BAB I
PENDAHULUAN

Bab ini menyajikan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup, dan sistematika penulisan skripsi mengenai klasifikasi tingkat risiko keluhan ibu hamil dan bayi berdasarkan teks dari forum kesehatan daring menggunakan model IndoBERT-CNN dan IndoBERT-BiLSTM. Latar belakang membahas urgensi pengembangan sistem klasifikasi risiko kesehatan ibu hamil dan bayi berbasis teks di Indonesia serta penelitian-penelitian terkait yang menjadi dasar dilakukannya penelitian ini. Berdasarkan latar belakang tersebut, dirumuskan masalah penelitian beserta tujuan dan manfaat yang ingin dicapai melalui pengembangan kedua model klasifikasi.

1.1 Latar Belakang

Kesehatan ibu hamil dan bayi merupakan salah satu prioritas utama dalam transformasi pelayanan kesehatan di Indonesia. Berdasarkan hasil *Long Form* Sensus Penduduk 2020, Angka Kematian Ibu (AKI) tercatat sebesar 189 per 100.000 kelahiran hidup, sementara Angka Kematian Bayi (AKB) berada pada level 16,85 per 1.000 kelahiran hidup (Badan Pusat Statistik, 2023). Meskipun menunjukkan tren penurunan dalam satu dekade terakhir, angka tersebut masih tergolong tinggi dan belum mencapai target nasional tahun 2024 sebesar 183 per 100.000 kelahiran hidup (Kementerian Kesehatan Republik Indonesia, 2023). Selain itu, Indonesia masih menghadapi tantangan besar untuk mencapai target *Sustainable Development Goals* (SDGs) 2030 dalam mengurangi kematian maternal menjadi di bawah 70 per 100.000 kelahiran hidup (Syairaji dkk., 2024). Kondisi ini menunjukkan bahwa deteksi dini terhadap risiko kesehatan ibu hamil masih menjadi tantangan serius yang perlu ditangani.

Salah satu faktor utama yang menyebabkan masih tingginya angka kematian ibu di Indonesia adalah keterlambatan dalam penegakan diagnosis dan perujukan ke fasilitas kesehatan yang memadai. Kementerian Kesehatan Republik Indonesia (2023) mengidentifikasi tiga keterlambatan utama yang menjadi penyebab ibu hamil berisiko tidak tertolong, yaitu terlambat mengambil keputusan, terlambat sampai di tempat rujukan, dan terlambat mendapat penanganan. Kondisi ini menunjukkan bahwa deteksi dini terhadap

tingkat risiko keluhan kesehatan ibu hamil dan bayi menjadi hal yang krusial untuk mencegah komplikasi yang berujung pada kematian.

Seiring berkembangnya teknologi informasi, masyarakat semakin memanfaatkan platform digital untuk mengakses layanan kesehatan, termasuk menyampaikan keluhan kesehatan secara daring. Salah satu platform konsultasi kesehatan daring yang banyak digunakan di Indonesia adalah Alodokter, yang memiliki lebih dari 30 juta pengguna aktif dengan seluruh jawaban dokter diberikan oleh dokter terverifikasi (Alodokter, 2020). Platform ini menghasilkan data teks dalam jumlah besar berupa pasangan pertanyaan pengguna dan jawaban dokter yang dapat dimanfaatkan untuk analisis kesehatan berbasis teks, dan telah digunakan dalam penelitian klasifikasi teks kesehatan berbahasa Indonesia sebelumnya oleh Vihikan & Trisna (2024).

Penelitian pada domain teks kesehatan telah dilakukan baik di tingkat internasional maupun nasional. Lu dkk. (2022) melakukan klasifikasi kondisi penyakit berdasarkan catatan medis pasien *(medical notes)* menggunakan berbagai model *deep learning*, yaitu *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), *Bidirectional Long Short-Term Memory* (BiLSTM), *Bidirectional Encoder Representations from Transformers* (BERT), dan *Transformer Encoder*. Pada konteks bahasa Indonesia, Abdillah dkk. (2023) mengembangkan model *Biomedical Named Entity Recognition* (BioNER), yaitu tugas ekstraksi informasi untuk mengenali entitas medis seperti gejala, penyakit, obat, dan istilah biomedis dari teks konsultasi kesehatan daring. Meskipun berfokus pada tugas *Named Entity Recognition* dan bukan klasifikasi teks, penelitian tersebut menunjukkan bahwa teknik pemrosesan bahasa alami pada domain kesehatan berbahasa Indonesia telah dikembangkan. Studi tersebut menunjukkan bahwa data konsultasi kesehatan berbahasa Indonesia dapat diolah untuk menghasilkan informasi yang bermanfaat pada domain kesehatan.

Secara umum, beberapa penelitian menunjukkan bahwa model berbasis transformer mengungguli arsitektur *deep learning* konvensional. Penelitian oleh Vihikan & Trisna (2024) menemukan bahwa IndoBERT mencapai akurasi 87,9% dan *F1-score* 87,9% pada klasifikasi pertanyaan kesehatan dari Alodokter, mengungguli CNN (akurasi 87,6%, F1 87,1%) dan BiLSTM (akurasi 87,3%, F1 85,7%). Hasil serupa ditemukan oleh Abdillah dkk.

(2023) yang menunjukkan model berbasis *transformer* mencapai *F1-score* 76,91% dan mengungguli LSTM pada teks konsultasi kesehatan berbahasa Indonesia.

Meskipun demikian, CNN dan BiLSTM tetap relevan dalam kondisi tertentu. Lu dkk. (2022) menemukan bahwa CNN mampu menghasilkan performa yang sebanding dengan *Transformer Encoder* pada kondisi ketidakseimbangan kelas tertentu dengan efisiensi komputasi yang jauh lebih tinggi. Sementara itu, Al-Smadi (2024) dan Murfi dkk. (2024) menunjukkan bahwa kombinasi *transformer* dengan LSTM maupun CNN menghasilkan performa yang lebih baik dibandingkan *transformer* tunggal. Al-Smadi (2024) memperoleh *micro-F1* 84% dengan DeBERTa-BiLSTM, sedangkan Murfi dkk. (2024) mencatat akurasi 87,68% dengan BERT-LSTM-CNN pada analisis sentimen berbahasa Indonesia.

Berdasarkan berbagai penelitian tersebut, model berbasis *transformer* menunjukkan performa yang unggul dalam memahami konteks bahasa, termasuk pada domain kesehatan. Di sisi lain, CNN dan BiLSTM masih banyak digunakan sebagai lapisan tambahan karena memiliki karakteristiknya dalam mengekstraksi informasi dari teks. CNN efektif dalam menangkap pola lokal antar token atau frasa, sedangkan BiLSTM mampu memodelkan dependensi kontekstual dua arah dalam urutan teks.

Meskipun berbagai penelitian telah menerapkan model *transformer*, CNN, dan BiLSTM pada tugas klasifikasi teks kesehatan maupun tugas NLP lainnya, belum ditemukan penelitian yang secara khusus untuk klasifikasi tingkat risiko keluhan tentang ibu hamil dan bayi berasarkan teks berbahasa Indonesia. Oleh karena itu, penelitian ini mengembangkan dua model klasifikasi tingkat risiko keluhan ibu hamil dan bayi berbasis teks menggunakan IndoBERT yang dikombinasikan dengan CNN dan BiLSTM. Kedua model dilatih menggunakan data teks keluhan dari forum kesehatan daring Alodokter dan dievaluasi kinerjanya dalam mengklasifikasikan keluhan ke dalam tiga tingkatan risiko, yaitu rendah, sedang, dan tinggi. Penggunaan tiga tingkatan kelas ini mengacu pada penelitian klasifikasi kesehatan maternal, sebagaimana diterapkan pada *dataset* UCI *Maternal Health Risk* oleh Ahmed (2020) yang juga menggunakan tiga kelas serupa: *low risk, mid risk,* dan *high risk*. Hasil perbandingan kedua model dapat memberikan rekomendasi arsitektur yang paling efektif untuk tugas klasifikasi risiko kesehatan berbasis teks bahasa Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana optimasi model IndoBERT-CNN dan IndoBERT-BiLSTM dilakukan dalam mengklasifikasikan risiko ibu hamil dan bayi berdasarkan teks keluhan?
2. Bagaimana perbandingan kinerja model IndoBERT-CNN dan IndoBERT-BiLSTM dalam mengklasifikasikan risiko ibu hamil dan bayi berdasarkan teks keluhan?

1.3 Tujuan dan Manfaat Penelitian

Tujuan utama penelitian ini adalah membangun model klasifikasi tingkat risiko keluhan ibu hamil dan bayi berbasis teks menggunakan IndoBERT-CNN dan IndoBERT-BiLSTM. Adapun tujuan khusus penelitian ini adalah:

1. Mengimplementasikan optimasi model IndoBERT-CNN dan IndoBERT-BiLSTM dalam mengklasifikasikan risiko ibu hamil dan bayi berdasarkan teks keluhan.
2. Membandingkan kinerja model IndoBERT-CNN dan IndoBERT-BiLSTM untuk menentukan model yang memberikan performa terbaik dalam mengklasifikasikan risiko ibu hamil dan bayi berdasarkan teks keluhan.

Adapun manfaat yang diharapkan dari penelitian ini adalah tersedianya model klasifikasi risiko keluhan ibu hamil dan bayi berdasarkan teks keluhan yang dapat membantu pengguna memperoleh identifikasi awal tingkat risiko dari keluhan yang disampaikan. Dengan adanya sistem klasifikasi tersebut, pengguna dapat memperoleh gambaran awal mengenai tingkat urgensi keluhannya sehingga lebih terbantu dalam menentukan langkah penanganan atau konsultasi lanjutan yang sesuai. Selain itu, hasil penelitian ini dapat menjadi referensi bagi peneliti di bidang *Natural Language Processing* (NLP) yang ingin mengembangkan sistem klasifikasi teks kesehatan berbahasa Indonesia dengan karakteristik data yang berbeda.

1.4 Ruang Lingkup Penelitian

Ruang lingkup penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan teks pertanyaan tentang ibu hamil dan bayi beserta jawabannya yang dikumpulkan dari forum kesehatan daring Alodokter (alodokter.com) melalui teknik *web scraping* yang dilakukan pada 24 November 2025

dengan data yang terkumpul yaitu pada rentang 17 April 2017−21 November 2025. Proses *scraping* dibatasi pada halaman komunitas dengan *tag* hamil, kehamilan, ibu hamil, gangguan kehamilan, bayi, dan anak.

2. Klasifikasi risiko dibagi menjadi tiga kelas, yaitu rendah, sedang, dan tinggi, yang ditentukan berdasarkan isi jawaban dokter. Penggunaan ketiga kelas ini mengacu pada *dataset* UCI *Maternal Health Risk* oleh Ahmed (2020).
3. Penelitian ini tidak membandingkan berbagai model klasifikasi lain seperti *Multi Layer Perceptron* (MLP), *Long Short-Term Memory* (LSTM), atau model-model berbasis jaringan saraf lainnya.
4. Penelitian ini tidak membandingkan kinerja model yang menggunakan teknik augmentasi data dengan model tanpa augmentasi data.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini menyajikan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup, dan sistematika penulisan skripsi mengenai klasifikasi tingkat risiko keluhan ibu hamil dan bayi menggunakan model IndoBERT-CNN dan IndoBERT-BiLSTM.

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan tinjauan pustaka yang berhubungan dengan penelitian sebagai landasan teori dan analisis permasalahan. Tinjauan pustaka meliputi penelitian terdahulu terkait *multiclass classification*, pra pemrosesan data, augmentasi data, BERT, IndoBERT, CNN, BiLSTM, *hyperparameter,* optimasi *hyperparameter*, dan metrik evaluasi serta *confusion matrix*.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan yang dilakukan dalam penelitian, meliputi pengumpulan data, pelabelan data, pra pemrosesan data, pembagian data, augmentasi data, tokenisasi, pengembangan model IndoBERT-CNN dan IndoBERT-BiLSTM, serta evaluasi kinerja model.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menguraikan hasil dan analisis dari setiap tahapan penelitian yang telah dilakukan. Pembahasan dimulai dari hasil pengumpulan data, pelabelan data,

hasil pra pemrosesan teks, pembagian data, hingga hasil augmentasi data. Selanjutnya dibahas hasil tokenisasi dan pengembangan model IndoBERT-CNN dan IndoBERT-BiLSTM yang meliputi hasil optimasi *hyperparameter* serta hasil pelatihan model. Bab ini juga menyajikan hasil evaluasi kinerja kedua model untuk melihat pola kesalahan klasifikasi pada masing-masing kelas. Bagian akhir bab ini membahas perbandingan kinerja kedua model beserta interpretasi temuan berdasarkan hasil analisis yang telah dilakukan.

BAB V KESIMPULAN

Bab ini menyajikan kesimpulan dari keseluruhan uraian yang telah dijabarkan pada bab-bab sebelumnya serta saran untuk pengembangan penelitian selanjutnya. Kesimpulan mencakup ringkasan temuan utama terkait kinerja model IndoBERT-CNN dan IndoBERT-BiLSTM dalam mengklasifikasikan tingkat risiko keluhan ibu hamil dan bayi, serta hasil perbandingan kedua model berdasarkan metrik evaluasi yang digunakan. Saran difokuskan pada peningkatan metodologi penelitian dan eksplorasi lebih lanjut dalam konteks klasifikasi risiko kesehatan maternal dan bayi berbasis teks bahasa Indonesia.