BAB 2

TINJAUAN PUSTAKA

Penelitian mengenai NLP, khususnya *Sentiment Analysis*, ABSA, dan NER telah berkembang pesat dalam beberapa tahun terakhir. Berbagai penelitian telah memanfaatkan pendekatan *machine learning* maupun *deep learning* untuk mengekstraksi informasi dari data media sosial pada berbagai domain, seperti ulasan produk, pariwisata, layanan publik, hingga kebijakan pemerintah (Hua et al., 2024). Meskipun demikian, setiap penelitian memiliki karakteristik yang berbeda dari sisi domain, metode ekstraksi aspek, pembangunan dataset, maupun model yang digunakan. Oleh karena itu, kajian terhadap penelitian terdahulu diperlukan untuk mengidentifikasi perkembangan penelitian serta menemukan kesenjangan penelitian (*research gap*) yang menjadi dasar penelitian ini.

Pada domain kebijakan publik, sebagian besar penelitian masih menggunakan pendekatan *Sentiment Analysis* untuk mengklasifikasikan opini masyarakat ke dalam kategori positif, negatif, atau netral. Penelitian oleh Mumtaz et al. (2025) menunjukkan bahwa algoritma SVM mampu mengklasifikasikan opini masyarakat terhadap kebijakan efisiensi anggaran dengan performa yang baik, namun pendekatan tersebut hanya menghasilkan satu label sentimen untuk setiap *Tweet* sehingga belum mampu mengidentifikasi aspek kebijakan yang menjadi penyebab munculnya sentimen (Mumtaz et al., 2025). Temuan serupa juga ditunjukkan oleh penelitian Alghifari dan Azis (2025), Hafika et al. (2025), dan Hafika dan Waruwu (2025) yang masih berfokus pada klasifikasi sentimen secara umum tanpa menghubungkan sentimen dengan aspek kebijakan tertentu (Alghifari & Azis, 2025; Hafika, Hafiz, et al., 2025; Hafika, Waruwu, et al., 2025). Selain itu, Hadiansyah (2025) memanfaatkan opini masyarakat pada media sosial sebagai salah satu indikator evaluasi implementasi kebijakan, namun proses analisis masih dilakukan secara manual sehingga kurang efisien untuk data dalam jumlah besar (Hadiansyah, 2025).

Seiring berkembangnya penelitian di bidang NLP, pendekatan ABSA mulai banyak digunakan untuk mengatasi keterbatasan *Sentiment Analysis* konvensional. Berbeda dengan analisis sentimen tingkat dokumen atau kalimat, ABSA mampu menghubungkan sentimen dengan aspek tertentu sehingga menghasilkan informasi yang lebih rinci. Penelitian Putra et al. (2024) menunjukkan bahwa penerapan ABSA pada ulasan pariwisata mampu memberikan informasi yang lebih komprehensif mengenai aspek yang dibahas dibandingkan analisis sentimen konvensional (Putra et al., 2024). Sementara itu, penelitian Sun et al. (2019) menunjukkan bahwa model berbasis Transformer mampu meningkatkan performa ABSA melalui representasi kontekstual yang lebih baik (Sun et al., 2019), sedangkan Devlin et al. (2019) memperkenalkan arsitektur BERT yang menjadi dasar perkembangan berbagai model NLP modern (Devlin et al., 2019). Meskipun demikian, berbagai kajian literatur menunjukkan bahwa penerapan ABSA pada domain kebijakan publik masih relatif terbatas dibandingkan domain komersial seperti ulasan produk maupun layanan (Harwenda et al., 2025).

Perkembangan penelitian selanjutnya mengarah pada penggunaan NER sebagai pendekatan *sequence labeling* untuk melakukan ekstraksi aspek pada tingkat token. Pendekatan ini memungkinkan model tidak hanya menentukan polaritas sentimen, tetapi juga mengenali batas (*boundary*) setiap aspek dalam suatu opini. Penelitian *Detecting Emotion on Indonesian Online Chat Text Using Text Sequential Labeling* menunjukkan bahwa pendekatan *sequence labeling* mampu mengekstraksi informasi secara lebih rinci dibandingkan klasifikasi tingkat kalimat karena proses pelabelan dilakukan pada setiap token (Hasudungan & Kodhra, 2022) Namun demikian, penelitian tersebut masih berfokus pada deteksi emosi, sehingga belum diterapkan untuk ekstraksi aspek pada domain kebijakan publik. Selain itu, penelitian mengenai BiLSTM-CRF pada teks tidak terstruktur juga menunjukkan bahwa pendekatan *sequence labeling* efektif digunakan pada data yang memiliki karakteristik bahasa tidak baku dan struktur kalimat yang beragam, seperti media sosial (Rahmat et al., 2026).

Selain metode ekstraksi aspek, berbagai penelitian juga menunjukkan bahwa kualitas dataset merupakan salah satu faktor yang sangat menentukan keberhasilan pengembangan model ABSA maupun NER. Penelitian AnnoABSA menekankan pentingnya penyusunan *Annotation Guideline*, proses anotasi yang sistematis, serta evaluasi kualitas anotasi agar menghasilkan dataset yang konsisten (Hellwig et al., 2026). Kajian mengenai *low-resource languages* juga menunjukkan bahwa keterbatasan dataset beranotasi masih menjadi tantangan utama dalam pengembangan model NLP berbahasa Indonesia (Pha et al., 2026). Di sisi lain, perkembangan *deep learning* telah menghasilkan berbagai arsitektur seperti SimpleRNN, GRU, BiLSTM, dan Transformer yang banyak digunakan pada tugas NER. Penelitian mengenai BiLSTM-CRF menunjukkan bahwa peningkatan ukuran dataset dapat meningkatkan kemampuan model dalam mempelajari pola entitas, meskipun diikuti oleh peningkatan kebutuhan komputasi. Penelitian lain juga menunjukkan bahwa keberhasilan model tidak hanya dipengaruhi oleh arsitektur yang digunakan, tetapi juga oleh kualitas anotasi, karakteristik domain, dan konsistensi dataset yang digunakan sebagai data pelatihan.

Berdasarkan berbagai penelitian terdahulu tersebut dapat disimpulkan bahwa perkembangan penelitian bergerak dari *Sentiment Analysis* menuju ABSA dengan pendekatan NER untuk menghasilkan informasi yang lebih rinci mengenai aspek dan sentimen dalam suatu opini. Meskipun demikian, masih terdapat beberapa kesenjangan penelitian, yaitu (1) penerapan ABSA berbasis NER pada domain kebijakan publik berbahasa Indonesia masih relatif terbatas, (2) penelitian yang membangun dataset melalui penyusunan *Annotation Guideline*, validasi pakar, dan pengukuran *Inter-Rater Reliability* masih belum banyak dilakukan, (3) evaluasi beberapa arsitektur *deep learning* menggunakan dataset yang dibangun melalui prosedur anotasi yang sama masih terbatas, serta (4) pengaruh penambahan ukuran dataset terhadap kemampuan generalisasi model pada domain kebijakan publik juga belum banyak dilaporkan. Oleh karena itu, penelitian ini membangun dataset ABSA berbasis NER melalui proses anotasi dan validasi yang sistematis, kemudian mengevaluasi empat arsitektur *deep learning*, yaitu SimpleRNN-CRF,

BiGRU-CRF, BiLSTM-CRF, dan Transformer Encoder-CRF, serta menganalisis pengaruh penambahan ukuran dataset terhadap performa model.

Untuk memperjelas posisi penelitian ini dibandingkan dengan penelitian-penelitian sebelumnya, ringkasan perbandingan penelitian terdahulu disajikan pada Tabel 2.1.

Tabel 2. 1 Daftar Penelitian Terdahulu

No	Peneliti	Domain	Pendekatan	Kontribusi Penelitian	Keterbatasan	Kesenjangan Penelitian yang Diatasi
1.	(Mumtaz et al., 2025)	Kebijakan Efisiensi Anggaran	Sentiment Analysis (SVM)	Menunjukkan bahwa analisis sentimen mampu menggambarkan kecenderungan opini masyarakat terhadap kebijakan efisiensi anggaran.	Analisis masih pada tingkat kalimat sehingga belum mampu mengidentifikasi aspek kebijakan yang menjadi sumber sentimen	Mengembangkan ABSA berbasis NER sehingga aspek dan sentimen dapat diekstraksi secara bersamaan.
2.	(Alghifari & Azis, 2025)	Kebijakan Efisiensi Anggaran	Sentiment Analysis (Naïve Bayes)	Membuktikan bahwa metode klasifikasi tradisional dapat digunakan untuk menganalisis opini	Belum mampu memahami hubungan antara aspek dan sentimen serta belum	Menggunakan pendekatan sequence labeling untuk mengekstraksi aspek beserta sentimennya.

No	Peneliti	Domain	Pendekatan	Kontribusi Penelitian	Keterbatasan	Kesenjangan Penelitian yang Diatasi
				masyarakat terhadap kebijakan pemerintah.	melakukan ekstraksi aspek.	
3.	(Putra et al., 2024)	Pariwisata	Aspect-Based Sentiment Analysis	Menunjukkan bahwa ABSA menghasilkan informasi yang lebih rinci dibandingkan analisis sentimen konvensional.	Domain penelitian berupa pariwisata sehingga karakteristik aspek lebih eksplisit dibandingkan kebijakan publik.	Menerapkan ABSA pada domain kebijakan publik yang memiliki aspek lebih konseptual dan implisit.
4.	(Sun et al., 2019)	Aspect-Based Sentiment Analysis	Transformer (BERT)	Menunjukkan bahwa representasi kontekstual meningkatkan performa ABSA.	Menggunakan dataset benchmark internasional dan tidak membahas pembangunan dataset.	Mengevaluasi beberapa arsitektur *deep learning* menggunakan dataset yang dibangun sendiri pada domain kebijakan publik

No	Peneliti	Domain	Pendekatan	Kontribusi Penelitian	Keterbatasan	Kesenjangan Penelitian yang Diatasi
5.	(Hasudungan & Kodhra, 2022)	Deteksi Emosi Bahasa Indonesia	Sequence Labeling	Membuktikan bahwa *sequence labeling* efektif melakukan ekstraksi informasi pada tingkat token untuk teks berbahasa Indonesia.	Fokus pada deteksi emosi, bukan ekstraksi aspek kebijakan.	Mengadaptasi pendekatan *sequence labeling* untuk ABSA berbasis NER pada opini kebijakan publik.
6.	(Hellwig et al., 2026)	Aspect-Based Sentiment Analysis	*Annotation Guideline* & Dataset Construction	Menunjukkan pentingnya pedoman anotasi dan konsistensi pelabelan dalam pembangunan dataset ABSA.	Tidak berfokus pada domain kebijakan publik berbahasa Indonesia maupun validasi pakar.	Membangun dataset menggunakan Annotation Guideline, validasi pakar, dan Inter-Rater Reliability.
7.	(Rahmat et al., 2026)	Named Entity Recognition Bahasa Indonesia	BiLSTM-CRF	Menunjukkan bahwa peningkatan ukuran dataset berpengaruh terhadap performa model NER.	Tidak membahas ekstraksi aspek pada domain kebijakan publik.	Menganalisis pengaruh penambahan ukuran dataset terhadap performa empat model *deep learning* pada ABSA berbasis NER.

No	Peneliti	Domain	Pendekatan	Kontribusi Penelitian	Keterbatasan	Kesenjangan Penelitian yang Diatasi
8.	Penelitian yang diusulkan	Kebijakan Efisiensi Anggaran pada Media Sosial X	Aspect-Based Sentiment Analysis berbasis Named Entity Recognitio	Membangun dataset menggunakan Annotation Guideline, anotasi manual melalui Prodigy, validasi pakar menggunakan *Fleiss' Kappa*, mengevaluasi empat arsitektur (SimpleRNN-CRF, BiGRU-CRF, BiLSTM-CRF, Transformer Encoder-CRF), serta menganalisis pengaruh penambahan ukuran dataset menggunakan *Exact Match* dan *Partial Match*.		Memberikan kontribusi berupa pembangunan gold-standard dataset pada domain kebijakan publik, evaluasi komprehensif empat arsitektur *deep learning*, serta analisis pengaruh ukuran dataset terhadap kemampuan generalisasi model.

Berdasarkan Tabel 2.1. dapat diketahui bahwa penelitian mengenai analisis opini masyarakat telah berkembang dari *Sentiment Analysis* menuju ABSA dengan pendekatan NER. Perkembangan tersebut menunjukkan bahwa penelitian tidak lagi hanya berfokus pada klasifikasi sentimen secara umum, tetapi juga pada ekstraksi aspek untuk memperoleh informasi yang lebih rinci. Selain itu, berbagai penelitian telah memanfaatkan arsitektur *deep learning* untuk meningkatkan performa model, namun sebagian besar masih menggunakan domain komersial, dataset yang berbeda-beda, serta proses anotasi yang belum dilakukan secara seragam.

Berdasarkan hasil kajian tersebut, masih terdapat beberapa kesenjangan penelitian, yaitu terbatasnya penerapan ABSA berbasis NER pada domain kebijakan publik berbahasa Indonesia, belum banyaknya penelitian yang membangun dataset melalui *Annotation Guideline*, validasi pakar, dan *Inter-Rater Reliability,* serta masih terbatasnya evaluasi beberapa arsitektur *deep learning* menggunakan dataset yang dibangun melalui prosedur anotasi yang sama. Oleh karena itu, penelitian ini membangun dataset ABSA berbasis NER yang tervalidasi, kemudian mengevaluasi performa SimpleRNN-CRF, BiGRU-CRF, BiLSTM-CRF, dan Transformer Encoder-CRF, serta menganalisis pengaruh penambahan ukuran dataset terhadap kemampuan generalisasi model menggunakan evaluasi *Exact Span Matching* dan *Partial Span Matching.*