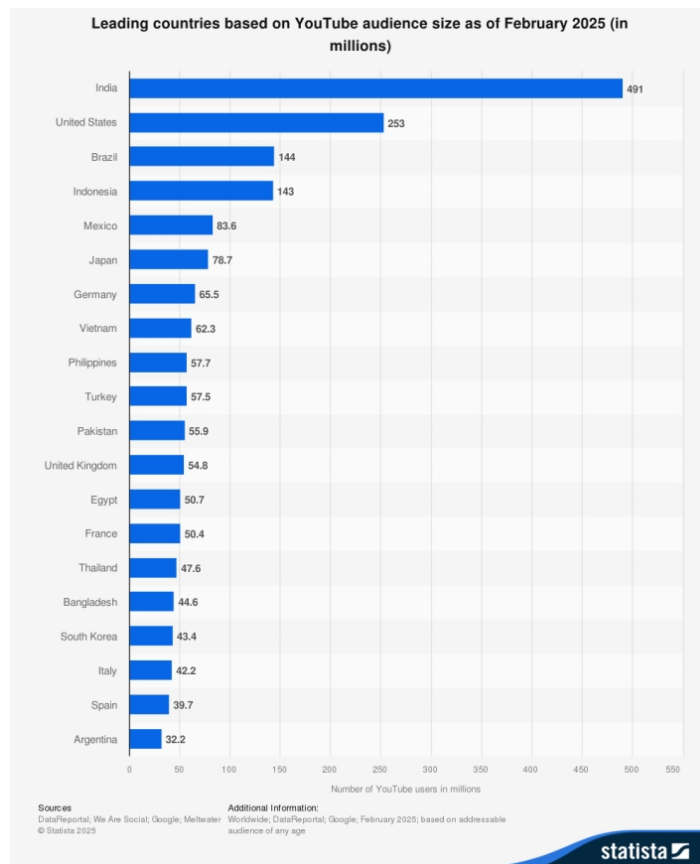


BAB I

PENDAHULUAN

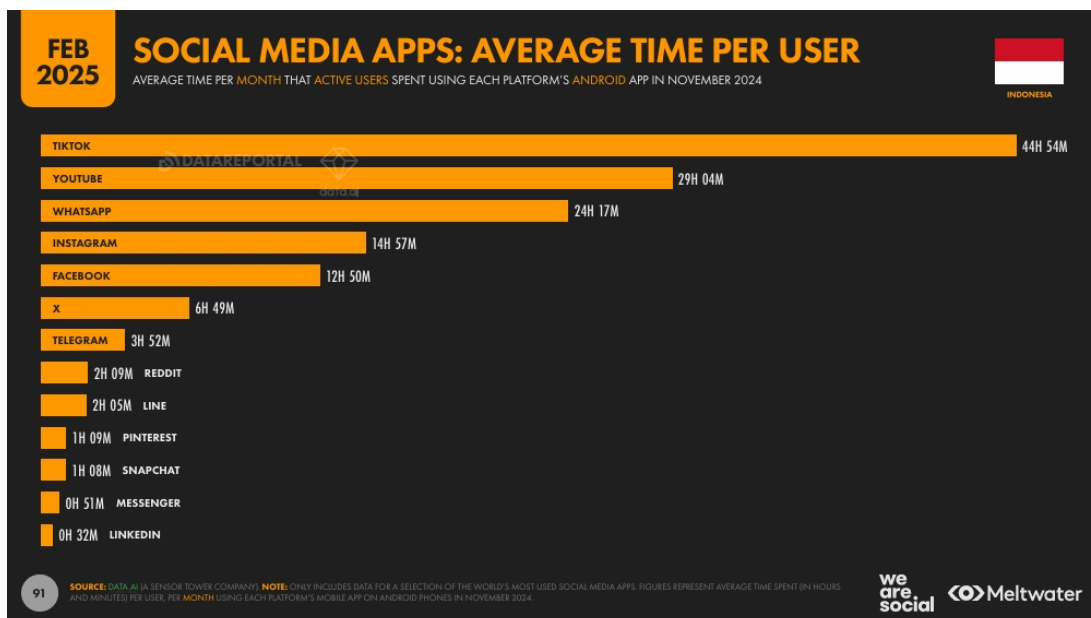
1.1 Latar Belakang

Perkembangan teknologi digital dalam beberapa tahun terakhir telah mengubah secara signifikan cara masyarakat Indonesia berkomunikasi, berinteraksi, dan mengekspresikan pendapat terhadap berbagai isu publik. Salah satu platform digital yang paling dominan adalah YouTube, yang kini berperan sebagai ruang diskusi terbuka antara masyarakat dan pemerintah. Berdasarkan data statistik per Februari 2025, jumlah pengguna aktif YouTube di Indonesia mencapai sekitar 143 juta orang, menjadikannya salah satu platform media sosial dengan basis pengguna terbesar di tanah air serta menempatkan Indonesia di posisi keempat secara global [1]. Kondisi ini memperkuat peran YouTube sebagai sumber utama untuk memantau opini publik, terutama ketika terjadi peristiwa sosial dan politik yang memicu reaksi emosional masyarakat secara luas.



Gambar 1.1 Statistik Pengguna Youtube di Dunia per Februari 2025

Selain memiliki jumlah pengguna yang besar, YouTube juga menunjukkan tingkat keterlibatan (*engagement*) yang tinggi di kalangan masyarakat Indonesia. Menurut laporan Digital 2025 Indonesia yang diterbitkan oleh *We Are Social* dan Meltwater, rata-rata pengguna aktif di Indonesia menghabiskan waktu 29 jam 4 menit per bulan untuk menggunakan aplikasi YouTube, menjadikannya platform kedua paling banyak diakses setelah TikTok (44 jam 54 menit per bulan) [2]. Durasi penggunaan yang tinggi ini mencerminkan intensitas interaksi masyarakat terhadap konten video digital, sekaligus menegaskan posisi YouTube sebagai media kedua dalam penyebaran informasi, hiburan, serta ekspresi opini publik di ruang digital nasional.



Gambar 1.2 Statistik *Screentime* Sosial Media di Indonesia

Fenomena dominasi YouTube ini menjadi semakin relevan dalam konteks dinamika sosial-politik di Indonesia. Pada tahun 2025, salah satu peristiwa yang memicu respons publik secara luas adalah insiden demonstrasi pengemudi ojek online, di mana Affan Kurniawan menjadi korban terlindas kendaraan taktis polisi saat aksi berlangsung [3]. Kejadian tersebut memicu gelombang kemarahan masyarakat, terlebih karena bertepatan dengan munculnya berita kenaikan gaji anggota DPR, yang dinilai tidak sensitif terhadap kondisi ekonomi rakyat [4]. Sebagai bentuk klarifikasi dan respons atas situasi tersebut, Sekretariat Presiden Republik Indonesia kemudian menggelar konferensi pers Presiden pada 31

Agustus 2025, yang disiarkan secara langsung melalui kanal YouTube resmi Sekretariat Presiden, dan menjadi pusat perhatian publik di ruang digital.

	A	B
507	2025-08-31T12:04:44Z	Berikan rakyatmu nafas biar kami merasakan kekayaan hasil alam negeri pertiwi
508	2025-08-31T12:04:43Z	Panglima Tertinggi di negeri ini akhirnya berpidato.. Ini yg di tunggu tunggu.. Tapi harus semua dibenahi dan dilaksanakan dgn segera jgn hnya beretorika saja.. 1.benahi dan bersihkan tikus tikus di sarang DPR dan pemerintahan. 2.normalkan pajak. 3.karyawan rakyat diberi gaji saja jgn ada tunjangan tunjangan.. Includ dlm 1 gaji. 4.segera saahkan tuu perampasan aset bagi koruptor dgn nilai 100 jt ke atas. 5.naikan UMR pekerja swasta jgn hny menaikkan gaji pegawai pemerintah saja. Ini usulan rakyat dari desa terpencil..
509	2025-08-31T12:04:39Z	Mohon maaf pak presiden, semoga pemerintahan Indonesia dalam lindungan Allah SWT, sehat dan panjang umur
510	2025-08-31T12:04:34Z	Sayangnya TELAT MIKIR lo ya? Jarena Kakau bicara tentang masalah indikasi Makar dan Terorisme... Itu indikasinya sdh jelas lo ya? Sayangnya Pemerintahan dan aparat penegak hukum tidak berani bertindak tegas. Salah satunya adanya Gerombolan yg ingin mendirikan idiologi khilafah yg selalu berkoar2 menantang perang. Masa iya Pemerintahan Buta dan tuli terhadap indikasi makar dan terorisme
511	2025-08-31T12:04:30Z	Tolong jangan ada suami dan istri yg sama2 anggota DPR. DPR adalah penyaluran aspirasi rakyat, bukan tempat menumpuk kekayaan SELAMATKAN INDONESIA BAPAK.... Sebelumnya mohon

Gambar 1.3 Komentar Publik di YouTube

Pada Gambar 1.3 Komentar yang muncul di kolom komentar YouTube menunjukkan variasi emosi yang sangat kaya. Contohnya, komentar seperti “Mohon maaf pak presiden, semoga pemerintahan Indonesia dalam lindungan Allah SWT, sehat dan panjang umur” mencerminkan emosi love, sementara kalimat “Panglima tertinggi di negeri ini akhirnya berpidato... Ini yang ditunggu-tunggu rakyat!” menunjukkan happiness. Sebaliknya, komentar seperti “Sayangnya TELAT MIKIR lo ya? Pemerintahan dan aparat penegak hukum tidak berani bertindak tegas” mengandung emosi anger, sedangkan “Sayangnya pemerintah buta dan tuli terhadap indikasi makar dan terorisme” mencerminkan fear, dan “Berikan rakyatmu nafas biar kami merasakan kekayaan hasil alam negeri pertiwi” menunjukkan sadness. Lima kategori emosi ini dipilih karena mewakili spektrum utama emosi publik dalam konteks sosial-politik dan telah digunakan secara luas dalam penelitian klasifikasi emosi pada media sosial [5], [6].

Analisis emosi publik menjadi semakin penting karena dapat memberikan gambaran objektif mengenai persepsi masyarakat terhadap pernyataan negara.

Informasi ini dapat digunakan sebagai masukan strategis bagi pemerintah, maupun peneliti sosial untuk memahami dinamika opini publik secara lebih akurat. Namun, tantangan utama dalam menganalisis komentar media sosial terletak pada karakteristik bahasa yang informal, penuh singkatan, emoji, sarkasme, serta komentar yang ambigu, sehingga pendekatan statistik konvensional sering kali tidak memadai [7].

Untuk mengatasi permasalahan ini, penelitian ini menggunakan pendekatan berbasis model deep learning BERT (Bidirectional Encoder Representations from Transformers). BERT menjadi pilihan tepat karena bersifat bidirectional mekanisme self-attention-nya mampu memahami konteks dari dua arah (kiri dan kanan) memberikan representasi semantik yang lebih dalam dan kontekstual [8], [9]. Berbagai penelitian terkini juga menunjukkan bahwa BERT merupakan model yang sangat unggul dibandingkan pendekatan tradisional seperti LSTM, CNN, dan Word2Vec. Penelitian [10] melaporkan bahwa BERT mampu meningkatkan akurasi klasifikasi emosi hingga 12–18% dibandingkan model berbasis RNN. Temuan ini diperkuat oleh penelitian [7] yang menegaskan bahwa contextual embedding dua arah milik BERT memungkinkan model ini menangkap nuansa emosional yang kompleks dalam teks pendek seperti komentar YouTube. Selain itu, pendekatan berbasis transformer terbukti lebih tahan terhadap variasi bahasa informal dan lebih efektif dalam menangkap makna implisit dalam komunikasi daring [11], [12].

Analisis emosi publik juga membutuhkan kemampuan mengenali entitas penting dalam teks, seperti tokoh publik, institusi, atau lokasi. Misalnya, komentar “Saya bangga dengan Presiden X” dan “Saya kecewa dengan Menteri X” menunjukkan arah emosi terhadap entitas yang berbeda. Oleh karena itu, integrasi teknik Named Entity Recognition (NER) dengan BERT menjadi penting untuk memperkaya representasi makna teks. Sejumlah penelitian menunjukkan bahwa kombinasi BERT dan NER mampu meningkatkan akurasi analisis emosi karena memperkuat pemahaman konteks semantik dengan informasi entitas yang relevan [13], [14].

Selain itu, ekspresi emosi dalam komentar publik sering kali bersifat multidimensional, di mana satu kalimat dapat mengandung lebih dari satu emosi.

Misalnya, komentar “saya sedih tapi tetap bangga” dapat mencerminkan sadness dan love secara bersamaan. Oleh karena itu, penelitian ini menggunakan pendekatan multilabel emotion classification yang memungkinkan satu teks memiliki lebih dari satu label emosi. Pendekatan ini lebih representatif terhadap ekspresi emosional masyarakat Indonesia yang kaya dan kontekstual [15].

Meskipun penelitian tentang klasifikasi emosi pada media sosial telah banyak dilakukan, sebagian besar studi terdahulu masih berfokus pada teks pendek dalam bahasa Inggris, dengan sumber data dominan berasal dari platform global seperti Twitter dan Reddit [16], [17]. Selain itu, sebagian besar pendekatan masih menggunakan single-label emotion classification yang membatasi representasi emosi publik yang kompleks [18]. Penelitian yang secara spesifik mengeksplorasi komentar publik dalam konteks kebijakan sosial-politik Indonesia, khususnya dari kanal resmi pemerintah di YouTube, masih sangat terbatas [19].

Lebih jauh lagi, pendekatan berbasis entitas dalam multilabel emotion classification pada media sosial telah mulai dikembangkan menggunakan teknik graph-based representation [20]. Namun, integrasi eksplisit antara BERT dan Named Entity Recognition, khususnya dalam konteks teks berbahasa Indonesia, masih sangat terbatas. Padahal, pendekatan ini memiliki potensi besar untuk menghasilkan representasi semantik yang lebih kaya dan relevan, terutama dalam konteks analisis emosi publik terhadap pernyataan pemerintah. Oleh karena itu, penelitian ini menawarkan kebaruan dengan mengembangkan model klasifikasi emosi multilabel berbasis integrasi BERT dan NER untuk komentar publik berbahasa Indonesia, sehingga mampu memberikan pemetaan emosi yang lebih komprehensif dan kontekstual terhadap isu-isu sosial-politik nasional.

Dengan demikian, penelitian ini berfokus pada pengembangan model analisis emosi multilabel pada komentar publik terhadap konferensi pers Presiden yang disiarkan melalui kanal YouTube Sekretariat Presiden. Integrasi BERT dan NER diharapkan mampu meningkatkan akurasi klasifikasi emosi dengan menangkap konteks linguistik dua arah sekaligus mengenali entitas penting dalam teks, sehingga memberikan pemetaan emosi publik yang lebih komprehensif terhadap pernyataan pemerintah.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja integrasi BERT dan *Named Entity Recognition* (NER) dalam menganalisis emosi multilabel pada komentar publik YouTube terkait konferensi pers Presiden Republik Indonesia?
2. Bagaimana hasil pembuatan sistem analisis emosi multilabel berbasis integrasi BERT dan *Named Entity Recognition* (NER) pada komentar publik YouTube terkait konferensi pers Presiden Republik Indonesia?

1.3 Batasan Masalah

Batasan penelitian ini meliputi:

1. Data penelitian diperoleh dari komentar publik pada satu video konferensi pers Presiden yang diunggah melalui kanal YouTube resmi Sekretariat Presiden pada tanggal 31 Agustus 2025.
2. Data penelitian diambil dalam rentang waktu 31 Agustus 2025 hingga 12 Oktober 2025, dengan jumlah komentar 10.521.
3. Komentar yang dianalisis berbahasa Indonesia dan tidak mengandung unsur spam, promosi, emoji, atau tautan.
4. Analisis dilakukan pada level teks komentar tanpa memperhitungkan faktor eksternal lain seperti profil pengguna, atau metadata lainnya.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Menganalisis kinerja integrasi model BERT dan *Named Entity Recognition* (NER) dalam melakukan klasifikasi emosi multilabel pada komentar publik di kanal YouTube Sekretariat Presiden Indonesia.
2. Membuat sistem analisis emosi multilabel berbasis integrasi BERT dan *Named Entity Recognition* (NER) yang dapat mengidentifikasi lima kategori emosi utama (*love*, *happiness*, *anger*, *fear*, dan *sadness*) pada komentar publik YouTube terkait konferensi pers Presiden Republik Indonesia.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Penelitian ini dapat memberikan pemahaman yang lebih baik kepada masyarakat dan peneliti bahwa teknologi Natural Language Processing (NLP), khususnya model BERT yang diintegrasikan dengan Named Entity Recognition (NER), dapat digunakan untuk mengenali dan menganalisis berbagai bentuk emosi publik yang muncul dalam komentar YouTube Presiden Republik Indonesia.
2. Dengan menerapkan pendekatan multilabel emotion classification yang berfokus pada lima kategori emosi utama (love, happiness, anger, fear, dan sadness), serta mengombinasikan kemampuan kontekstual BERT dengan ekstraksi entitas dari NER, penelitian ini diharapkan dapat menghasilkan model dengan performa terbaik dalam memahami konteks emosional komentar publik di media sosial.
3. Hasil penelitian ini diharapkan dapat menjadi referensi dalam pengembangan model analisis emosi berbasis deep learning yang lebih akurat untuk bahasa Indonesia, serta menjadi dasar bagi penelitian lanjutan dalam bidang analisis sentimen publik, opini masyarakat, dan sistem pemantauan emosi otomatis terhadap pernyataan pemerintah.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun dalam lima bab utama, yang secara sistematis menggambarkan alur penelitian dari pendahuluan hingga kesimpulan. Adapun sistematika penulisannya adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi gambaran umum mengenai penelitian yang mencakup latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, serta sistematika penulisan skripsi. Bab ini memberikan landasan awal untuk memahami konteks dan arah penelitian yang dilakukan.

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan kajian literatur dan penelitian terdahulu yang relevan untuk memperkuat dasar teoritis penelitian. Di dalamnya mencakup hasil-hasil penelitian sebelumnya, teori-teori pendukung,

serta konsep-konsep utama yang menjadi dasar dalam perumusan metodologi penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan secara rinci tahapan-tahapan yang dilakukan untuk mencapai tujuan penelitian. Isinya mencakup rancangan penelitian, teknik pengumpulan data, proses *pre-processing* data, pemodelan menggunakan BERT, teknik klasifikasi multilabel, serta metode evaluasi kinerja model. Bab ini juga menjelaskan alat, bahan, dan lingkungan penelitian yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian yang diperoleh dari proses pengolahan data dan pembangunan model klasifikasi emosi. Selain itu, dibahas pula hasil evaluasi kinerja model dan interpretasi temuan penelitian, termasuk analisis terhadap ekspresi emosi publik yang teridentifikasi dari komentar video konferensi pers Presiden.

BAB V PENUTUP

Bab ini memuat kesimpulan penelitian yang merangkum hasil temuan secara keseluruhan dan memberikan jawaban atas rumusan masalah. Selain itu, bab ini juga memuat saran yang dapat menjadi dasar untuk pengembangan penelitian di masa mendatang.

DAFTAR PUSTAKA

Berisi seluruh referensi dan sumber literatur yang digunakan dalam penelitian, disusun sesuai dengan format penulisan yang berlaku.

LAMPIRAN

Lampiran memuat dokumen pendukung penelitian seperti data mentah, kode pemrograman, hasil uji model, dan dokumentasi lain yang relevan dengan proses penelitian.