SEMINAR NASIONAL 2026 Universitas Ma Chung
Prosiding Seminar Nasional Universitas Ma Chung 2026
(Sistem Informasi dan Informatika)
pp. 150-164
E-ISSN: 2808-7771

KONSTRUKSI *DATASET* ANALISIS SENTIMEN BERBASIS ASPEK MENGGUNAKAN BERTopic DAN *HUMAN-IN-THE-LOOP*

STUDI KASUS UMPAN BALIK TERBUKA SISWA SMKN 1 BOJONGGENTENG

Taufik Alfikar[1*]

[1]Magister Informatika, Universitas Nusa Putra, Jl. Raya Cibolang No.21, Cisaat, Sukabumi, Indonesia, 43152

***Email korespondensi:** taufik.alfikar_mif24@nusaputra.ac.id

Received: 27 05 2026 – Revised: 29 07 2026 - Accepted: 04 08 2026 - Published: 05 08 2026

Abstrak. Aspect-Based Sentiment Analysis (ABSA) telah banyak diterapkan untuk mengidentifikasi sentimen terhadap aspek spesifik dalam data tekstual. Namun, ketersediaan dataset ABSA berbahasa Indonesia pada domain pendidikan, khususnya Sekolah Menengah Kejuruan (SMK) masih relatif terbatas. Penelitian ini menyajikan konstruksi dataset ABSA berbahasa Indonesia dari umpan balik terbuka siswa pada sebuah SMK menggunakan pendekatan human-in-the-loop yang mengintegrasikan BERTopic untuk identifikasi kandidat topik secara otomatis dan anotasi manual untuk menghasilkan skema aspek yang kontekstual. Dataset awal terdiri dari 756 respons siswa, kemudian melalui prapemrosesan yang meliputi normalisasi emoji, slang dan typo, serta segmentasi kalimat, diperoleh 717 dokumen dan 1.317 kalimat. BERTopic dioptimalkan melalui grid search kombinasi parameter UMAP-HDBSCAN dan menghasilkan 20 topik laten dengan nilai coherence 0,609, topic diversity 0,974, dan noise ratio 15,4%. Topik tersebut kemudian divalidasi melalui anotasi human-in-the-loop menjadi sembilan kategori aspek bermakna secara kontekstual. Dua orang anotator independen memberikan label sentimen positif, negatif, dan netral. Kesepakatan awal antar-anotator mencapai nilai Cohen's Kappa 0,776 (Baik), meningkat menjadi 0,859 dan 0,912 (Hampir Sempurna) setelah proses adjudikasi. Analisis menunjukkan bahwa aspek Pengalaman Akademik dan Kompetensi Kejuruan didominasi sentimen positif, sedangkan aspek Fasilitas dan Pembelajaran serta Evaluasi dan Harapan memperoleh proporsi sentimen negatif tertinggi. Dataset ini menyediakan sumber daya untuk pengembangan model ABSA berbahasa Indonesia, serta mendukung evaluasi mutu pendidikan yang berbasis data.

Kata kunci: analisis sentimen berbasis aspek, konstruksi dataset, BERTopic, anotasi human-in-the-loop, pendidikan vokasi

PENDAHULUAN

Dataset ini dikembangkan untuk menjawab kebutuhan sumber data dalam menganalisis pengalaman belajar siswa pada pendidikan vokasi, khususnya pada jenjang Sekolah Menengah Kejuruan (SMK) untuk meningkatkan mutu pendidikan. Evaluasi kualitas pembelajaran umumnya dilakukan menggunakan survei kuantitatif berbasis skala Likert untuk mengukur persepsi dan tingkat kepuasan siswa. Namun pendekatan tersebut belum mampu menangkap nuansa dan konteks pengalaman belajar secara mendalam (Dervenis, 2024; Fargues et al., 2023; Kim & Qin, 2023). Umpan balik terbuka berbentuk

teks mulai dimanfaatkan karena dapat memberikan informasi yang lebih kaya terkait opini, emosi dan pengalaman belajar (Alotaibi, 2024; Fargues et al., 2023).

Analisis sentimen pada bidang pendidikan telah berkembang, namun ekstraksi aspek dari umpan balik masih menjadi tantangan. Kim & Qin (2023) menggunakan K-Means dengan penentuan jumlah klaster secara manual yang berpotensi membatasi identifikasi topik yang muncul dalam data. Sementara penelitian Alencia Hariyani et al. (2019) menerapkan multi-label classification dengan skema label yang telah ditentukan sebelumnya, memiliki keterbatasan menangkap aspek yang berkembang pada data.

Meskipun pendekatan ekstraksi aspek secara otomatis berbasis topic modeling BERTopic telah dilakukan oleh Alotaibi (2024), penelitian tersebut dilakukan pada konteks bahasa Arab sehingga penerapannya pada data berbahasa Indonesia, khususnya dalam domain pendidikan vokasi SMK masih belum diketahui. Perbedaan karakteristik bahasa, konteks lokal serta penggunaan bahasa informal dapat memengaruhi kualitas ekstraksi aspek dan interpretasi model.

Dataset dibangun menggunakan framework hybrid yang mengintegrasikan topic modeling berbasis BERTopic untuk menghasilkan kandidat topik secara otomatis dan human-in-the-loop untuk melakukan validasi serta membentuk kategori aspek yang lebih kontekstual dan sesuai dengan karakteristik data umpan balik siswa SMK.

Dataset ini berisi kumpulan umpan balik siswa SMK berbahasa Indonesia yang telah dianotasi berdasarkan aspek dan sentimen. Proses konstruksi dataset mempertimbangkan karakteristik bahasa siswa, termasuk variasi bahasa informal, penggunaan slang serta struktur kalimat tidak baku. Proses anotasi dan validasi dilakukan melalui mekanisme adjudikasi untuk memastikan konsistensi, reliabilitas serta kesesuaian label dengan konteks pengalaman belajar siswa SMK di Indonesia. Dataset dapat dimanfaatkan sebagai sumber data untuk pengembangan dan evaluasi model aspect-based sentiment analysis (ABSA) terhadap umpan balik siswa, evaluasi kualitas pembelajaran serta pengambilan keputusan berbasis data dalam peningkatan mutu pendidikan vokasi.

TINJAUAN PUSTAKA

Penelitian mengenai analisis sentimen (SA) dan aspect-based sentiment analysis (ABSA) dalam konteks pendidikan berkembang dalam beberapa tahun terakhir. Namun

sebagian besar studi berfokus pada pendidikan tinggi, sementara pemodelan aspek serta konteks pendidikan vokasi khusunya SMK di Indonesia masih relatif terbatas.

Dalam konteks pendidikan vokasi, Pi et al. (2025) menekankan pentingnya analisis sentimen berbasis kecerdasan buatan untuk menangkap kondisi emosional siswa selama pembelajaran praktik. Mereka mengusulkan pendekatan multimodal yang menggabungkan teks, audio dan video untuk merepresentasikan emosi siswa yang lebih komprehensif . Pendekatan tersebut tidak berfokus pada ekstraksi aspek maupun pembangunan dataset terstruktur.

Dalam konteks Indonesia, Nurfauzi et al. (2020) menerapkan analisis sentimen untuk mendukung pembelajaran kewirausahaan melalui Business Center. Hasil penelitian menunjukkan bahwa analisis sentimen dapat membantu siswa memahami dinamika pasar digital. Penelitian tersebut masih terbatas pada klasifikasi polaritas sentimen dan belum membahas analisis aspek maupun konstruksi dataset.

Pada konteks pendidikan tinggi, pendekatan berbasis leksikon dan machine learning masih banyak digunakan. Dervenis et al. (2024) membandingkan metode VADER dengan Random Forest dan Artificial Neural Network (ANN). Hasil penelitian menunjukkan bahwa model berbasis embedding dan ANN lebih mampu menangkap konteks dibandingkan pendekatan leksikon, terutama pada kelas sentimen netral yang bersifat ambigu. Sementara Dervenis (2024) menggunakan VADER untuk menganalisis kompetensi komunikasi dosen berdasarkan umpan balik mahasiswa, namun menemukan keterbatasan dalam menangani variasi bahasa dan konteks domain. Sedangkan penelitian Kim & Qin (2023) menunjukkan bahwa metode berbasis leksikon cenderung lemah dalam menangani sentimen negatif, sementara pendekatan klaster membutuhkan penentuan jumlah klaster secara manual sehingga berpotensi menghasilkan klaster kurang representatif. Penelitian-penelitian tersebut menunjukkan bahwa analisis sentimen di bidang pendidikan masih didominasi oleh klasifikasi polaritas tanpa eksplorasi struktur aspek yang lebih rinci.

Penelitian Alencia Hariyani et al. (2019) menggabungkan multi-label classification untuk mengekstraksi aspek layanan pendidikan berdasarkan label yang telah ditentukan sebelumnya. Meskipun efektif, pendekatan tersebut bergantung pada anotasi manual skala besar dan belum mampu menangkap aspek baru yang muncul secara alami. Sedangkan penelitian Alotaibi (2024) menerapkan BERTopic dengan embedding transformer multibahasa untuk analisis aspek pada survei evaluasi berbahasa Arab. Hasil penelitiannya

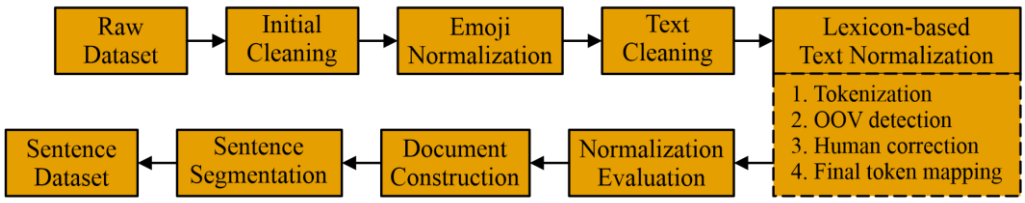
menunjukkan bahwa pendekatan berbasis topic modeling dapat membantu mengidentifikasi kandidat aspek secara otomatis. Namun, Penelitian tersebut dilakukan dalam konteks bahasa Arab, sehingga belum ada bukti bahwa pendekatan yang sama dapat bekerja optimal pada bahasa Indonesia yang memiliki struktur bahasa berbeda, khususnya pada konteks SMK.

Terdapat beberapa kesenjangan berdasarkan kajian literatur. Pertama, penelitian ABSA dalam konteks pendidikan masih didominasi oleh studi di pendidikan tinggi, sementara pendidikan vokasi seperti SMK di Indonesia masih jarang diteliti. Kedua, pendekatan masih bergantung pada label aspek yang telah ditentukan dengan penentuan jumlah klaster. Ketiga, pendekatan berbasis topic modeling seperti BERTopic sering menghasilkan topik yang ambigu. Keempat, belum tersedia dataset publik berbahasa Indonesia yang berisi anotasi aspek secara sistematis untuk konteks pendidikan vokasi.

METODOLOGI

Prapemrosesan

Prapemrosesan dilakukan untuk menstandarkan teks yang mengandung noise, slang, emoji dan singkatan. Tahapan yang dilakukan seperti disajikan pada Gambar 1 untuk menghasilkan data yang lebih konsisten tanpa kehilangan makna asli.



Gambar 1. Tahapan prapemrosesan teks.

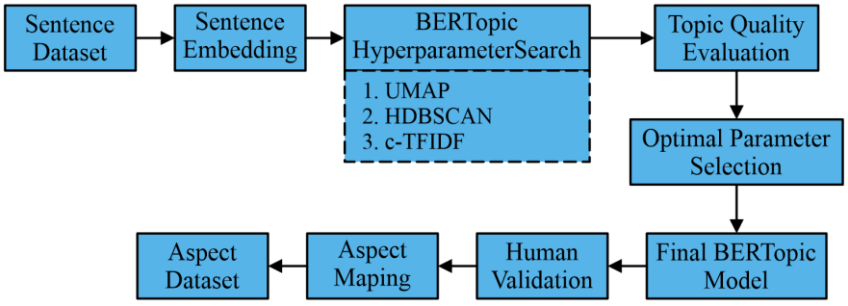
a. Normalisasi Emoji: Emoji diidentifikasi menggunakan pustaka Python berbasis Unicode dan dikonversi menjadi representasi teks deskriptif. Contohnya emoji "😁" dirubah menjadi "emoji_bahagia". Menurut Amalia et al. (2025) menggantikan emoji dengan teks deskripsi meningkatkan performa klasifikasi teks.
b. Pembersihan Teks: Dilakukan penghapusan noise seperti URL, mention dan simbol tidak relevan serta lowercase (Amalia et al., 2025; Hadiprakoso et al., 2023). Tanda baca, spasi berlebih serta karakter berulang dinormalisasi untuk meningkatkan konsistensi teks (Aliero et al., 2023).

c. Normalisasi Teks Berbasis Kamus: Tokenisasi dilakukan untuk mendeteksi OOV berbasis kamus KBBI https://github.com/damzaky/kumpulan-kata-bahasa-indonesia-KBBI/blob/master/legacy/indonesian-words.txt dan kamus slang (Wibowo et al., 2021). Kata tidak baku dipetakan secara manual ke bentuk standar dan digabung menjadi kamus normalisasi akhir, kemudian diterapkan menggunakan regex ke dalam struktur dataset berbasis dokumen (Saputro & Hermawan, 2021). Dataset disusun dengan id unik untuk menjaga keterlacakan (Nowakowska, 2025).

d. Segmentasi Kalimat: Dokumen dipisahkan menjadi kalimat menggunakan pendekatan rule-based (Petrus et al., 2023). Proses mencakup normalisasi delimiter, pemisahan berdasarkan tanda baca dan konjungsi (El Khair, 2022) serta filtering kalimat tidak lengkap (Alfian et al., 2025; Fitria, 2023; Wibowo et al., 2020). Kalimat terlalu panjang di-chunk, sedangkan minimal dua kata diterapkan untuk menjaga validitas (Petrus et al., 2023).

Ekstraksi Aspek Tanpa Supervisi

Penelitian ini menggunakan BERTopic untuk ekstraksi aspek secara unsupervised dari kalimat (Grootendorst, 2022; Wahyuni et al., 2025). Model ini mengombinasikan embedding, reduksi dimensi, clustering dan c-TF-IDF untuk menghasilkan topik yang koheren pada teks informal (Khairunnisa & Siregar*, 2025; Medvecki et al., 2024). Proses keseluruhan tahapan diilustrasikan pada Gambar 2.



Gambar 2. Tahapan Ekstraksi Aspek.

a. Representasi Teks (Text Embedding): Kalimat direpresentasikan menggunakan paraphrase-multilingual-MiniLM-L12-v2, yang memetakan teks ke vektor 384-dimensi dan mendukung Bahasa Indonesia (Tat & Aydogan, 2024) untuk menangkap kesamaan semantik lintas bahasa.

b. Reduksi Dimensi (Dimensionality Reduction): UMAP (*Uniform Manifold Approximation and Projection*) digunakan untuk mereduksi dimensi embedding, mengatasi curse of dimensionality dan meningkatkan kualitas pengelompokan (Hill et al., 2025; Wahyuni et al., 2025).

c. Klasterisasi: HDBSCAN (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*) digunakan untuk membentuk klaster topik pada ruang berdimensi rendah, termasuk deteksi noise secara otomatis (Tat & Aydogan, 2024).

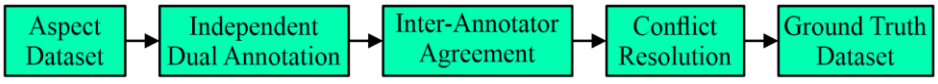
d. Representasi Topik: Representasi topik diperoleh menggunakan metode c-TF-IDF (*class-based Term Frequency–Inverse Document Frequency*) untuk mengekstraksi kata kunci representatif pada setiap klaster (Grootendorst, 2022), sebagaimana pada Persamaan (1).

$$W_{t,c} = tf_{t,c} \cdot \log\left(1 + \frac{A}{tf_t}\right) \quad (1)$$

e. Optimasi Hiperparameter: dilakukan menggunakan metode grid search sebanyak 96 kombinasi pada parameter UMAP dan HDBSCAN. Evaluasi menggunakan coherence, topic diversity dan noise ratio untuk menyeimbangkan kualitas, keberagaman dan proporsi noise (Medvecki et al., 2024). Parameter optimal yang diperoleh kemudian divalidasi secara kualitatif melalui pemeriksaan terhadap kata kunci representatif dan contoh kalimat pada setiap topik yang dihasilkan.

Konstruksi Ground Truth

Dataset ground truth dibangun melalui anotasi manual komentar siswa ke dalam tiga kelas sentimen: positif, negatif, dan netral. Anotasi dilakukan oleh dua annotator dengan pedoman label untuk mengurangi subjektivitas (Chamid et al., 2024). Anotator pertama menilai berdasarkan konteks kebijakan sekolah, sedangkan anotator kedua fokus pada aspek linguistik dan makna implisit. Alur proses pembentukan dataset ground truth disajikan pada Gambar 3.



Gambar 3. Contoh gambar boxplot sebaran data pengabdian.

Konsistensi antar anotator diukur menggunakan Cohen's Kappa (Persamaan 2) dengan mempertimbangkan kesepakatan di luar peluang acak. Interpretasi nilai kesepakatan seperti disajikan pada Tabel 1. Perbedaan label diselesaikan melalui proses

adjudikasi berbasis diskusi terfasilitasi hingga diperoleh label akhir yang disepakati (Braun, 2024).

$$K = \frac{Pr(a) - Pr(e)}{1 - \Pr(e)} \quad (2)$$

Tabel 1. Interpretasi Cohen's Kappa

Cohen's Kappa	Interpretasi
< 0,00	Sangat Buruk
0,00 – 0,20	Rendah
0,21 – 0,40	Cukup
0,41 – 0,60	Sedang
0,61 – 0,80	Baik
0,81 – 1,00	Hampir Sempurna

HASIL DAN PEMBAHASAN

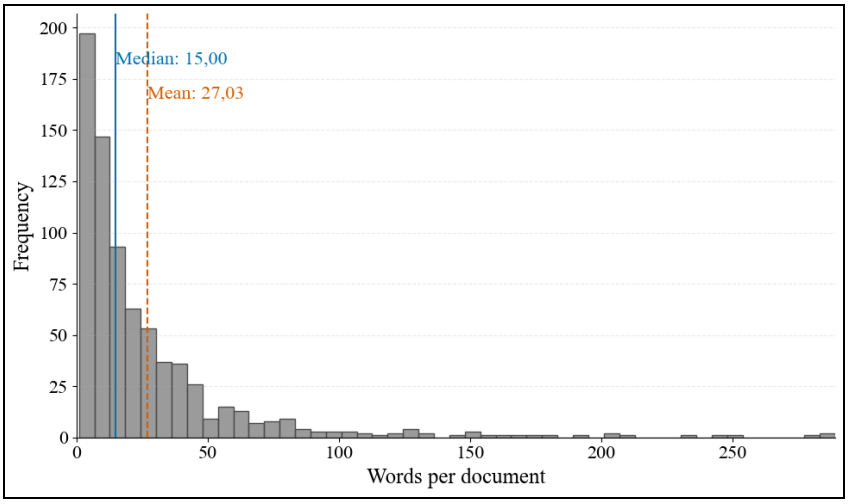
Karakteristik Dataset

Dataset awal terdiri atas 756 respons terbuka yang dikumpulkan dari siswa salah satu Sekolah Menengah Kejuruan (SMK) di Indonesia, dengan seluruh respons lengkap tanpa missing values. Setelah melalui proses prapemrosesan, diperoleh 717 dokumen dan 1.317 unit kalimat. Tabel 2 menyajikan ringkasan statistik dataset.

Tabel 2. Statistik dataset

Komponen	Nilai
Jumlah responden	756
Total dokumen	717
Dokumen duplikat	44
Rata-rata panjang dokumen	29,53
Panjang dokumen minimum	2
Panjang dokumen maksimum	301
Total kalimat	1.317
Rata-rata panjang kalimat	16,05
Panjang kalimat minimum	2
Panjang kalimat maksimum	50

Rata-rata panjang dokumen awal adalah 27,03 kata dan median 15 kata, dengan distribusi right-skewed seperti terlihat pada Gambar 4. Hal ini menunjukkan bahwa sebagian besar respons relatif pendek, meskipun terdapat variasi panjang antar dokumen, dengan beberapa respons yang jauh lebih panjang dibandingkan mayoritas data.

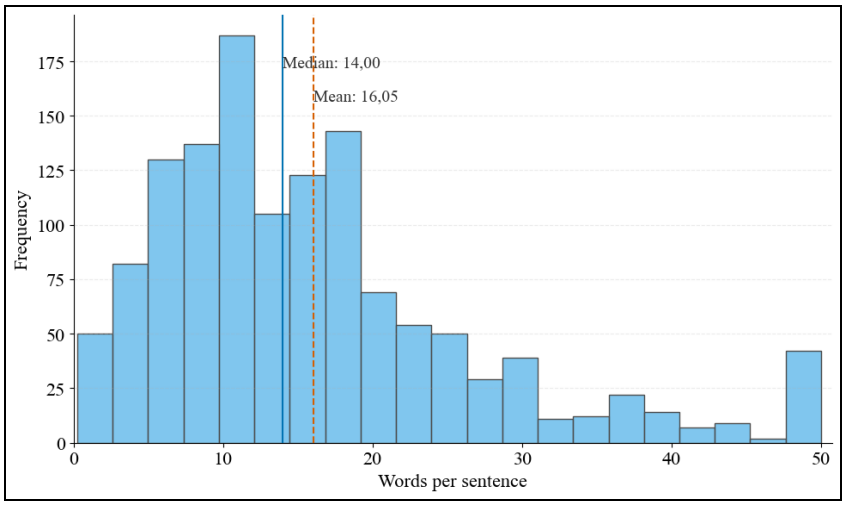


Gambar 4. Distribusi panjang dokumen pada teks mentah.

Dataset juga mengandung variasi linguistik yang cukup tinggi, meliputi variasi ejaan, bahasa gaul, singkatan, emoji, dan ungkapan informal. Sebelum proses normalisasi, teridentifikasi 85 kemunculan emoji (20 jenis emoji unik) serta 1.522 token out-of-vocabulary (OOV). Emoji dikonversi menjadi token teks deskriptif, sedangkan bahasa gaul dan variasi ejaan dinormalisasi menggunakan kombinasi leksikon bahasa Indonesia dan kamus bahasa gaul yang disusun secara manual.

Setelah prapemrosesan, dilakukan segmentasi kalimat berbasis tanda baca dan konjungsi kontras. Proses ini mengurangi variasi kosakata dan menghasilkan 1.317 kalimat dengan rata-rata panjang 16,05 kata. Distribusi panjang kalimat seperti terlihat pada Gambar 5 menjadi lebih homogen dibandingkan distribusi dokumen awal (Gambar 4), sehingga lebih sesuai untuk representasi embedding. Temuan ini menunjukkan bahwa umpan balik siswa SMK memiliki karakteristik noisy, informal dan multiaspek.

Prosiding Seminar Nasional Universitas Ma Chung 2026
(Sistem Informasi dan Informatika)
pp. 150-164
E-ISSN: 2808-7771



Gambar 5. Distribusi panjang kalimat.

Ekstraksi Aspek Menggunakan BERTopic

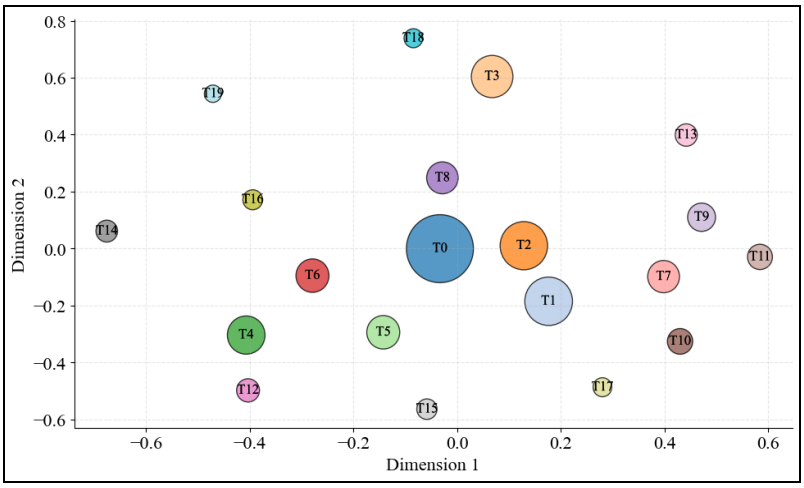
Proses ekstraksi aspek dilakukan menggunakan BERTopic (*Bidirectional Encoder Representations from Transformers Topic Modeling*) dengan memanfaatkan *multilingual transformer embedding* serta optimasi parameter UMAP dan HDBSCAN. Optimasi hiperparameter pada 96 kombinasi konfigurasi UMAP–HDBSCAN menghasilkan 20 topik laten, skor topic coherence 0,609, topic diversity 0,974, dan noise ratio 15,4%. Nilai diversity yang tinggi menunjukkan minimnya tumpang tindih antartopik, sedangkan nilai coherence mengindikasikan bahwa topik yang terbentuk memiliki makna semantik yang baik meskipun didominasi umpan balik siswa yang bersifat informal. Sebanyak 15,4% kalimat dikelompokkan sebagai noise, menunjukkan bahwa HDBSCAN tidak memaksakan kalimat yang ambigu atau kurang berkaitan ke dalam topik yang terbentuk.

Tabel 3. Konfigurasi parameter optimal

Komponen	Parameter	Nilai
UMAP	n_neighbors	10
	n_components	10
	min_dist	0.0
HDBSCAN	min_cluster_size	15
	min_samples	10

Visualisasi *intertopic distance map* pada Gambar 6 menunjukkan bahwa sebagian besar topik membentuk klaster yang terpisah dengan baik, sementara beberapa topik

memiliki kedekatan secara semantik. Topik dominan terlihat pada T0, T1, dan T2 yang ditunjukkan oleh ukuran lingkaran yang lebih besar dibandingkan dengan topik lainnya.



Gambar 6. Distribusi panjang kalimat.

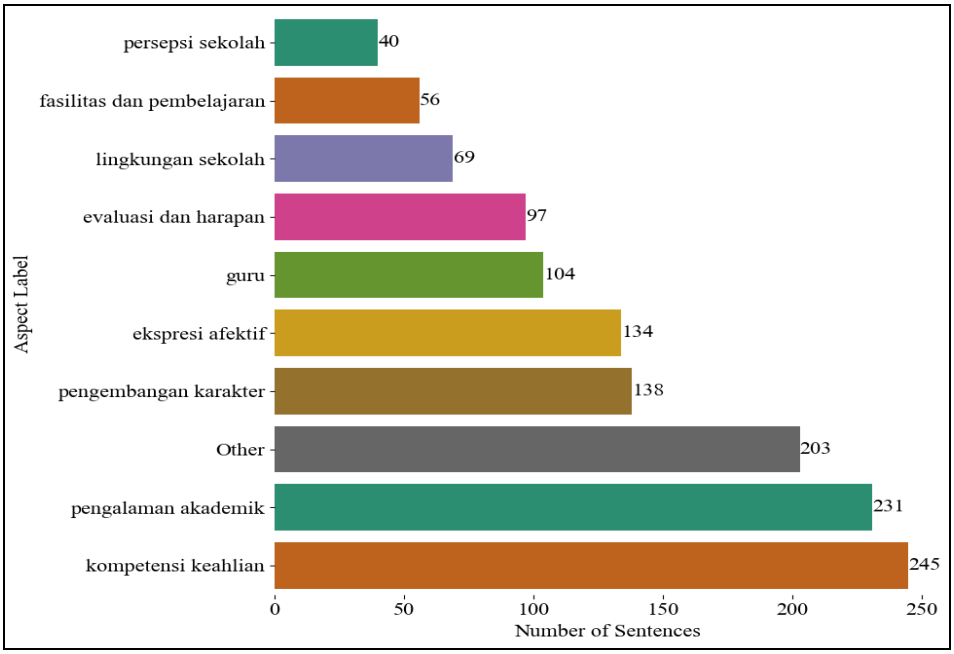
Topik yang dihasilkan oleh BERTopic tidak secara langsung digunakan sebagai aspek akhir. Peneliti terlebih dahulu menelaah kata kunci yang representatif serta contoh kalimat dari masing-masing topik, kemudian menggabungkan topik-topik yang memiliki kesamaan makna ke dalam taksonomi aspek akhir. Proses validasi yang dilakukan secara iteratif ini bertujuan memastikan bahwa kategori aspek yang dihasilkan benar-benar merepresentasikan konteks pendidikan di SMK, bukan sekadar hasil pengelompokan topik berdasarkan pola statistik. Hasil proses tersebut disajikan pada Tabel 4.

Tabel 4. Hasil interpretasi topik dan pemetaan aspek

ID Topik	Label Topik	Label Aspek
T0	Pengalaman belajar di sekolah	Pengalaman akademik
T1	Disiplin dan pengembangan karakter	Pengembangan karakter
T2	Ekspresi emosional siswa	Ekspresi afektif
T3	Desain komunikasi visual	Kompetensi keahlian
T4	Praktik kerja lapangan	Kompetensi keahlian
T5	Fasilitas praktik dan metode pembelajaran	Fasilitas dan pembelajaran
T6	Kualitas pengajaran guru	Guru
T7	Keluhan fasilitas dan tata kelola	Evaluasi dan harapan
T8	Lingkungan sekolah	Lingkungan sekolah

ID Topik	Label Topik	Label Aspek
T9	Persepsi terhadap sekolah	Persepsi sekolah
T10	Karakter guru	Guru
T11	Tata busana	Kompetensi keahlian
T12	Akuntansi	Kompetensi keahlian
T13	Kritik dan saran pengembangan	Evaluasi dan harapan
T14	Teknik kendaraan ringan	Kompetensi keahlian
T15	Kepercayaan diri dan kemandirian	Pengembangan karakter
T16	Harapan peningkatan	Evaluasi dan harapan
T17	Interaksi sosial	Lingkungan sekolah
T18	Ekspresi informal siswa	Ekspresi afektif
T19	Kejelasan penyampaian materi	Guru

20 topik tersebut dipetakan menjadi sembilan aspek seperti disajikan pada Gambar 6, yaitu pengalaman akademik, kompetensi keahlian, guru, fasilitas dan pembelajaran, pengembangan karakter, lingkungan sekolah, ekspresi afektif, persepsi sekolah serta evaluasi dan harapan. Aspek kompetensi keahlian menjadi aspek paling dominan dengan 245 kalimat, diikuti pengalaman akademik sebanyak 231 kalimat.

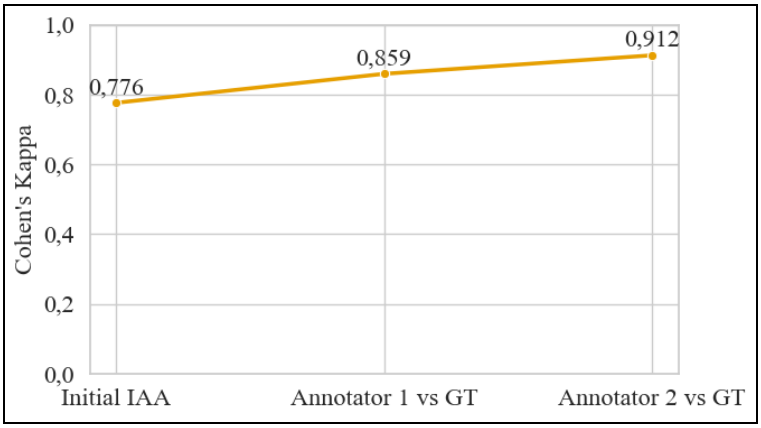


Gambar 7. Distribusi jumlah kalimat berdasarkan kategori aspek.

Hasil ini menunjukkan bahwa siswa lebih banyak membahas pengalaman belajar praktis, praktik kerja lapangan dan pengembangan keterampilan vokasional dibandingkan aspek administratif sekolah. Temuan tersebut konsisten dengan karakteristik pendidikan kejuruan yang berorientasi pada pengembangan kompetensi kerja.

Penyusunan Data Acuan (Ground Truth) dan Kualitas Anotasi

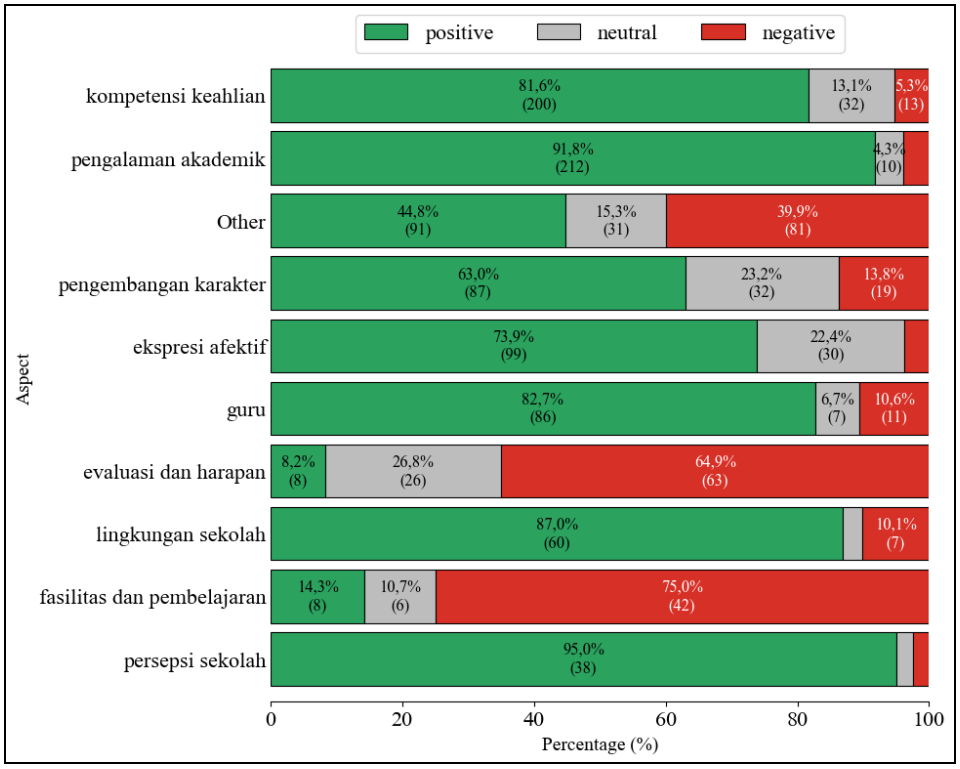
Untuk membangun dataset berlabel, dilakukan anotasi oleh dua anotator independen terhadap 1.317 kalimat menggunakan tiga kelas sentimen, yaitu positif, netral dan negatif. Kesepakatan awal menghasilkan Cohen's Kappa sebesar 0,776, termasuk kategori Baik. Namun sebanyak 149 kalimat (11,31%) mengalami konflik anotasi. Setelah adjudikasi, nilai Cohen's Kappa meningkat menjadi 0,859 dan 0,912 (Gambar7), kategori Hampir Sempurna. Peningkatan ini menunjukkan bahwa sebagian besar konflik berasal dari kalimat ambigu, sentimen implisit, atau pernyataan yang mengandung polaritas campuran. Dengan demikian, kombinasi dual annotation dan adjudication berhasil menghasilkan ground truth yang reliabel untuk eksperimen klasifikasi sentimen.



Gambar 8. Distribusi jumlah kalimat berdasarkan kategori aspek.

Hasil Analisis Sentimen Berbasis Aspek

Aspek pengalaman akademik, kompetensi keahlian, guru, lingkungan sekolah dan persepsi sekolah didominasi sentimen positif. Sebaliknya, aspek evaluasi dan harapan serta fasilitas dan pembelajaran menunjukkan proporsi sentimen negatif yang lebih tinggi dibandingkan aspek lainnya seperti terlihat pada Gambar 8. Temuan ini mengindikasikan bahwa siswa memiliki persepsi positif terhadap pengalaman belajar, kompetensi keahlian yang diperoleh serta kualitas interaksi dengan guru. Namun fasilitas praktik dan beberapa aspek pembelajaran masih menjadi sumber utama kritik dan harapan perbaikan.



Gambar 9. Distribusi jumlah kalimat berdasarkan kategori aspek.

KESIMPULAN

Hasil penelitian menunjukkan bahwa karakteristik umpan balik siswa SMK bersifat multiaspek, informal dan kaya informasi semantik. BERTopic menghasilkan 20 topik yang cukup koheren (0,609) dan beragam (0,974), meskipun masih terdapat noise sebesar 15,4%. Topik tersebut divalidasi dan dipetakan menjadi sembilan aspek yang lebih kontekstual dengan pendidikan vokasi SMK. Proses anotasi sentimen menggunakan dua anotator independen yang dilanjutkan dengan adjudikasi meningkatkan reliabilitas dataset, ditunjukkan oleh nilai Cohen's Kappa yang mencapai kategori Hampir Sempurna (hingga 0,912). Aspek Pengalaman Akademik, Kompetensi Keahlian, Guru dan Lingkungan Sekolah didominasi sentimen positif, sedangkan aspek Fasilitas dan Pembelajaran serta Evaluasi dan Harapan menjadi sumber sentimen negatif dan masukan perbaikan. Pendekatan ini menghasilkan dataset yang terstruktur, representatif dan kontekstual untuk pendidikan vokasi, khususnya SMK. Kombinasi topic modeling dan validasi manusia menghasilkan dataset berkualitas serta memberikan wawasan yang lebih rinci.

DAFTAR PUSTAKA

Alencia Hariyani, C., Nizar Hidayanto, A., Fitriah, N., Abidin, Z., & Wati, T. (2019). Mining Student Feedback to Improve the Quality of Higher Education through Multi Label Classification, Sentiment Analysis, and Trend Topic. *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 359–364. https://doi.org/10.1109/ICITISEE48480.2019.9003818

Alfian, M., Yuhana, U. L., Siahaan, D., & Munazharoh, H. (2025). Annotation Error Detection and Correction for Indonesian POS Tagging Corpus. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, *16*(1), 41–52. https://doi.org/10.24843/LKJITI.2025.v16.i01.p04

Aliero, A. A., Adebayo, B. S., Aliyu, H. O., Tafida, A. G., Kangiwa, B. U., & Dankolo, N. M. (2023). Systematic Review on Text Normalization Techniques and Its Approach to Non-Standard Words. *International Journal of Computer Applications*, *185*(33), 44–55. https://doi.org/10.5120/ijca2023923106

Alotaibi, R. (2024). Aspect-Based Sentiment Analysis of Open-Ended Responses in Student Course Evaluation Surveys. *Thermal Science*, *28*(6 Part B), 5037–5047. https://doi.org/10.2298/TSCI2406037A

Amalia, J., Tambunan, S. R., Purba, S. E. M., & Simanjuntak, W. V. (2025). Enhancing Hate Speech Detection: Leveraging Emoji Preprocessing with Bi-LSTM Model. *Journal of Information Systems and Informatics*, *7*(2), 1799–1813. https://doi.org/10.51519/journalisi.v7i2.1147

Braun, D. (2024). I Beg to Differ: How Disagreement Is Handled in the Annotation of Legal Machine Learning Datasets. *Artificial Intelligence and Law*, *32*(3), 839–862. https://doi.org/10.1007/s10506-023-09369-4

Chamid, A. A., Widowati, & Kusumaningrum, R. (2024). Labeling Consistency Test of Multi-Label Data for Aspect and Sentiment Classification Using the Cohen Kappa Method. *Ingénierie Des Systèmes d Information*, *29*(1), 161–167. https://doi.org/10.18280/isi.290118

Dervenis, C. (2024). Assessing Teacher Competencies in Higher Education: A Sentiment Analysis of Student Feedback. *International Journal of Information and Education Technology*, *14*(4), 533–541. https://doi.org/10.18178/ijiet.2024.14.4.2074

Dervenis, C., Kanakis, G., & Fitsilis, P. (2024). Sentiment Analysis of Student Feedback: A Comparative Study Employing Lexicon and Machine Learning Techniques. *Studies in Educational Evaluation*, *83*, 101406. https://doi.org/10.1016/j.stueduc.2024.101406

El Khair, H. (2022). Contrastive Conjunctions in Japanese and Indonesian. *Proceeding of International Conference on Business, Economics, Social Sciences, and Humanities*, *3*, 392–400. https://doi.org/10.34010/icobest.v3i.163

Fargues, M., Kadry, S., Lawal, I. A., Yassine, S., & Rauf, H. T. (2023). Automated Analysis of Open-Ended Students' Feedback Using Sentiment, Emotion, and Cognition Classifications. *Applied Sciences*, *13*(4), 2061. https://doi.org/10.3390/app13042061

Fitria, T. N. (2023). Performance of Google Translate, Microsoft Translator, and DeepL Translator: Error Analysis of Translation Results. *Al-Lisan*, *8*(2), 115–138. https://doi.org/10.30603/al.v8i2.3442

Grootendorst, M. (2022). *BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure*. http://arxiv.org/abs/2203.05794

Hadiprakoso, R. B., Setiawan, H., Yasa, R. N., & Girinoto. (2023). *Text Preprocessing for Optimal Accuracy in Indonesian Sentiment Analysis Using a Deep Learning Model With Word Embedding*. 020050. https://doi.org/10.1063/5.0126116

Hill, A., Goo, K., & Agarwal, P. (2025). Recommending the Right Academic Programs: An Interest Mining Approach Using BERTopic. *Data Mining and Knowledge Discovery*, *39*(3), 20. https://doi.org/10.1007/s10618-024-01087-y

Khairunnisa, R., & Siregar*, J. H. (2025). Social Network and Sentiment Analysis for Enhancing Social CRM in Indonesian Educational Technology Platforms. *Jurnal Riset Informatika*, *7*(4), 258–269. https://doi.org/10.34288/jri.v7i4.383

Kim, H., & Qin, G. (2023). Summarizing Students' Free Responses for an Introductory Algebra-Based Physics Course Survey Using Cluster and Sentiment Analysis. *IEEE Access*, *11*, 89052–89066. https://doi.org/10.1109/ACCESS.2023.3305260

Medvecki, D., Bašaragin, B., Ljajić, A., & Milošević, N. (2024). *Multilingual Transformer and BERTopic for Short Text Topic Modeling: The Case of Serbian*. https://doi.org/10.1007/978-3-031-50755-7_16

Nowakowska, M. (2025). A Comprehensive Approach to Preprocessing Data for Bibliometric Analysis. *Scientometrics*, *130*(9), 5191–5225. https://doi.org/10.1007/s11192-025-05415-x

Nurfauzi, Y., Suyanto, S., Sukidjo, S., Munsyi, M., & Ahdhianto, E. (2020). The Role of Business Center Using Sentiment Analysis to Foster Entrepreneurial Spirit in Vocational High School. *Universal Journal of Educational Research*, *8*(11), 5151–5157. https://doi.org/10.13189/ujer.2020.081114

Petrus, J., Ermatita, E., Sukemi, S., & Erwin, E. (2023). An AdapTabel Sentence Segmentation Based on Indonesian Rules. *IAES International Journal of Artificial Intelligence (IJ-AI)*, *12*(3), 1491. https://doi.org/10.11591/ijai.v12.i3.pp1491-1499

Pi, X., Chen, Y., & Liu, Q. (2025). Research on the Application of AI-based Sentiment Analysis in Vocational Education. *Proceedings of the 2025 International Conference on Digital Education and Information Technology*, 130–135. https://doi.org/10.1145/3732299.3732325

Saputro, T. H., & Hermawan, A. (2021). The Accuracy Improvement of Text Mining Classification on Hospital Review through The Alteration in The Preprocessing Stage. *International Journal of Computer and Information Technology(2279-0764)*, *10*(4). https://doi.org/10.24203/ijcit.v10i4.138

Tat, O., & Aydogan, I. (2024). Discovering Hidden Patterns: Applying Topic Modeling in Qualitative Research. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, *15*(3), 247–259. https://doi.org/10.21031/epod.1539694

Wahyuni, W., Lestari, T. P., Apriliana, M., & Gumelta, R. (2025). Implementation of BERTopic for Topic Modeling Analysis of the Free Nutritious Meal Program Based on YouTube Comments. *Journal of Applied Informatics and Computing*, *9*(4), 1964–1971. https://doi.org/10.30871/jaic.v9i4.9754

Wibowo, H. A., Nityasya, M. N., Akyürek, A. F., Fitriany, S., Aji, A. F., Prasojo, R. E., & Wijaya, D. T. (2021). IndoCollex: A Testbed for Morphological Transformation of Indonesian Word Colloquialism. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 3170–3183. https://doi.org/10.18653/v1/2021.findings-acl.280

Wibowo, H. A., Prawiro, T. A., Ihsan, M., Aji, A. F., Prasojo, R. E., Mahendra, R., & Fitriany, S. (2020). Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language with Iterative Forward Translation. *2020 International Conference on Asian Language Processing (IALP)*, 310–315. https://doi.org/10.1109/IALP51396.2020.9310459