracy: 0.9515
Validation Accuracy: 0.9384
Epoch 8, Loss: 0.9220, Accuracy: 0.9540
Validation Accuracy: 0.9381
Epoch 9, Loss: 0.9104, Accuracy: 0.9540
Validation Accuracy: 0.9425
Epoch 10, Loss: 0.9068, Accuracy: 0.9570
Validation Accuracy: 0.9425
Epoch 11, Loss: 0.9058, Accuracy: 0.9589
Validation Accuracy: 0.9425
Epoch 12, Loss: 0.9052, Accuracy: 0.9568
Validation Accuracy: 0.9445
Epoch 13, Loss: 0.9069, Accuracy: 0.9556
...
Validation Accuracy: 0.9459
Epoch 20, Loss: 0.8852, Accuracy: 0.9535
Validation Accuracy: 0.9418
Rekomendasi dokter: dokter umum,dokter paru

```

Gambar 4. 4 Hasil Akurasi

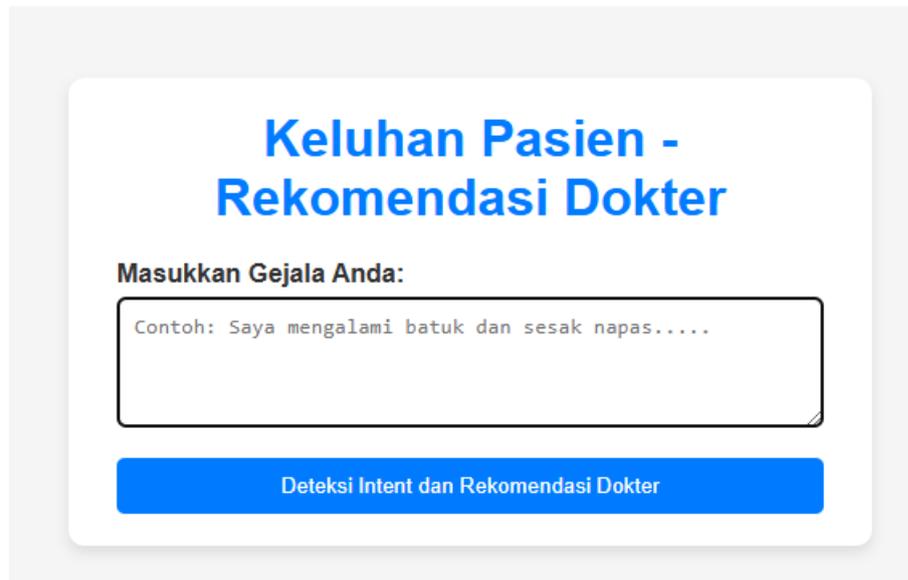
Pada gambar 4.4 merupakan hasil dari evaluasi model *BioBert* . Terlihat akurasi pada model menunjukkan hasil yang bagus.

4.3 Hasil Implementasi Antarmuka

Hasil dari penelitian ini adalah proses implementasi Sistem *Intent Detection* dalam mengklasifikasikan keluhan pasien untuk merekomendasikan dokter spesialis menggunakan metode *BioBert* . Sistem yang dikembangkan menggunakan *BioBert* , model berbasis *BERT* yang telah disesuaikan untuk domain biomedis. Model ini dilatih untuk mendeteksi *intent* berdasarkan teks keluhan pasien dan mengarahkannya ke spesialisasi medis yang sesuai.

4.2.1 Halaman Utama Sistem Rekomendasi Dokter

Pada penerapan desain antarmuka halaman awal sistem Rekomendasi Dokter, bagian appbar terdapat title dari aplikasi. Pada bagian bawah title terdapat boks teks yang digunakan untuk menginputkan teks gejala penyakit yang dialami oleh user. Lalu dibawahnya terdapat button untuk memproses inputan yang menghasilkan rekomendasi dokter dari gejala penyakit. Halaman awal sistem Rekomendasi Dokter dapat dilihat pada gambar 4.1



**Keluhan Pasien -
Rekomendasi Dokter**

Masukkan Gejala Anda:

Contoh: Saya mengalami batuk dan sesak napas.....

Deteksi Intent dan Rekomendasi Dokter

Gambar 4. 5 Halaman Utama Sistem Rekomendasi Dokter.

Terlihat pada gambar 4.5 Adalah tampilan awal yang akan di akses oleh user. Pada tampilan tersebut, user bisa menginputkan teks gejala penyakit yang dirasakan.

4.2.2 Halaman Hasil Sistem Rekomendasi

Hasil dari penarapan desain antarmuka halaman hasil rekomendasi dokter tampilannya hampir sama dengan halaman awal. Yang membedakan adalah didalam halaman hasil terdapat output hasil rekomendasi dokter dari gejala penyakit yang diinputkan dan telah di proses oleh sistem. Untuk tampilan hasilnya dapat dilihat pada gambar 4. 5



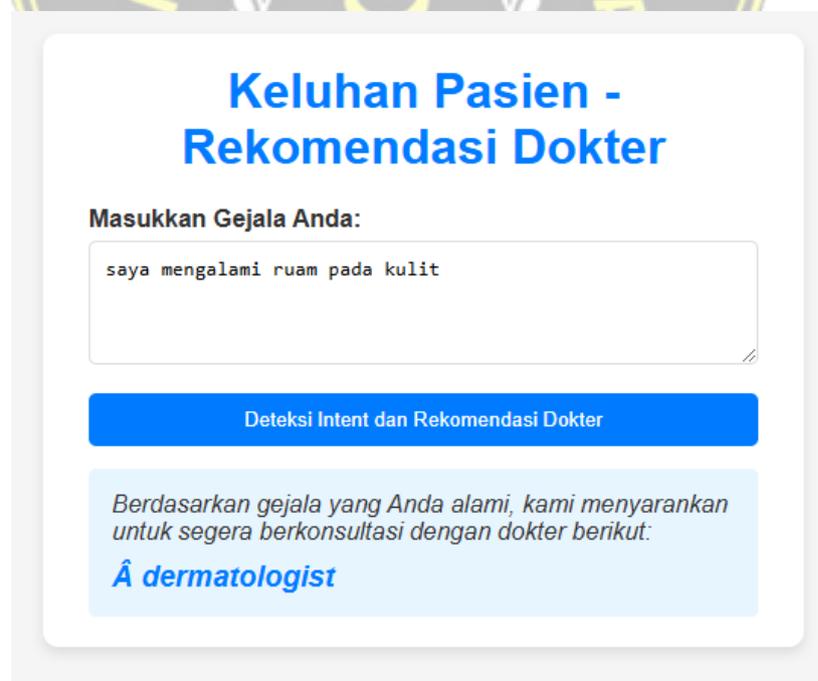
Keluhan Pasien - Rekomendasi Dokter

Masukkan Gejala Anda:

Deteksi Intent dan Rekomendasi Dokter

Gambar 4. 6 menginputkan teks

Pada gambar 4.6 Dapat dilihat proses demo dari inputan yang dimasukan oleh user lalu hasil dari sistem rekomendasinya dapat dilihat pada gambar 4.6



Keluhan Pasien - Rekomendasi Dokter

Masukkan Gejala Anda:

Deteksi Intent dan Rekomendasi Dokter

Berdasarkan gejala yang Anda alami, kami menyarankan untuk segera berkonsultasi dengan dokter berikut:

Â dermatologist

Gambar 4. 7 Hasil Rekomendasi dokter

Pada gambar 4.7 Terlihat hasil inputan yang dimasukan oleh user berhasil menampilkan rekomendasi dokter yang disarankan oleh sistem.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dari penelitian ini dapat ditarik kesimpulan dari hasil implementasi Sistem *Intent Detection* dalam mengklasifikasikan keluhan pasien untuk merekomendasikan dokter spesialis menggunakan metode *BioBert* dapat menghasilkan rekomendasi yang sesuai dan akurat. Sistem ini dikembangkan menggunakan BioBert, model berbasis BERT yang telah disesuaikan untuk domain biomedis. Model ini dilatih untuk mendeteksi intent berdasarkan teks keluhan pasien dan mengarahkannya ke spesialisasi medis yang sesuai. Dengan hasil akurasi yang menunjukkan angka 94%, sistem ini mampu mendeteksi intent keluhan pasien dengan tepat.

5.2 Saran

Berdasarkan hasil yang telah dicapai pada proses penelitian ini, peneliti menyarankan untuk melakukan

1. Pengujian lebih lanjut dengan dataset yang lebih besar guna memastikan keakuratan sistem.
2. Karena pada sistem ini hanya memberikan rekomendasi dokter dari gejala penyakit yang dirasakan, disarankan penelitian selanjutnya dapat memperluas ruang lingkupnya dengan menambahkan jenis penyakit atau obat yang dapat direkomendasikan
3. Pada penelitian ini sistem hanya mengerti bahasa yang dirasa baku, bahasa-bahasa yang sulit dimengerti tidak bisa terdeteksi. Penelitian selanjutnya disarankan untuk memperluas cakupannya sehingga user dapat menginputkan semua gejala yang dialami dan mendapat rekomendasi lebih spesifik.

DAFTAR PUSTAKA

- Al Fadil Syahputra, S., Mita Azizah, N., Aiman, J., Ainun Nikmah, D., & Rosyani, P. (2024). Identifikasi dan Prediksi Umur Berdasarkan Citra Wajah Menggunakan Deep Learning Algoritma Convolutional Neural Network(CNN). *Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan*, 2(1), 87–95.
- Awalina, A. (n.d.). *Indonesia 's Fake News Detection using Transformer Network*.
- Bei, F., & Sudin, S. (2021). Analisis Sentimen Aplikasi Tiket Online Di Play Store Menggunakan Metode Support Vector Machine (Svm). *Sismatik*, 01(01), 91–97.
- Berge, G. T., Granmo, O. C., Tveit, T. O., Munkvold, B. E., Ruthjersen, A. L., & Sharma, J. (2023). Machine learning-driven clinical decision support system for concept-based searching: a field trial in a Norwegian hospital. *BMC Medical Informatics and Decision Making*, 23(1), 1–15.
<https://doi.org/10.1186/s12911-023-02101-x>
- Bian, J., Jiang, R., Zhai, W., Huang, T., Zhou, H., & Zhu, S. (2023). *DMNER: Biomedical Entity Recognition by Detection and Matching*.
<http://arxiv.org/abs/2306.15736>
- Eldi, E., & Syaputra, H. (2020). Implementasi Chatbot Untuk Mendukung Sistem Informasi Pada Rumah Sakit Muhammadiyah Palembang. *Jurnal Nasional Ilmu Komputer*, 1(3), 139–148. <https://doi.org/10.47747/jurnalnik.v1i3.160>
- Fatharani, F., Kania, K. P., Hutahaean, J., & Wulan, S. R. (2022). Deteksi Intensi Chatbot Berbahasa Indonesia dengan Menggunakan Metode Capsule Network. *Journal of Information System Research (JOSH)*, 3(4), 590–596.
<https://doi.org/10.47065/josh.v3i4.1821>
- Halim, J. D. (2024). *Klasifikasi Sel Darah Putih Menggunakan Vision Transformer (ViT)*. 6(November), 291–307.
- Han, K., Xiao, A., Wu, E., Guo, J., Xu, C., & Wang, Y. (2021). 已看-TNT-Transformer in Transformer. *Advances in Neural Information Processing Systems*, 19(NeurIPS), 15908–15919.

- Homepage, J., Rahayu, S., Safaat Harahap, N., Agustian, S., Studi Teknik Informatika, P., & Sains dan Teknologi, F. (2024). *MALCOM: Indonesian Journal of Machine Learning and Computer Science Application of Langchain Technology to the Fiqh Question Answering System of Four Madhhab Penerapan Teknologi LangChain pada Question Answering System Fikih Empat Madzhab*. 4(3), 974–983.
- Joey Ferelestian, V., Susanto, B., & Senapartha, I. K. D. (2023). Pengembangan Telegram Chatbot Informasi Mahasiswa Menggunakan Wit.ai. *Jurnal Terapan Teknologi Informasi*, 7(2), 89–97.
<https://doi.org/10.21460/jutei.2023.72.257>
- Kumar, A. (2021). Disease Prediction and Doctor Recommendation System using Machine Learning Approaches. *International Journal for Research in Applied Science and Engineering Technology*, 9(VII), 34–44.
<https://doi.org/10.22214/ijraset.2021.36234>
- Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBert : A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4), 1234–1240.
<https://doi.org/10.1093/bioinformatics/btz682>
- Mahardhika Chandra, Rizki Pratama, Fathan Azka Pradana, & Alvita Bonita. (2022). Chatbot Interaksi Rumah Sakit menggunakan FFNN. *Indonesian Journal of Data and Science*, 3(1), 62–68.
<https://doi.org/10.56705/ijodas.v3i1.36>
- Mohiuddin, K., Welke, P., Alam, M. A., Martin, M., Alam, M. M., Lehmann, J., & Vahdati, S. (2023). Retention Is All You Need. *International Conference on Information and Knowledge Management, Proceedings, Nips*, 4752–4758.
<https://doi.org/10.1145/3583780.3615497>
- Nur Oktavia, A., Iqbal, M., Saputra, R. W., Zulfikar, M. I., & Saifudin, A. (2024). Implementasi Metode Natural Language Processing Dalam Studi Analisis. *Jurnal Ilmiah Ilmu Komputer dan Multimedia*, 2(1), 154–159.
<https://jurnalmahasiswa.com/index.php/biikma>
- Nurul Huda, Zuhroidah, I., Toha, M., & Sujarwadi, M. (2021). Pelatihan

Pertolongan Pertama Pada Kecelakaan (P3K) Pada Guru Pembina Dan Anggota Pmr. *Jurnal kreativitas pengabdian kepada masyarakat*, 4(2615–0921), 323–328. <https://doi.org/10.33024/jkpm.v4i2.3746>

Santoso, B. B., & Saian, P. O. N. (2023). Implementasi Flask Framework pada Development Modul Reporting Aplikasi Sistem Informasi Helpdesk di PT.XYZ). *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, 7(2), 217–226. <https://doi.org/10.35870/jtik.v7i2.718>

Setiawan, D., Karuniawati, E. A. D., & Janty, S. I. (2023). Peran Chat Gpt (Generative Pre-Training Transformer) Dalam Implementasi Ditinjau Dari Dataset. *INNOVATIVE: Journal of Social Science Research*, 3(3), 9527–9539. <https://j-innovative.org/index.php/Innovative/article/view/3286>

