BAB II
LANDASAN TEORI

2.1 Tinjauan Pustaka

Natural Language Processing (NLP), atau pengolahan bahasa alami, telah mengalami perkembangan pesat dalam beberapa dekade terakhir, terutama dalam bidang analisis sentimen. Berbagai metode *machine learning*, mulai dari yang sederhana seperti Naïve Bayes dan SVM, hingga yang lebih kompleks seperti *deep learning* dengan *Recurrent Neural Networks* (RNN) dan *Convolutional Neural Network* (CNN), telah banyak diadopsi dalam analisis sentimen. Selain itu, metode terbaru seperti *transformer networks*, yang meliputi model BERT dan *Generative pre-trained transformer* (GPT), telah menunjukkan peningkatan performa yang substansial dalam tugas-tugas NLP, termasuk analisis sentimen, karena kemampuannya untuk memahami konteks secara lebih mendalam (Min, dkk, 2021).

Penggunaan metode pengolahan bahasa alami untuk analisis sentimen juga dapat digunakan dalam survei kepuasan pengguna, salah satunya pada sebuah kelas kursus (*workshop*) yang dilakukan dari pengajar kepada muridnya (Pinto, 2023). Data yang digunakan adalah data komentar dari pengguna dan saran dari guru. Pada penelitian ini, model BERT digunakan dengan tujuan klasifikasi emosi ke dalam beberapa klasifikasi seperti puas, tidak puas dan sangat puas. Hasil dari penggunaan model BERT ini kemudian dikomparasikan dengan penelitian terdahulu dengan data yang sama tetapi menggunakan algoritma yang berbeda seperti *Lexicon with Multilayer Perceptron, SVM with Information Gain, Lexicon with Multilayer Perceptron,* dan *Aspect Based Sentiment Analysis.* Dari hasil perbandingan tersebut, penggunaan model BERT unggul dari yang lain dengan perolehan akurasi sebesar 94%. Namun pada penelitian ini tidak diketahui nilai yang lain dalam pengukuran kinerja model seperti F1-*score* ataupun *precision* jika dibandingkan dengan model yang lain. Selain itu, pada hasil penelitian ini hanya menggunakan klasifikasi biner (positif dan negatif). Padahal dalam banyak kasus di dunia nyata, terdapat sejumlah besar opini yang bersifat netral atau informatif. Memaksa data netral untuk masuk ke dalam kelas positif atau negatif dapat

menyebabkan ambiguitas dan berpotensi menurunkan performa model. Oleh karena itu, penelitian ini mengusulkan model klasifikasi tiga kelas (positif, negatif, netral) untuk menciptakan pemetaan sentimen yang lebih presisi dan meningkatkan keandalan sistem secara keseluruhan.

Penelitian tentang analisis sentimen dengan metode yang sama yaitu menggunakan BERT juga telah digunakan pada data ulasan film IMBD dan ulasan Amazon *Fine Food* (Selvakumar dan Lakshmanan, 2022). Pada penelitian ini, negasi kata dan inklusi dijadikan turunan pada proses analisis sentimen. Pada proses pelatihan data dilakukan variasi pada pembagian proses pelatihan data dan uji data mulai dari 50:50, 60:40, dan 70:30. Pada variasi 70:30 menghasilkan hasil yang lebih baik. Dibandingkan dengan model lain seperti *machine learning* maupun *deep learning*, BERT memperoleh nilai akurasi yang paling baik sebesar 81%. Namun, pembuktian pada penelitian tersebut hanya dilakukan terhadap model-model generasi sebelumnya, sehingga menjadikan pertanyaan mengenai bagaimana performa BERT jika dihadapkan dengan arsitektur *deep learning* yang berbasis *transformer* lain seperti RoBERTa atau GPT. Sedangkan yang akan dilakukan pada penelitian ini akan membandingan dengan variasi model BERT lain yaitu RoBERTa.

Pada penelitian selanjutnya dengan menggunakan model *transformer* seperti BERT, analisis sentimen berbasis aspek juga dapat menggunakan model *transformer* lain seperti GPT untuk dataset dari berbagai ulasan produk atau layanan yang mengandung informasi tentang aspek tertentu, seperti kualitas produk, layanan pelanggan, atau harga (Chumakov, dkk, 2023). Metode ini dapat digunakan memanfaatkan kekuatan model GPT dalam memahami konteks dengan tujuan mengidentifikasi aspek-aspek tertentu dalam teks dan menentukan sentimen yang terkait dengan masing-masing aspek. Penggunaan model GPT, khususnya OpenAI GPT-3 text-davinci-003, sangat efektif dalam tugas *Aspect Sentiment Triplet Extraction* (ASTE), terutama dalam mode *few-shot* pada teks berbahasa Rusia. Model ini unggul dalam ekstraksi triplet dengan struktur output yang jelas dan sederhana. GPT-2 juga menunjukkan performa yang baik, terutama setelah *fine-tuning*, meskipun menggunakan lebih sedikit parameter. Namun pada

penelitian ini evaluasi hanya ditunjukan pada nilai F-*score*, sehingga tidak memberikan gambaran evaluasi yang lebih luas jika disandingkan parameter evaluasi yang lain. Selain itu, pada penelitian ini hanya menggunakan satu model generatif untuk menangani seluruh tugas ASTE. Sedangkan pendekatan yang akan dilakukan pada penelitian lebih modular dengan melakukan ekstraksi aspek terlebih dahulu menggunakan *rule-based* dan BERTopic untuk selanjutnya dilakukan klasifikasi sentimen.

Metode analisis sentimen berbasis aspek dengan leksikon juga dapat digunakan dalam menganalisis data ulasan aplikasi pintar (*smart app)* milik pemerintah United Emirate Arab (UEA) (Alqaryouti, dkk, 2024). Penelitian ini juga cukup relevan jika disandingkan dengan penelitian yang akan dilakukan di proposal ini karena memiliki objek yang sama yaitu pada layanan pemerintahan. Tujuan utama dari penelitian ini adalah membantu entitas pemerintah mendapatkan wawasan tentang kebutuhan dan harapan pelanggan mereka melalui analisis ulasan aplikasi pintar. Penelitian ini menggunakan pendekatan analisis sentimen berbasis aspek yang menggabungkan leksikon khusus domain dan aturan untuk mengekstrak aspek penting dari ulasan dan mengklasifikasikan sentimen yang terkait dengan aspek tersebut. Penelitian ini menggunakan pendekatan hibrida yang menggabungkan leksikon domain dan aturan untuk menganalisis ulasan aplikasi pemerintahan pintar. Model yang diusulkan berhasil menangani berbagai tantangan dalam analisis sentimen, termasuk aspek eksplisit dan implisit. Namun, metode semacam ini memiliki keterbatasan dalam memahami konteks kalimat yang kompleks dan nuansa bahasa yang tidak terdefinisi dalam aturan. Pada komparasi evaluasi juga hanya disandingkan dengan *machine-learning* klasing yaitu SVM. Oleh karena itu, penelitian ini bertujuan untuk menjembatani celah tersebut dengan mengintegrasikan model *transformer* BERT yang memiliki kemampuan pemahaman kontekstual yang superior, serta BERTopic untuk ekstraksi aspek yang lebih dinamis.

Model algoritma *deep learning* seperti BERT semakin berkembang menyesuaikan bahasa yang digunakan pada proses analisis sentimen contohnya adalah IndoBERT. Model BERT ini digunakan untuk melakukan analisis sentimen

pada teks bahasa Indonesia yang ada di sosial media (Nabiilah, dkk., 2023). Model ini mengklasifikasikan sentimen lebih spesifik di media sosial yaitu dengan mengklasifikasikan teks yang mengandung komen negatif yang diklasifikasikan ke dalam beberapa kelas. Pada nilai F1-*score* model IndoBERT menghasilkan nilai yang paling unggul jika dikomparasikan dengan model BERT lain seperti Multilingual BERT dan IndoRoBERTa. Walaupun performa yang dihasilkan menunjukan indikator yang cukup baik, namun pada penelitian tersebut data yang digunakan dalam proses pelatihan model dilakukan pelabelan menggunakan *lexicon-based*. Hal ini dapat dikatakan kontradiksi dalam metodologi, dikarenakan model BERT dirancang untuk memahami konteks lebih mendalam, sedangkan jika data latihnya menggunakan leksikon, walaupun performanya baik tetapi bisa gagal dalam memahami konteks pada kata/kalimatnya. Dalam peneilitian yang akan dilakukan pada penelitian ini akan disandingkan metode *lexicon-based* murni dengan model BERT dalam memahami konteks kata.

Pada penelitian lain model BERT seperti DeBERTa dikombinasikan dengan *Gated Recurrent Unit* (GRU) untuk melakukan analisis sentimen (Assiri, dkk, 2024). Proses analisis sentimen dengan kombinasi model DeBERTa dan GRU dilakukan pada dataset twitter dengan jumlah 9640 *tweet.* Pada penelitian ini juga menggunakan leksikon dalam memberikan skor sentimen. Pada proses pelatihan model dilakukan dengan konfigurasi epoch 30 dan pembagian dataset pada perbandingan 6:2:2 meliputi data latih, data uji dan data validasi. Hasil dari DeBERTa dan GRU dengan konfigurasi tersebut menghasilkan akurasi tertinggi yaitu 96.87%. Namun dari hasil ini belum diketahui jika menggunakan variasi pada konfigurasi saat pelatihan data. Selain itu, sama halnya dengan penelitan terkait sebelumnya, dalam proses pelabelan data pada penelitian tersebut menggunakan metode leksikon yang minim dalam memahami konteks dalam analisis sentimen.

Analisis sentimen menggunakan model IndoBERT juga digunakan dalam menganalisis sentimen kebijakan keanekaragaman hayati menggunakan data twitter (Uliniansyah, dkk, 2024). Dataset yang digunakan sejumlah 500.000 data *tweet* berbahasa Indonesia. Penelitian ini menggunakan konfigurasi fungsi aktivasi yang berbeda untuk lapisan terhubung pertama yaitu tanpa aktivasi,

Gaussian Error Linear Unit (GELU), dan norm hiperbolik tangen (norm tanh). Dari ketiga konfigurasi ini, fungsi aktivasi GELU menunjukkan hasil yang paling baik dengan akurasi sebesar 0.8263 dan F1-*score* 0.7959. Pada sekenario kedua yaitu dengan menggunakan *logistic regression* dan *Support Vector Classifier* (SVC) hanya memperoleh nilai akurasi 0.72115 dan F1-*score* 0.71967. Namun dalam penelitian tersebut, penentuan topik hanya sebatas menggunakan aturan kata (*rule-based*). Motode ini memiliki kelemahan dalam menemukan topik-topik lain yang lebih bervariasi dan mungkin penting. Sedangkan penelitian yang akan dilakukan pada penelitian ini akan dikomparasikan dengan BERTopic, model yang dapat menemukan topik-topik lebih luas dan komperhensif.

Beberapa penelitian sebelumnya juga telah memanfaatkan model IndoBERT maupun IndoRoBERTa pada berbagai domain. Penelitian dengan data ulasan hotel menunjukkan bahwa meskipun data sudah dikategorikan berdasarkan aspek, tidak semua aspek berhasil diekstraksi sehingga hasil analisis masih terbatas (Yulianti dan Nissa, 2024). Pada ulasan pelanggan Agoda, IndoBERT mampu menghasilkan akurasi tinggi sebesar 92,52%, namun model tersebut gagal menangani kelas minoritas yang terlihat dari nilai F1-Score untuk kelas netral sebesar 0,00 dan hanya 28,57% pada kelas negatif (Singgalen, 2025). Sementara itu, penelitian dengan komentar YouTube menghadapi permasalahan ketidakseimbangan data yang berdampak pada rendahnya performa model, di mana akurasi dan F1-Score masih berada di bawah 70% meskipun telah dilakukan pembuangan data (Widarmanti, dkk, 2022). Penelitian lain dengan data artikel berita menunjukkan bahwa gaya bahasa formal pada berita mampu memberikan pemahaman konteks yang lebih baik, namun cakupan analisis yang dilakukan masih terbatas pada struktur dokumen tanpa mengeksplorasi aspek spesifik dari teks (Rihadatul'Aisy dan Pamungkas, 2025).

Penelitian terkait yang sudah dijelaskan pada paragraf-paragraf sebelumnya dapat dilihat pada Tabel 2.1.

Tabel 2.1 Penelitian terkait

No.	Peneliti	Metode	Domain dan dataset	Keterbatasan dan celah penelitian
1.	(Pinto, dkk, 2023)	BERT	25.000 komentar pengguna dan 506 saran guru dari kursus workshop	Penelitian ini hanya menggunakan klasifikasi biner (positif/negatif).
2.	(Selvakumar dan Lakshmanan, 2022)	BERT	50.000 ulasan IMBD dan 125.060 ulasan Amazon *fine food*	Komparasi model yang dilakukan pada penelitian ini hanya sebatas dengan model *machine learning*.
3.	(Chumakov, dkk., 2023)	*Aspect Sentiment Triplet Extraction* (ASTE), GPT	Dataset bahasa Rusia dan Inggris	Penelitian ini hanya menggunakan satu model generatif untuk seluruh tugas ASTE.
4.	(Alqaryouti, dkk., 2024)	*Lexicon based* dan *rule based*	Ulasan aplikasi pintar (*smarta pp)* pemerintah UEA	Penelitian ini menggunakan metode *lexicon based* dan *rule-based* kurang maksimal dalam memahami konteks kata/kalimat.
5.	(Nabiilah, dkk., 2022)	Multilingual BERT, IndoBERT dan IndoRoBERTa	Sosial media (Instagram, Twitter dan Kaskus)	Penelitian ini menggunakan dataset yang sudah dilabeli dalam menentukan topik (*supervised*)
6.	(Assiri, A. dkk., 2024)	BERT, DeBERTa dan *Gated Recurrent*	9.640 Data *tweet*	Penelitian ini menggunakan metode *Valence Aware Dictionary and Sentiment Reasoner (VADER)*,

Tabel 2.2 Penelitian terkait (lanjutan)

No.	Peneliti	Metode	Domain dan dataset	Keterbatasan dan celah penelitian
		Unit (GRU)		metode ini sama halnya seperti leksikon yang kurang maksimal dalam memahami konteks kata/kalimat.
7.	(Uliniansyah, , M. T., dkk., 2024)	IndoBERT, *rule-based*	15.000 Data *tweet*	Penelitian ini melakukan klasifikasi sentimen pada level keseluruhan data, namun dalam menentukan aspek hanya menggunakan aturan kata (*rule-based*)
8.	(Yulianti dan Nissa, 2024)	IndoBERT	2.854 Data ulasan hotel	Penelitian ini menggunakan data yang sudah dikategorikan aspeknya, sehingga tidak tidak semua aspek terekstraksi.
9.	(Singgalen, 2025)	IndoBERT	5.796 Ulasan pelanggan Agoda	Pada penelitian ini, IndoBERT menghasilkan akurasi yang tinggi sebesar 92,52%, tetapi model gagal menangani kelas minoritas yang dapat dilihat pada nilai F1-*Score* untuk kelas netral adalah 0.00 dan untuk kelas negatif hanya 28.57%
10.	(Widarmanti, dkk, 2022)	IndoRoBERTa	9.000 Komentar YouTube	Pada penelitian ini dilakukan pembuangan data yang tidak seimbang, padahal data masih bisa diekstraksi untuk mendapatkan informasi. Metrik F1-*score* maupun akurasi juga masih di bawah 70%.

Tabel 2.3 Penelitian terkait (lanjutan)

No.	Peneliti	Metode	Domain dan dataset	Keterbatasan dan celah penelitian
11.	(Rihaadatul' Aisy dan Pamungkas, 2025)	IndoBERT dan IndoRoBERTa	3.538 Artikel berita	Penggunaan data pada artikel berita memiliki gaya bahasa yang formal, berbeda dengan data teks ulasan atau komentar yang memilliki pemahaman konteks yang cukup kompleks. Pada penelitian ini juga melakukan klasifikasi pada level dokumen dalam keseluruhan artikel

2.2 Dasar Teori

Dasar teori pada penelitian ini akan menjelaskan beberapa teori dan konsep utama yang menjadi pijakan dalam melakukan analisis sentimen berbasis aspek menggunakan model BERT dalam konteks aplikasi E-SKM Jawa Tengah.

2.2.1 Analisis Sentimen

Analisis sentimen adalah cabang dari pemrosesan bahasa alami (NLP) yang bertujuan untuk menentukan sikap atau opini penulis terhadap suatu subjek. Dalam konteks ini, analisis sentimen mengklasifikasikan teks menjadi kategori sentimen seperti positif, negatif, atau netral (Prager, 2006). Ada beberapa pendekatan yang digunakan dalam analisis sentimen, termasuk pendekatan berbasis kamus, pendekatan berbasis pembelajaran mesin, dan pendekatan berbasis *deep learning*. Pendekatan berbasis kamus menggunakan daftar kata yang telah diberi label sentimen positif atau negatif untuk menentukan sentimen dari sebuah teks. Pendekatan berbasis pembelajaran mesin melibatkan pelatihan model pada data latih yang telah diberi label sentimen, menggunakan algoritma seperti SVM, atau model *deep learning* seperti *Recurrent Neural Networks* (RNN) dan *Convolutional Neural Networks* (CNN) (Kharde dan Sonawane, 2016). Pendekatan berbasis *deep learning*, termasuk penggunaan model seperti *Long Short-Term Memory* (LSTM) dan *transformer* seperti BERT, mampu menangkap

konteks dan makna yang lebih kompleks dari teks (Minaee, dkk, 2021). Namun, analisis sentimen menghadapi beberapa tantangan signifikan. Ambiguitas dan polisemi kata, di mana kata yang sama dapat memiliki makna yang berbeda tergantung pada konteksnya, merupakan salah satu tantangan utama. Selain itu, memahami sentimen dari ekspresi kiasan, sarkasme, atau ironi sering kali sulit karena makna yang dimaksudkan sering kali berbeda dari kata-kata yang digunakan. Analisis sentimen juga memerlukan pemahaman konteks domain spesifik, karena sentimen terhadap kata yang sama dapat berbeda tergantung pada domainnya.

Analisis sentimen memiliki berbagai aplikasi praktis, termasuk menilai opini pelanggan terhadap produk atau layanan, mengukur sentimen masyarakat terhadap isu-isu tertentu atau tokoh publik melalui media sosial, dan memahami persepsi publik terhadap kampanye pemasaran atau perubahan produk. Teknik ini membantu perusahaan dan organisasi dalam membuat keputusan yang lebih baik berdasarkan data opini dan sentimen yang dikumpulkan (Nanli, dkk, 2012).

2.2.2 Analisis Sentimen Berbasis Aspek

Analisis Sentimen Berbasis Aspek (*Aspect-Based Sentiment Analysis* atau ABSA) adalah teknik dalam pemrosesan bahasa alami (*Natural Language Processing* atau NLP) yang bertujuan untuk mengidentifikasi sentimen terhadap aspek-aspek spesifik dari entitas yang disebutkan dalam teks. Teknik ini memungkinkan analisis yang lebih mendetail dan kontekstual mengenai pandangan atau opini terhadap berbagai fitur atau elemen dari suatu produk, layanan, atau subjek tertentu. ABSA sering digunakan dalam ulasan produk, media sosial, dan layanan pelanggan untuk memberikan wawasan yang lebih granular mengenai kepuasan atau ketidakpuasan pengguna terhadap aspek-aspek tertentu (Madhoushi, dkk, 2019).

Salah satu pendekatan awal dalam ABSA adalah dengan mengidentifikasi kata-kata yang terkait dengan aspek tertentu menggunakan metode berbasis aturan atau leksikon. Pendekatan teknik ABSA digunakan untuk mengekstrak aspek-aspek dari ulasan produk dengan mencari kata benda dan frasa kata benda, serta menentukan sentimen terkait dengan kata-kata ini melalui analisis kata sifat yang

sering muncul bersamaan (Amplayo, dkk, 2022). Pendekatan ini memberikan dasar bagi pengembangan model yang lebih canggih dalam ABSA, termasuk penggunaan teknik pembelajaran mesin dan jaringan saraf.

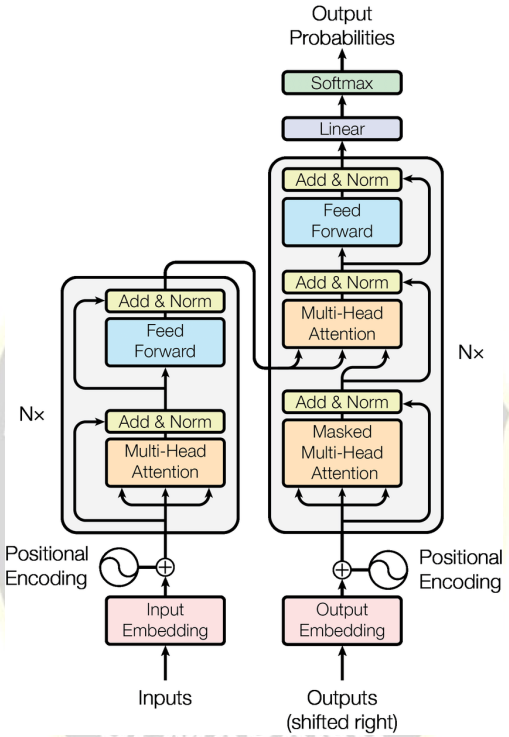
Perkembangan lebih lanjut dalam ABSA melibatkan penggunaan teknik pembelajaran mesin, seperti *Support Vector Machines* (SVM) dan Naive Bayes, untuk mengklasifikasikan sentimen terhadap aspek tertentu (Do, dkk, 2019). Teknik ini menggunakan fitur-fitur seperti kata-kata, frasa, dan n-gram untuk membangun model yang dapat memprediksi sentimen berdasarkan teks yang diberikan. Lebih baru lagi, pendekatan berbasis *deep learning* telah menunjukkan kinerja yang lebih baik dalam ABSA dengan menggunakan model seperti *Recurrent Neural Networks* (RNN) dan *Long Short-Term Memory* (LSTM) yang mampu menangkap hubungan jangka panjang dalam teks (Tang, dkk, 2016).

Sebagai salah satu bidang yang terus berkembang dalam NLP, ABSA memainkan peran penting dalam membantu organisasi dan individu untuk lebih memahami opini publik terhadap berbagai aspek produk atau layanan mereka. Penelitian terbaru menunjukkan bahwa integrasi model-model *transformer* seperti BERT dan RoBERTa dalam ABSA dapat memberikan hasil yang lebih akurat dan efisien dalam menangkap nuansa sentimen yang kompleks (Sun, dkk, 2019). Dengan demikian, ABSA terus menjadi alat yang esensial dalam analisis sentimen yang mendetail dan berbasis konteks.

2.2.3 BERT (*Bidirectional Encoder Representations from Transformers*)

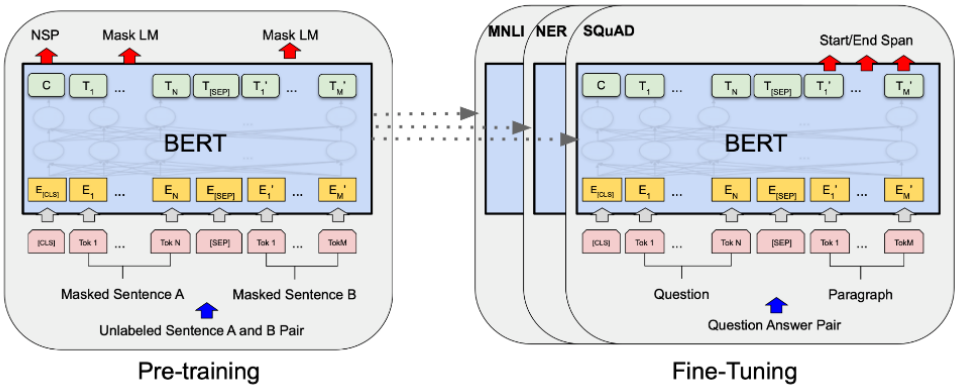
BERT (*Bidirectional Encoder Representations from Transformers*) adalah model pembelajaran mesin yang diperkenalkan oleh Google pada tahun 2018. Algoritma ini mendefinisikan ulang cara kerja pemrosesan bahasa alami (NLP) dengan memperkenalkan metode yang memanfaatkan konteks dua arah dari teks yang dipelajari. Sebelum BERT, model NLP umumnya bersifat satu arah (*unidirectional*), yang berarti mereka hanya memahami konteks dari kiri ke kanan atau sebaliknya. BERT mengatasi keterbatasan ini dengan memanfaatkan pendekatan dua arah penuh (*bidirectional*), yang memungkinkan pemahaman kontekstual yang lebih baik (Devlin, dkk, 2019).

Model BERT didasarkan pada sebuah arsitektur *transformer* seperti pada Gambar 2.1. *Transformer* adalah model yang menggunakan mekanisme perhatian (*attention mechanism*) untuk menangkap hubungan antara kata-kata dalam sebuah kalimat, terlepas dari jarak antara kata-kata tersebut. BERT terdiri dari lapisan *encoder* dari *transformer*, yang secara simultan memproses semua kata dalam sebuah kalimat dan mempertimbangkan hubungan di antara kata-kata tersebut.



Gambar 2.1 Arsitektur *transformer* (Vaswani, dkk, 2017)

BERT dilatih dengan dua tugas utama, yaitu *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP). Pada MLM, sebagian kata dalam kalimat disembunyikan (masked), lalu model dilatih untuk memprediksi kata yang hilang berdasarkan konteks sekitarnya. Sementara itu, NSP digunakan untuk melatih model dalam memahami hubungan antar kalimat, yaitu dengan menentukan apakah suatu kalimat benar-benar mengikuti kalimat sebelumnya. Kombinasi kedua teknik ini membuat BERT unggul dalam memahami makna kata dalam konteks yang lebih luas serta relasi antar kalimat. Arsitektur BERT dapat dilihat pada Gambar 2.2.



Gambar 2.2 Arsitektur BERT (Devlin, dkk, 2019)

2.2.4 IndoBERT

Perkembangan pada model BERT dalam analisis sentimen juga menyesuaikan bahasa yang digunakan, salah satunya yaitu IndoBERT. Pengembangan IndoBERT dilakukan dengan menyesuaikan model BERT asli untuk bahasa Indonesia. IndoBERT dilatih menggunakan korpus teks berbahasa Indonesia yang besar, yang mencakup berbagai sumber teks seperti berita, media sosial, dan dokumen resmi. Proses tokenisasi WordPiece dalam IndoBERT juga disesuaikan agar sesuai dengan karakteristik bahasa Indonesia. Hal ini memungkinkan IndoBERT untuk memahami dan memproses teks berbahasa Indonesia dengan lebih baik (Wilie dkk, 2020).

Perkembangan model IndoBERT secara signifikan meningkatkan kinerja pada berbagai tugas NLP untuk bahasa Indonesia dibandingkan dengan model sebelumnya. IndoBERT mencapai hasil terbaik dalam beberapa *benchmark* NLP bahasa Indonesia, selain itu IndoBERT menunjukkan kemampuan lebih baik dalam menangkap konteks dan makna dalam teks berbahasa Indonesia. Keberhasilan ini menunjukkan potensi besar IndoBERT dalam mengatasi berbagai tantangan NLP untuk bahasa Indonesia (Koto, dkk, 2020).

IndoBERT telah digunakan dalam berbagai aplikasi, termasuk analisis sentimen, *named entity recognition* (NER), penerjemahan mesin, dan pemahaman

teks. Dalam analisis sentimen, IndoBERT mampu mengidentifikasi sentimen positif, negatif, atau netral dari teks dengan akurasi tinggi. Untuk pengenalan entitas bernama, IndoBERT dapat mengidentifikasi entitas seperti nama orang, organisasi, dan lokasi dalam teks secara efektif. Selain itu, IndoBERT juga berperan dalam membantu penerjemahan teks dari dan ke bahasa Indonesia serta menjawab pertanyaan berdasarkan teks yang diberikan.

2.2.5 IndoRoBERTa

IndoRoBERTa merupakan model bahasa berbasis arsitektur RoBERTa (*Robustly Optimized BERT Pretraining Approach*) yang telah diadaptasi secara khusus untuk bahasa Indonesia. RoBERTa sendiri merupakan pengembangan lebih lanjut dari model BERT dengan beberapa modifikasi penting, seperti penghapusan tugas *Next Sentence Prediction* (NSP), penggunaan jumlah data pelatihan yang lebih besar, serta waktu pelatihan yang lebih lama untuk meningkatkan kinerja model (Liu, dkk, 2019). IndoRoBERTa dikembangkan oleh tim IndoNLU dengan memanfaatkan kumpulan data berbahasa Indonesia yang sangat luas dan beragam, mencakup teks dari berita daring, media sosial, forum diskusi, hingga dokumen publik (Wilie, dkk, 2020).

Pelatihan yang dilakukan secara mendalam pada data yang relevan dengan konteks lokal memungkinkan IndoRoBERTa untuk memahami struktur dan makna kalimat dalam bahasa Indonesia secara lebih akurat. Dalam berbagai studi empiris, model ini menunjukkan kinerja yang unggul pada berbagai tugas pemrosesan bahasa alami, seperti klasifikasi sentimen, pengenalan entitas bernama (*Named Entity Recognition*), dan analisis topik, jika dibandingkan dengan model berbasis embedding statis atau model BERT yang tidak dilatih secara khusus pada bahasa Indonesia (Wilie, dkk, 2020). Dengan demikian, IndoRoBERTa menjadi salah satu model yang sangat relevan untuk digunakan dalam penelitian pemrosesan bahasa alami berbahasa Indonesia, termasuk dalam konteks analisis sentimen berbasis aspek.

2.2.6 BERTopic

BERTopic merupakan teknik pemodelan topik modern yang memadukan embedding berbasis *transformer* dengan algoritma *clustering* untuk menghasilkan topik yang lebih kontekstual, *interpretable*, dan kohesif dibandingkan pendekatan klasik seperti LDA. Mekanisme utama BERTopic melibatkan pengubahan teks menjadi representasi embedding melalui model seperti BERT atau Sentence-BERT, reduksi dimensi menggunakan UMAP, *clustering* dengan HDBSCAN, serta ekstraksi kata kunci topik menggunakan class-based TF-IDF (Grootendorst, 2022).

Dalam studi evaluasi pemodelan topik pada data media sosial, BERTopic menunjukkan performa superior dalam menghasilkan klaster yang lebih terstruktur dan mudah diinterpretasikan dibandingkan dengan LDA dan NMF (Kaur dan Wallace, 2024). Penelitian lain pada dokumen catatan klinis juga menunjukkan bahwa BERTopic mampu mendeteksi topik-topik kesehatan dengan konsistensi temporal yang baik dan relevansi yang tinggi, meskipun sensitif terhadap jumlah klaster (Meaney, dkk, 2022).

Implementasi BERTopic dalam tugas teks berbahasa Indonesia telah berhasil diterapkan untuk analisis opini publik terhadap kebijakan keanekaragaman hayati, dengan hasil menunjukkan nilai koherensi yang tinggi setelah penyesuaian *preprocessing* pada karakteristik bahasa lokal seperti *hashtag* dan *mention* (Hidayati dan Shaleha, 2024). Selain itu, penerapannya dalam domain keuangan pada korpus berita menunjukkan bahwa BERTopic dapat menghasilkan topik yang koheren secara semantik dan memiliki efisiensi komputasi yang lebih baik dibandingkan dengan LDA maupun Top2Vec (Chen, dkk, 2023).

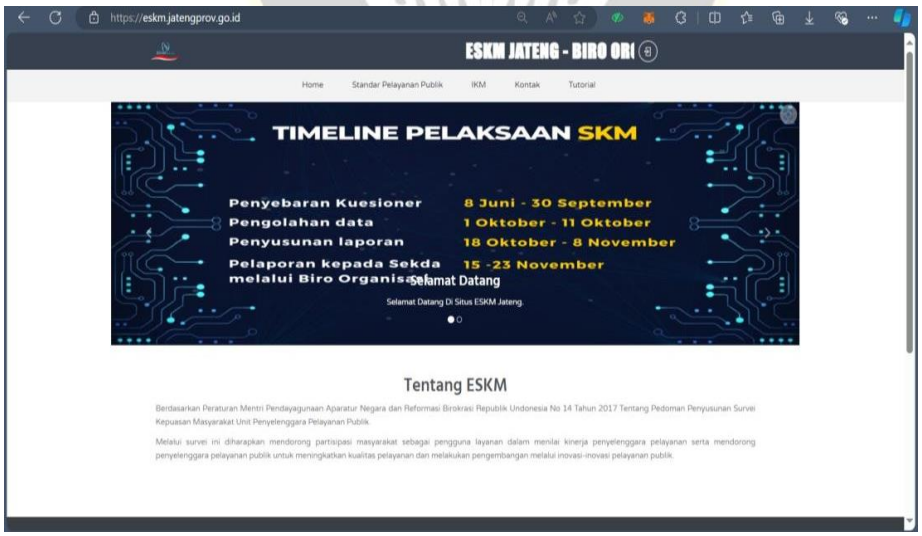
2.2.7 Survei Kepuasan Masyarakat (SKM)

Survei kepuasan masyarakat adalah instrumen penting yang digunakan oleh instansi pemerintah untuk mengukur tingkat kepuasan masyarakat terhadap layanan publik. Tujuannya adalah untuk mendapatkan gambaran mengenai kualitas pelayanan dari sudut pandang masyarakat, mengidentifikasi kelemahan dan kekuatan dalam pelayanan, serta memberikan masukan untuk perbaikan dan peningkatan kualitas layanan. Landasan hukum survei ini diatur dalam Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi (MenPANRB)

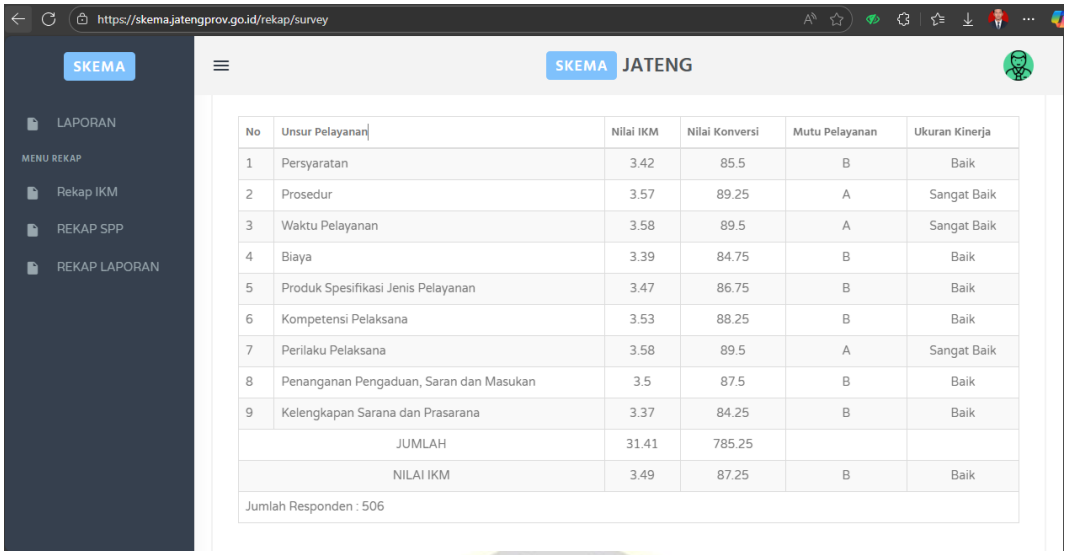
Republik Indonesia Nomor 16 Tahun 2014, yang memberikan pedoman tentang pelaksanaan survei oleh instansi pemerintah.

Metodologi survei mencakup penentuan populasi dan sampel secara acak untuk memastikan representativitas, penggunaan kuesioner sebagai instrumen survei, serta pengumpulan data melalui berbagai metode seperti wawancara langsung dan survei online. Data yang dikumpulkan kemudian diolah dan dianalisis untuk menghasilkan indeks kepuasan masyarakat. Beberapa indikator yang diukur dalam survei ini antara lain kesesuaian persyaratan, kemudahan prosedur, waktu pelayanan, kualitas sarana dan prasarana, kompetensi pelaksana, perilaku pelaksana, keadilan dalam pelayanan, dan kepuasan terhadap biaya.

Survei kepuasan masyarakat memberikan berbagai manfaat signifikan, termasuk peningkatan kualitas pelayanan, peningkatan kepercayaan publik terhadap pemerintah, dan perencanaan yang lebih baik berdasarkan data survei. Dengan dasar hukum yang kuat dan metodologi yang sistematis, survei ini dapat memberikan gambaran yang jelas mengenai persepsi masyarakat dan menjadi dasar untuk perbaikan layanan yang berkelanjutan. Survei kepuasan masyarakat di lingkungan Pemerintah Provinsi Jawa Tengah dilakukan melalui aplikasi E-SKM Jateng.



Gambar 2.3 Halaman utama aplikasi E SKM



No	Unsur Pelayanan	Nilai IKM	Nilai Konversi	Mutu Pelayanan	Ukuran Kinerja
1	Persyaratan	3.42	85.5	B	Baik
2	Prosedur	3.57	89.25	A	Sangat Baik
3	Waktu Pelayanan	3.58	89.5	A	Sangat Baik
4	Biaya	3.39	84.75	B	Baik
5	Produk Spesifikasi Jenis Pelayanan	3.47	86.75	B	Baik
6	Kompetensi Pelaksana	3.53	88.25	B	Baik
7	Perilaku Pelaksana	3.58	89.5	A	Sangat Baik
8	Penanganan Pengaduan, Saran dan Masukan	3.5	87.5	B	Baik
9	Kelengkapan Sarana dan Prasarana	3.37	84.25	B	Baik
JUMLAH		31.41	785.25		
NILAI IKM		3.49	87.25	B	Baik
Jumlah Responden : 506					

Gambar 2.4 Nilai sembilan unsur pertanyaan pada aplikasi E-SKM

Pada Gambar 2.3 merupakan tampilan umum aplikasi E-SKM yang bisa diakses secara umum oleh publik. Masyarakat umum dapat melihat nilai IKM setiap instansi pada menu IKM. Selanjutnya, rincian IKM yang memuat nilai dari sembilan unsur pertanyaan secara lebih lengkap juga dapat diakses oleh admin masing-masing instansi seperti pada Gambar 2.4.