

Analisis Sentimen Berbasis Aspek Melalui Integrasi LDA dan Kombinasi XLM-RoBERTa - BiGRU

Pande Komang Tio Gemisatya¹, I Nyoman Saputra Wahyu Wijaya^{*2},

Ni Putu Novita Puspa Dewi³

Universitas Pendidikan Ganesha, Singaraja, Indonesia^{1,2,3}

pande.tio@student.undiksha.ac.id¹, wahyu.wijaya@undiksha.ac.id²,

novita.puspa.dewi@undiksha.ac.id³

*Corresponding author : pande.tio@student.undiksha.ac.id¹

Abstrak—Analisis sentimen pada media sosial umumnya hanya menghasilkan polaritas sentimen secara keseluruhan tanpa mengidentifikasi aspek spesifik yang menjadi fokus opini masyarakat. Kondisi tersebut menyebabkan informasi yang diperoleh menjadi kurang rinci, terutama pada opini yang membahas lebih dari satu aspek kebijakan pemerintah. Penelitian ini bertujuan mengembangkan sistem *Aspect-Based Sentiment Analysis* (ABSA) untuk menganalisis sentimen masyarakat terhadap kebijakan pemerintahan Prabowo Subianto. Sistem yang diusulkan menggunakan TF-IDF dan *Latent Dirichlet Allocation* (LDA) untuk ekstraksi aspek, kemudian menerapkan *multi-aspect annotation* dalam pembentukan *dataset*, serta memanfaatkan model *hybrid* XLM-RoBERTa dan *Bidirectional Gated Recurrent Unit* (BiGRU) untuk klasifikasi sentimen. Evaluasi model dilakukan menggunakan *Stratified 5-Fold Cross Validation*, sedangkan kualitas ekstraksi aspek dievaluasi menggunakan *Coherence Score*, *Topic Diversity*, dan *Silhouette Score*. Hasil penelitian menunjukkan bahwa model LDA berhasil mengidentifikasi tiga aspek utama dengan nilai *Coherence Score* sebesar 0,5427, *Topic Diversity* sebesar 0,9333, dan *Silhouette Score* sebesar 0,8496. Model klasifikasi menghasilkan *Accuracy* sebesar 95,35%, *Precision* sebesar 95,41%, *Recall* sebesar 95,35%, dan *F1-score* sebesar 95,35%. Hasil tersebut menunjukkan bahwa *framework* ABSA yang diusulkan mampu melakukan ekstraksi aspek dan klasifikasi sentimen secara otomatis dengan performa yang tinggi dan konsisten pada data opini masyarakat berbahasa Indonesia.

Abstract—Conventional sentiment analysis on social media generally produces only overall sentiment polarity without identifying the specific aspects discussed in public opinions. As a result, the extracted information becomes less informative, particularly for opinions addressing multiple government policy issues. This study proposes an *Aspect-Based Sentiment Analysis* (ABSA) system to analyze public sentiment toward the policies of the Prabowo Subianto administration. The proposed approach employs TF-IDF and *Latent Dirichlet Allocation* (LDA) for aspect extraction, applies *multi-aspect annotation* for dataset construction, and utilizes a *hybrid* XLM-RoBERTa and *Bidirectional Gated Recurrent Unit* (BiGRU) model for aspect-level sentiment classification. Model performance was evaluated using *Stratified 5-Fold Cross Validation*, while aspect extraction quality was assessed using *Coherence Score*, *Topic Diversity*, and *Silhouette Score*. The experimental results show that the LDA model successfully identified three main aspects with a *Coherence Score* of 0.5427, a *Topic Diversity* of 0.9333, and a *Silhouette Score* of 0.8496. The proposed sentiment classification model achieved an *Accuracy* of 95.35%, *Precision* of 95.41%, *Recall* of 95.35%, and an *F1-score* of 95.35%. These findings demonstrate that the proposed ABSA framework can automatically perform aspect extraction and aspect-level sentiment classification with high and consistent performance on Indonesian social media opinion data.

Keywords— *Aspect-Based Sentiment Analysis*, *XLM-RoBERTa*, *Bidirectional Gated Recurrent Unit*, *Latent Dirichlet Allocation*, *Sentiment Analysis*

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc/4.0/) license.



1. Pendahuluan

Pemerintahan Prabowo Subianto dan Gibran Rakabuming Raka yang dimulai pada 20 Oktober 2024 menghadirkan berbagai kebijakan strategis yang memunculkan respons publik yang beragam. Beberapa kebijakan, seperti rencana kenaikan Pajak Pertambahan Nilai (PPN) sebesar 12%, larangan pengecer menjual elpiji 3 kg, dan Program Makan Bergizi Gratis (MBG), menjadi perhatian masyarakat karena menuai dukungan maupun kritik terhadap proses implementasi, kualitas pelaksanaan, serta pengelolaan anggarannya [1][2]. Respons masyarakat terhadap kebijakan pemerintah tidak hanya mencerminkan perbedaan pandangan publik, tetapi juga dapat memengaruhi legitimasi pemerintahan, efektivitas komunikasi publik, serta keberhasilan implementasi kebijakan [3][4]. Oleh karena itu, pemantauan opini masyarakat menjadi bagian penting dalam proses evaluasi kebijakan publik.

Media sosial, khususnya platform X (sebelumnya *Twitter*), menjadi salah satu media utama bagi masyarakat untuk menyampaikan opini terhadap berbagai isu politik dan kebijakan pemerintah. Dengan tingkat penetrasi internet di Indonesia yang telah mencapai lebih dari 79,5% populasi, X menjadi ruang diskusi digital yang aktif dan menghasilkan data opini dalam jumlah besar [5]. Karakteristik *tweet* yang bersifat spontan, ringkas, serta didukung penggunaan tagar, balasan, dan *retweet* menjadikan platform ini sebagai sumber data yang representatif untuk menggambarkan persepsi masyarakat terhadap kebijakan pemerintah [6][7].

Meskipun demikian, analisis opini pada media sosial masih menjadi tantangan karena volume data yang sangat besar, penggunaan bahasa yang beragam, serta kemungkinan satu *tweet* mengandung opini terhadap lebih dari satu aspek kebijakan. Analisis sentimen konvensional umumnya hanya menghasilkan satu label sentimen terhadap keseluruhan teks sehingga belum mampu memberikan informasi mengenai aspek kebijakan yang menjadi penyebab munculnya sentimen tersebut. Kondisi ini memerlukan pendekatan yang mampu mengidentifikasi aspek yang dibahas sekaligus menentukan polaritas sentimen pada setiap aspek. Salah satu pendekatan yang memenuhi kebutuhan tersebut adalah Analisis Sentimen Berbasis Aspek atau *Aspect-Based Sentiment Analysis* (ABSA), yang memungkinkan analisis opini dilakukan secara lebih rinci dengan menghubungkan sentimen terhadap aspek-aspek tertentu dalam suatu teks [8][9].

Implementasi ABSA umumnya terdiri atas dua tahapan utama, yaitu ekstraksi aspek dan klasifikasi sentimen. Pada tahap ekstraksi aspek, penelitian ini menggunakan *Latent Dirichlet Allocation* (LDA) karena mampu menemukan aspek atau topik secara otomatis tanpa memerlukan anotasi aspek, sehingga sesuai untuk data media sosial yang umumnya tidak memiliki label aspek [10]. Selain itu, penelitian sebelumnya menunjukkan bahwa LDA mampu menghasilkan topik yang lebih koheren dibandingkan beberapa metode *topic modeling* lainnya sehingga banyak digunakan dalam proses ekstraksi aspek pada data teks tidak terstruktur [11].

Tahap berikutnya adalah klasifikasi sentimen menggunakan kombinasi XLM-RoBERTa dan *Bidirectional Gated Recurrent Unit* (BiGRU). XLM-RoBERTa dipilih karena mampu menghasilkan representasi kontekstual multibahasa yang kuat, termasuk untuk bahasa Indonesia, serta menunjukkan performa yang lebih baik dibandingkan beberapa model *transformer* lainnya pada tugas klasifikasi sentimen multibahasa [12]. Sementara itu, BiGRU digunakan untuk memperkuat pemodelan dependensi sekuensial dua arah sehingga hubungan antar kata dalam suatu kalimat dapat dipahami secara lebih komprehensif [13]. Kombinasi kedua metode tersebut diharapkan mampu menghasilkan representasi konteks yang lebih baik sehingga meningkatkan performa klasifikasi sentimen berbasis aspek.

Berbagai penelitian sebelumnya telah menunjukkan potensi metode-metode tersebut dalam pengembangan ABSA maupun analisis sentimen. Penelitian oleh Pathan dan Prakash [11] menunjukkan bahwa LDA menghasilkan topik yang lebih koheren dibandingkan beberapa metode *topic modeling* lainnya untuk proses ekstraksi aspek. Sementara itu, Prytula [12] melaporkan bahwa XLM-RoBERTa memberikan performa yang lebih baik dibandingkan BERT dan DistilBERT pada klasifikasi sentimen multibahasa. Selain itu, Trang [13] menunjukkan bahwa arsitektur RoBERTa-BiGRU mampu menghasilkan performa yang lebih tinggi dibandingkan beberapa arsitektur lain, seperti RoBERTa-LSTM dan CNN-LSTM, pada tugas analisis sentimen. Hasil-hasil penelitian tersebut menunjukkan bahwa LDA dan model berbasis *transformer* memiliki potensi yang besar untuk diterapkan dalam analisis sentimen berbasis aspek.

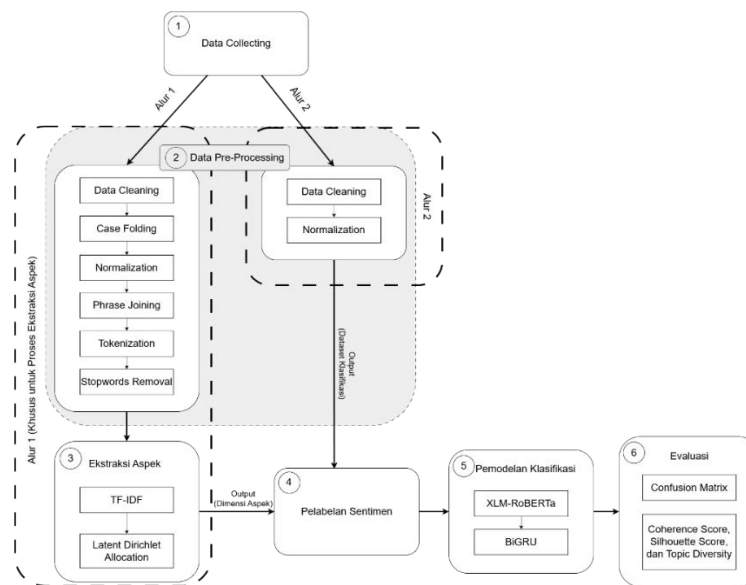
Meskipun demikian, berdasarkan kajian terhadap penelitian-penelitian sebelumnya, belum ditemukan penelitian yang mengintegrasikan LDA sebagai metode ekstraksi aspek dengan XLM-RoBERTa dan BiGRU sebagai model klasifikasi sentimen dalam sebuah *framework Aspect-Based Sentiment Analysis* terhadap opini masyarakat Indonesia mengenai kebijakan pemerintahan di platform X. Sebagian besar penelitian masih menggunakan kombinasi LDA dengan model *transformer* lain atau hanya berfokus pada klasifikasi sentimen tanpa mengintegrasikan proses ekstraksi aspek dan klasifikasi sentimen dalam satu alur pemrosesan. Kondisi tersebut menunjukkan adanya peluang untuk mengembangkan pendekatan yang menggabungkan keunggulan masing-masing metode sehingga mampu menghasilkan analisis opini yang lebih komprehensif.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sebuah *framework Aspect-Based Sentiment Analysis* untuk menganalisis opini masyarakat mengenai kebijakan pemerintahan Prabowo Subianto di platform X menggunakan kombinasi LDA, XLM-RoBERTa, dan BiGRU. *Framework* yang diusulkan mencakup proses pengumpulan data, *pre-processing* yang disesuaikan untuk ekstraksi aspek dan klasifikasi sentimen, ekstraksi aspek menggunakan LDA, klasifikasi sentimen menggunakan XLM-RoBERTa-BiGRU, serta evaluasi performa menggunakan metrik *topic modeling* dan klasifikasi. Kontribusi penelitian ini adalah menghasilkan pendekatan ABSA yang mampu mengidentifikasi aspek kebijakan beserta sentimen masyarakat secara otomatis pada data media sosial berbahasa Indonesia,

sehingga dapat memberikan informasi yang lebih komprehensif mengenai respons masyarakat terhadap kebijakan pemerintahan Prabowo Subianto.

2. Metodologi Penelitian

Penelitian ini mengusulkan sebuah *framework* ABSA untuk menganalisis sentimen masyarakat terhadap kebijakan pemerintahan Prabowo Subianto pada platform X. Berbeda dengan pendekatan ABSA yang memerlukan anotasi aspek secara manual, penelitian ini mengintegrasikan ekstraksi aspek menggunakan kombinasi TF-IDF dan LDA dengan klasifikasi sentimen berbasis XLM-RoBERTa dan BiGRU dalam satu alur pemrosesan. Selain itu, penelitian ini menerapkan dua alur *pre-processing* yang berbeda untuk menyesuaikan karakteristik *topic modeling* dan *transformer*, mekanisme *multi-aspect annotation* untuk menangani *tweet* yang membahas lebih dari satu aspek, serta *ensemble majority voting* untuk meningkatkan stabilitas prediksi sentimen. Keseluruhan tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

Proses penelitian diawali dengan pengumpulan data menggunakan teknik *data crawling* pada platform X melalui *library Tweet Harvest*. Pengambilan data dilakukan menggunakan kumpulan kata kunci yang merepresentasikan berbagai kebijakan pemerintahan Prabowo Subianto selama periode penelitian. Dataset hasil *crawling* kemudian diseleksi secara manual untuk menghilangkan data yang tidak relevan, seperti berita informatif, promosi, spam, maupun *tweet* yang tidak mengandung opini terhadap kebijakan pemerintah, sehingga diperoleh dataset yang layak digunakan pada proses analisis sentimen berbasis aspek. Dataset yang telah diseleksi selanjutnya diproses menggunakan dua alur *pre-processing* yang berbeda sesuai dengan kebutuhan masing-masing model. Alur pertama ditujukan untuk proses ekstraksi aspek menggunakan TF-IDF dan LDA, sedangkan alur kedua digunakan sebagai masukan model klasifikasi sentimen berbasis XLM-RoBERTa dan BiGRU. Pada Alur 1 diterapkan *data cleaning*, *case folding*, normalisasi kata, *phrase joining*, tokenisasi, dan *stopword removal* untuk menghasilkan representasi dokumen yang sesuai bagi proses *topic modeling*. Sementara itu, Alur 2 mempertahankan struktur kalimat melalui *data cleaning*, dan normalisasi tanpa melakukan tokenisasi maupun *stopword removal*, sehingga informasi kontekstual tetap terjaga ketika diproses oleh model *transformer*.

Hasil *pre-processing* Alur 1 digunakan sebagai masukan TF-IDF untuk mengubah dokumen menjadi representasi numerik berdasarkan tingkat kepentingan setiap kata dan frasa. Representasi tersebut kemudian diproses menggunakan LDA untuk membentuk distribusi topik pada setiap dokumen sekaligus menghasilkan distribusi probabilitas kata pada masing-masing topik. Berdasarkan probabilitas tersebut, setiap *tweet* diberikan label aspek sesuai topik dengan probabilitas tertinggi. Selain itu, penelitian ini menerapkan mekanisme *multi-aspect annotation*, yaitu memberikan lebih dari satu label aspek apabila probabilitas topik sekunder memenuhi ambang batas yang telah ditentukan. Pendekatan ini memungkinkan

satu *tweet* direpresentasikan menjadi beberapa pasangan aspek-sentimen sehingga model dapat mempelajari sentimen secara spesifik pada setiap aspek yang dibahas.

Setelah proses ekstraksi aspek selesai, dilakukan pelabelan sentimen secara manual terhadap setiap pasangan aspek-*tweet* ke dalam dua kategori, yaitu positif dan negatif. Dataset hasil pelabelan kemudian digunakan untuk membangun model klasifikasi sentimen berbasis XLM-RoBERTa dan BiGRU. Teks hasil *pre-processing* Alur 2 terlebih dahulu ditokenisasi menggunakan *tokenizer* bawaan XLM-RoBERTa untuk menghasilkan *input IDs* dan *attention mask*. Selanjutnya, XLM-RoBERTa menghasilkan *contextual embedding* yang merepresentasikan makna setiap token berdasarkan konteks kalimat. Representasi kontekstual tersebut kemudian diproses oleh BiGRU untuk mempelajari hubungan sekuensial antar token dari arah maju dan mundur sebelum dipetakan ke kelas sentimen menggunakan lapisan klasifikasi.

Pelatihan model dilakukan menggunakan *Stratified 5-Fold Cross Validation* agar distribusi kelas sentimen tetap seimbang pada setiap *fold*. Proses *fine-tuning* XLM-RoBERTa dilakukan pada empat lapisan teratas menggunakan *optimizer AdamW*, *Weighted Cross Entropy Loss*, *learning rate scheduler*, *weight decay*, serta mekanisme *early stopping* berdasarkan nilai *validation F1-score*. Selanjutnya, keluaran XLM-RoBERTa digunakan sebagai masukan BiGRU yang dilatih untuk memperkuat representasi sekuensial sebelum menghasilkan prediksi sentimen. Pada tahap inferensi, lima model terbaik hasil proses *cross validation* digabungkan menggunakan metode *ensemble majority voting*, sehingga prediksi akhir diperoleh berdasarkan keputusan mayoritas seluruh model dan menghasilkan prediksi yang lebih stabil dibandingkan penggunaan satu model tunggal.

Tabel 1. *Hyperparameter Model LDA, XLM-RoBERTa, dan BiGRU*

Parameter	Nilai
TF-IDF <i>max_features</i>	1000
TF-IDF <i>min_df</i>	3
TF-IDF <i>max_df</i>	0,7
TF-IDF <i>ngram</i>	(1,2)
Jumlah Topik LDA	3
<i>Learning Method</i> LDA	<i>Batch</i>
<i>Sequence Length</i>	128
<i>Fine-tuning</i> XLM-RoBERTa	4 lapisan teratas
<i>Hidden Size</i> XLM-RoBERTa	768
<i>Hidden Size</i> BiGRU	384 (setiap arah)
<i>Batch Size</i>	16
<i>Learning Rate</i> XLM-RoBERTa	1×10^{-5}
<i>Learning Rate</i> BiGRU	8×10^{-4}
<i>Optimizer</i>	AdamW
<i>Dropout</i> BiGRU	0,3
<i>Weight Decay</i>	0,01
<i>Loss Function</i>	<i>Weighted Cross Entropy Loss</i>
<i>Cross Validation</i>	<i>Stratified 5-Fold</i>

Kualitas model LDA dievaluasi menggunakan *Coherence Score*, *Silhouette Score*, dan *Topic Diversity* untuk mengukur kualitas topik yang dihasilkan dari aspek koherensi, keberagaman kata kunci antar topik, serta pemisahan antar kelompok topik. Selanjutnya, kinerja model dievaluasi menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan model dalam mengklasifikasikan sentimen pada setiap aspek. Seluruh tahapan pada penelitian ini membentuk sebuah *framework* ABSA yang terintegrasi, dimulai dari masukan berupa teks *tweet*, dilanjutkan dengan *pre-processing*, ekstraksi aspek menggunakan TF-IDF dan LDA, klasifikasi sentimen menggunakan XLM-RoBERTa dan BiGRU, hingga menghasilkan prediksi aspek dan sentimen secara otomatis dalam satu kali proses (*single-run pipeline*). Dengan demikian, kerangka kerja yang diusulkan memungkinkan sebuah sistem ABSA melakukan analisis terhadap masukan teks secara *end-to-end* tanpa memerlukan intervensi pengguna pada setiap tahapan pemrosesan.

3. Hasil dan Pembahasan

Penelitian ini mengembangkan sistem ABSA terhadap opini masyarakat mengenai kebijakan pemerintahan Prabowo Subianto dengan menggunakan *framework* yang terdiri atas ekstraksi aspek berbasis TF-IDF dan LDA, pembentukan pasangan aspek-sentimen melalui *multi-aspect annotation*, serta klasifikasi sentimen menggunakan model *hybrid* XLM-RoBERTa dan BiGRU. Seluruh tahapan dievaluasi secara berurutan

sesuai dengan alur penelitian sehingga hasil yang diperoleh menggambarkan performa masing-masing komponen dalam *framework* yang diusulkan.

Tahap pertama pada *framework* bertujuan mengidentifikasi aspek-aspek utama yang dibahas dalam kumpulan *tweet*. Dokumen yang telah melalui proses *pre-processing* direpresentasikan menggunakan TF-IDF sebelum diproses oleh algoritma LDA untuk membentuk distribusi topik. Evaluasi kualitas model dilakukan menggunakan tiga metrik, yaitu *Coherence Score*, *Silhouette Score*, dan *Topic Diversity*.

Tabel 2. Hasil Evaluasi Model LDA

Coherence Score	Silhouette Score	Topic Diversity
0.5427	0.8496	0.9333

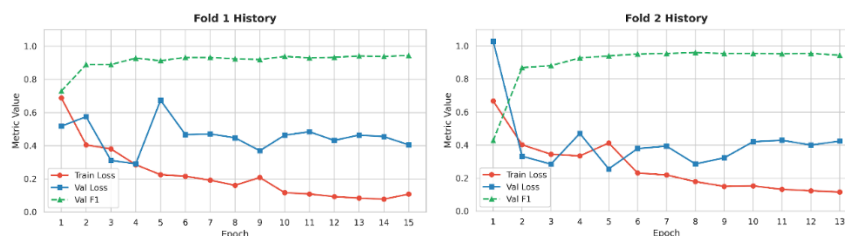
Berdasarkan hasil evaluasi pada Tabel 2, model LDA menghasilkan nilai *Coherence Score* sebesar 0.5427, *Silhouette Score* sebesar 0.8496, dan *Topic Diversity* sebesar 0.9333. Ketiga metrik tersebut menunjukkan bahwa model mampu menghasilkan topik yang relatif koheren, memiliki keberagaman kata kunci yang tinggi, serta menghasilkan pemisahan topik yang cukup baik. Selanjutnya, setiap topik diinterpretasikan berdasarkan kata-kata dengan probabilitas tertinggi sehingga diperoleh representasi aspek yang digunakan pada proses anotasi dataset.

Tabel 3. Interpretasi Aspek Berdasarkan Hasil LDA

Topik	Kata Kunci Utama	Interpretasi Aspek
0	nasional, ekonomi, melanjutkan, pembangunan, <i>food_estate</i> , dukung, makan_bergizi_gratis, ketahanan_pangan, ibu_kota_nusantara, efisiensi_anggaran	Pembangunan Nasional dan Ketahanan Pangan
1	pajak mencekik, mencekik tolak, tolak persen, tolak pajak, persen mencekik, pajak persen, tolak, mencekik, persen, pajak	Kebijakan Fiskal dan Pengelolaan Anggaran Negara
2	pembatalan, peningkatan kesejahteraan, pemerintah, peningkatan, makan_bergizi_gratis, guru, kesejahteraan, efisiensi_anggaran pendidikan, efisiensi_anggaran, pendidikan	Pendidikan, Kesejahteraan, dan Pembangunan SDM

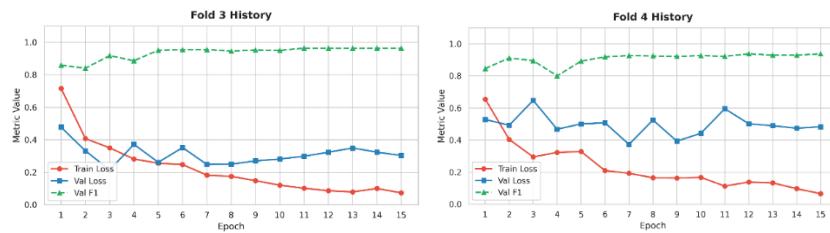
Hasil interpretasi menunjukkan bahwa model berhasil mengelompokkan pembahasan kebijakan pemerintah ke dalam tiga aspek utama, yaitu Pembangunan Nasional dan Ketahanan Pangan, Kebijakan Fiskal dan Pengelolaan Anggaran Negara, serta Pendidikan, Kesejahteraan, dan Pembangunan SDM. Ketiga aspek tersebut selanjutnya digunakan sebagai dasar dalam proses *multi-aspect annotation* sehingga setiap *tweet* dapat dipasangkan dengan satu atau lebih aspek sesuai distribusi probabilitas yang dihasilkan oleh model LDA.

Dataset hasil anotasi aspek kemudian digunakan pada tahap klasifikasi sentimen menggunakan model *hybrid* XLM-RoBERTa dan BiGRU. Evaluasi dilakukan menggunakan *Stratified 5-Fold Cross Validation* sehingga setiap *fold* menghasilkan model yang dilatih dan diuji pada pembagian data yang berbeda. Selama proses pelatihan, performa model dipantau menggunakan nilai *training loss*, *validation loss*, dan *validation F1-score*.



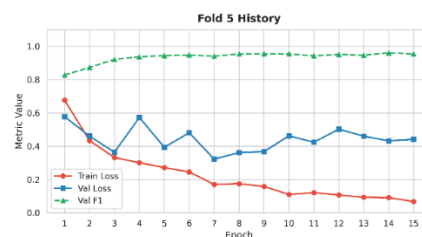
Gambar 2. Learning History Fold 1 dan Fold 2

Pada *Fold 1*, *training loss* mengalami penurunan secara bertahap selama proses pelatihan, sedangkan *validation loss* menunjukkan fluktuasi pada beberapa *epoch*. Nilai *validation F1-score* meningkat hingga mencapai nilai terbaik sebesar 0.9442 pada *epoch* ke-15 yang kemudian digunakan sebagai model terbaik pada *fold* tersebut. Pada *Fold 2*, model menunjukkan peningkatan *validation F1-score* yang cukup cepat sejak awal pelatihan dan mencapai nilai terbaik sebesar 0.9609 sebelum proses pelatihan dihentikan menggunakan mekanisme *early stopping*.



Gambar 3. Learning History Fold 3 dan Fold 4

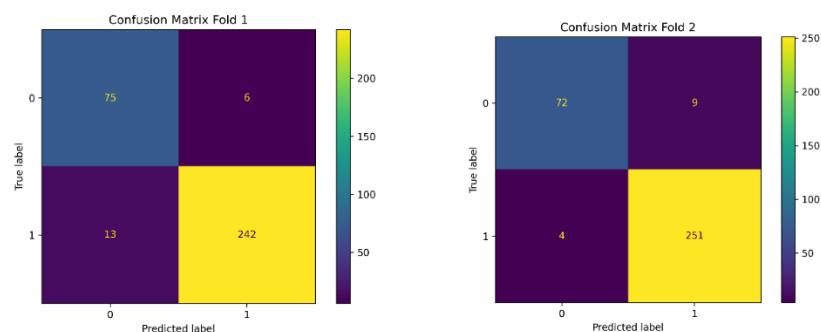
Pada *Fold 3*, penurunan *training loss* berlangsung secara konsisten hingga akhir pelatihan. Nilai *validation F1-score* mencapai nilai terbaik sebesar 0.9640, yang merupakan nilai tertinggi dibandingkan *fold* lainnya. *Fold 4* memperlihatkan pola pelatihan yang relatif stabil meskipun *validation loss* mengalami fluktuasi pada beberapa *epoch*. Nilai *validation F1-score* terbaik yang diperoleh adalah 0.9377.



Gambar 4. Learning History Fold 5

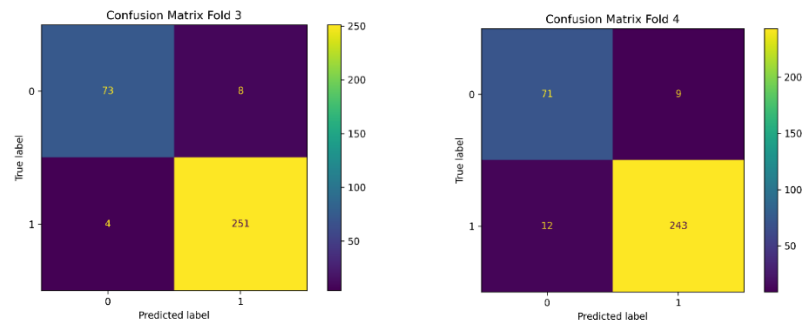
Pada *Fold 5*, model kembali menunjukkan penurunan *training loss* yang konsisten disertai peningkatan *validation F1-score* hingga mencapai nilai terbaik sebesar 0.9608 pada *epoch* ke-14. Secara umum, seluruh *fold* menunjukkan kecenderungan penurunan *training loss* selama proses pelatihan, sementara *validation F1-score* meningkat hingga mencapai titik terbaik pada masing-masing *fold*. Mekanisme *early stopping* dan penyimpanan model terbaik berdasarkan nilai *validation F1-score* memastikan bahwa model yang digunakan pada tahap evaluasi merupakan model dengan performa validasi terbaik.

Setelah proses pelatihan selesai, masing-masing model dievaluasi menggunakan *confusion matrix* untuk menggambarkan distribusi prediksi terhadap setiap kelas sentimen.



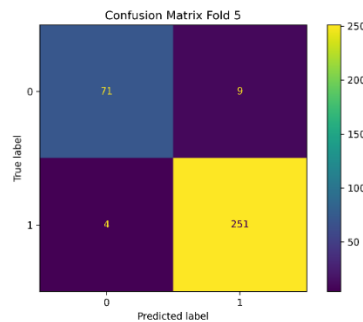
Gambar 5. Confusion Matrix Fold 1 dan Fold 2

Confusion matrix pada *Fold 1* menunjukkan bahwa sebagian besar data berhasil diklasifikasikan pada kelas yang sesuai dengan label sebenarnya, dengan jumlah kesalahan klasifikasi yang relatif kecil. Pada *Fold 2*, distribusi prediksi menunjukkan peningkatan jumlah prediksi yang benar dibandingkan *Fold 1*, terutama pada kelas sentimen positif.



Gambar 6. Confusion Matrix Fold 3 dan Fold 4

Fold 3 menghasilkan distribusi prediksi yang paling baik dibandingkan *fold* lainnya dengan jumlah kesalahan klasifikasi yang paling sedikit. *Confusion matrix* pada *Fold 4* menunjukkan bahwa sebagian besar data tetap berhasil diprediksi dengan benar meskipun jumlah kesalahan klasifikasi sedikit lebih tinggi dibandingkan *Fold 2* dan *Fold 3*.



Gambar 7. Confusion Matrix Fold 5

Fold 5 kembali menunjukkan distribusi prediksi yang seimbang dengan jumlah prediksi benar yang dominan pada kedua kelas sentimen. Seluruh *confusion matrix* menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data uji ke dalam kelas sentimen yang benar pada setiap *fold*, sehingga menghasilkan performa yang relatif konsisten selama proses validasi silang.

Setelah dilakukan evaluasi pada masing-masing *fold*, nilai *accuracy*, *precision*, *recall*, dan *F1-score* dari setiap model dirangkum untuk melihat konsistensi performa selama proses *Stratified 5-Fold Cross Validation*.

Tabel 4. Hasil Evaluasi Setiap Fold

Fold	Accuracy	Precision	Recall	F1-Score
1	0.943452	0.946026	0.943452	0.944230
2	0.961310	0.961041	0.961310	0.960881
3	0.964286	0.964035	0.964286	0.963973
4	0.937313	0.938288	0.937313	0.937705
5	0.961194	0.960915	0.961194	0.960757

Berdasarkan Tabel 4, performa model pada seluruh *fold* menunjukkan hasil yang relatif konsisten. Nilai *accuracy* berada pada rentang 93,73% hingga 96,43%, sedangkan nilai *F1-score* berada pada rentang 93,77% hingga 96,40%. *Fold 3* memperoleh performa terbaik dengan nilai *accuracy* sebesar 96,43% dan *F1-score* sebesar 96,40%, sedangkan *Fold 4* menghasilkan performa terendah dengan nilai *accuracy* sebesar 93,73% dan *F1-score* sebesar 93,77%. Meskipun demikian, selisih performa antar *fold* relatif kecil sehingga menunjukkan bahwa model menghasilkan performa yang stabil pada berbagai pembagian data pelatihan dan data pengujian. Untuk memperoleh gambaran performa keseluruhan model, dilakukan perhitungan rata-rata hasil evaluasi dari seluruh *fold*.

Tabel 5. Hasil Evaluasi Keseluruhan Model

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
95.35%	95.41%	95.35%	95.35%

Hasil evaluasi keseluruhan menunjukkan bahwa model menghasilkan *accuracy* sebesar 95,35%, *precision* sebesar 95,41%, *recall* sebesar 95,35%, dan *F1-score* sebesar 95,35%. Nilai yang relatif seimbang pada seluruh metrik menunjukkan bahwa model mampu mempertahankan keseimbangan antara kemampuan mengenali data positif maupun negatif selama proses klasifikasi.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa kombinasi XLM-RoBERTa dan BiGRU menghasilkan performa klasifikasi yang konsisten pada seluruh proses *cross validation*. Perbedaan performa antar *fold* relatif kecil sehingga menunjukkan bahwa model mampu mempertahankan kemampuan klasifikasi meskipun menggunakan pembagian data pelatihan dan data pengujian yang berbeda.

Hasil yang diperoleh juga menunjukkan performa yang kompetitif apabila dibandingkan dengan beberapa penelitian sebelumnya. Penelitian oleh Trang (2023) melaporkan akurasi sebesar 95,35% menggunakan kombinasi RoBERTa dan BiGRU pada dataset IMDb, sedangkan Ri Shin et al. (2021) memperoleh akurasi 90,15% menggunakan XLM-RoBERTa dan BiGRU pada ulasan film berbahasa Korea. Pada penelitian ABSA, Noprianto Kambayo et al. (2025) melaporkan akurasi 96,71% menggunakan kombinasi LDA dan DistilBERT, sementara Andono et al. (2022) memperoleh akurasi 97% menggunakan LDA, *Semantic Similarity*, dan BERT pada ulasan hotel. Dibandingkan dengan penelitian-penelitian tersebut, model yang diusulkan pada penelitian ini menghasilkan akurasi 95,35% dan *F1-score* 95,35%, sehingga menunjukkan performa yang sebanding dengan berbagai pendekatan terkini meskipun diterapkan pada data Twitter berbahasa Indonesia yang memiliki karakteristik lebih tidak terstruktur dan dibangun menggunakan *framework* ABSA otomatis mulai dari ekstraksi aspek hingga klasifikasi sentimen.

Secara umum, hasil perbandingan menunjukkan bahwa performa model yang diusulkan berada pada tingkat yang kompetitif dibandingkan penelitian-penelitian sebelumnya. Meskipun menggunakan domain, bahasa, serta karakteristik dataset yang berbeda, model berhasil menghasilkan performa klasifikasi yang sebanding dengan penelitian berbasis RoBERTa-BiGRU maupun penelitian ABSA yang menggunakan LDA sebagai metode ekstraksi aspek. Temuan ini menunjukkan bahwa *framework* yang diusulkan mampu mempertahankan performa klasifikasi yang tinggi sekaligus mengintegrasikan proses ekstraksi aspek dan klasifikasi sentimen dalam satu alur analisis yang utuh.

Berdasarkan keseluruhan hasil yang diperoleh, *framework Aspect-Based Sentiment Analysis* yang diusulkan mampu mengintegrasikan proses ekstraksi aspek dan klasifikasi sentimen dalam satu alur analisis yang utuh. Pada tahap ekstraksi aspek, kombinasi TF-IDF dan LDA berhasil mengidentifikasi tiga aspek utama yang merepresentasikan pembahasan kebijakan pemerintahan Prabowo Subianto. Aspek-aspek tersebut kemudian digunakan sebagai dasar pembentukan pasangan aspek-sentimen melalui mekanisme *multi-aspect annotation*, sehingga setiap *tweet* dapat direpresentasikan sesuai dengan distribusi topik yang dihasilkan.

Pada tahap klasifikasi sentimen, kombinasi XLM-RoBERTa dan BiGRU menghasilkan performa yang konsisten pada seluruh proses *Stratified 5-Fold Cross Validation*. Grafik *learning curve* menunjukkan bahwa model mengalami penurunan *training loss* selama proses pelatihan disertai peningkatan *validation F1-score* hingga mencapai performa terbaik pada masing-masing *fold*. Selain itu, hasil *confusion matrix* memperlihatkan bahwa sebagian besar data berhasil diklasifikasikan ke dalam kelas sentimen yang sesuai dengan label sebenarnya, yang kemudian tercermin pada nilai rata-rata akurasi sebesar 95,35% dan *F1-score* sebesar 95,35%.

Hasil penelitian juga menunjukkan bahwa pendekatan berbasis *single-run pipeline* ini memberikan kontribusi pada keseluruhan proses analisis. Tahap ekstraksi aspek menghasilkan informasi aspek secara otomatis tanpa memerlukan pelabelan aspek sejak awal, sedangkan model XLM-RoBERTa dan BiGRU memanfaatkan informasi aspek tersebut untuk melakukan klasifikasi sentimen secara lebih spesifik pada setiap kebijakan yang dibahas. Integrasi seluruh tahapan tersebut memungkinkan proses analisis dilakukan secara berurutan mulai dari teks mentah hingga menghasilkan pasangan aspek dan sentimen sebagai keluaran akhir.

Secara keseluruhan, hasil penelitian menunjukkan bahwa metode yang diusulkan mampu menghasilkan sistem ABSA dengan performa klasifikasi yang tinggi dan stabil. Selain menghasilkan model klasifikasi sentimen yang kompetitif, penelitian ini juga menunjukkan bahwa kombinasi TF-IDF, LDA, XLM-RoBERTa, dan BiGRU dapat diimplementasikan sebagai *framework* analisis sentimen berbasis aspek yang bekerja secara otomatis dari proses ekstraksi aspek hingga prediksi sentimen.

4. Kesimpulan

Penelitian ini berhasil mengembangkan sebuah *framework Aspect-Based Sentiment Analysis* untuk menganalisis opini masyarakat terhadap kebijakan pemerintahan Prabowo Subianto dengan mengintegrasikan ekstraksi aspek menggunakan TF-IDF dan LDA serta klasifikasi sentimen menggunakan model *hybrid* XLM-RoBERTa dan BiGRU. *Framework* yang diusulkan mampu mengidentifikasi aspek-aspek utama secara otomatis, membentuk pasangan aspek-sentimen melalui *multi-aspect annotation*, serta mengklasifikasikan sentimen pada setiap aspek dengan performa yang tinggi dan konsisten berdasarkan hasil evaluasi. Hasil penelitian menunjukkan bahwa pendekatan yang diusulkan dapat digunakan sebagai solusi untuk analisis sentimen berbasis aspek pada data media sosial berbahasa Indonesia dan memberikan dasar yang dapat dikembangkan lebih lanjut, terutama melalui pemanfaatan metode ekstraksi aspek berbasis *Transformer* atau *Large Language Model* guna meningkatkan kualitas identifikasi aspek pada penelitian selanjutnya.

5. Daftar Pustaka

- [1] H. K. Muhid, "Kebijakan Maju-Mundur Pemerintahan Prabowo: Kenaikan PPN 12 Persen, Elpiji 3 Kg di Eceran, hingga Anggaran Beasiswa," *Tempo*. Accessed: Mar. 28, 2025. [Online]. Available: <https://www.tempo.co/politik/kebijakan-maju-mundur-pemerintahan-prabowo-kenaikan-ppn-12-persen-elpiji-3-kg-di-eceran-hingga-anggaran-beasiswa-1207810>
- [2] K. Alfisyahri. MJ, "Catatan Kritis Terhadap Program Makan Bergizi Gratis," *CNBC Indonesia*. Accessed: May 26, 2025. [Online]. Available: <https://www.cnbcindonesia.com/opini/20250509141534-14-632444/catatan-kritis-terhadap-program-makan-bergizi-gratis>
- [3] T. Sarihati, M. Luthfie, and B. Kurniadi, *KOMUNIKASI POLITIK, MEDIA MASSA, DAN OPINI PUBLIK*. 2019.
- [4] K. N. Latifah, & Fatma, and U. Najicha, "IMPLIKASI MEDIA SOSIAL TERHADAP FORMULASI KEBIJAKAN PUBLIK," *Jurnal Kewarganegaraan*, vol. 6, no. 1, 2022.
- [5] "APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang," Asosiasi Penyelenggara Jasa Internet Indonesia. Accessed: Mar. 28, 2025. [Online]. Available: <https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang>
- [6] A. Bilbao-Jayo and A. Almeida, "Improving Political Discourse Analysis on Twitter with Context Analysis," *IEEE Access*, vol. 9, pp. 104846–104863, 2021, doi: 10.1109/ACCESS.2021.3099093.
- [7] M. Fachriza and Munawar, "Analisis Sentimen Kalimat Depresi Pada Pengguna Twitter Dengan Naive Bayes, Support Vector Machine, Random Forest," *KOMPUTEK: Jurnal Teknik Universitas Muhammadiyah Ponorogo*, vol. 7, no. 2, pp. 49–58, Oct. 2023, [Online]. Available: <http://studentjournal.umpo.ac.id/index.php/komputeK>
- [8] I. Perikos and A. Diamantopoulos, "Explainable Aspect-Based Sentiment Analysis Using Transformer Models," *Big Data and Cognitive Computing*, vol. 8, no. 11, Nov. 2024, doi: 10.3390/bdcc8110141.
- [9] W. Ahmad, H. U. Khan, F. K. Alarfaj, and M. Alreshoodi, "Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges," 2025, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2025.3555744.
- [10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation Michael I. Jordan," 2003.
- [11] A. F. Pathan and C. Prakash, "Unsupervised Aspect Extraction Algorithm for opinion mining using topic modeling," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 492–499, Nov. 2021, doi: 10.1016/j.gltp.2021.08.005.
- [12] M. Prytula, "Fine-tuning BERT, DistilBERT, XLM-RoBERTa and Ukr-RoBERTa models for sentiment analysis of ukrainian language reviews," *Artif. Intell.*, vol. 29, no. AI.2024.29(2), pp. 85–97, Jun. 2024, doi: 10.15407/jai2024.02.085.
- [13] N. T. M. Trang, "A Hybrid Model of RoBERTa and Bidirectional GRU for Enhanced Sentiment Analysis," *International Journal of Open Information Technologies*, vol. 11, no. 11, pp. 19–24, 2023.

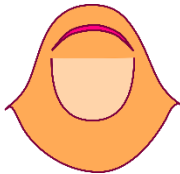
6. Penulis



Pande Komang Tio Gemisatya
Program Studi Ilmu Komputer, Fakultas Teknik dan Kejuruan,
Universitas Pendidikan Ganesha, Singaraja, Indonesia



I Nyoman Saputra Wahyu Wijaya
Dosen Fakultas Teknik dan Kejuruan, Universitas Pendidikan Ganesha,
Singaraja, Indonesia



Ni Putu Novita Puspa Dewi
Dosen Fakultas Teknik dan Kejuruan, Universitas Pendidikan Ganesha,
Singaraja, Indonesia