

BAB II

LANDASAN TEORI

2.1 Penelitian Terdahulu

Penelitian ini mengacu pada sejumlah studi sebelumnya sebagai dasar konseptual dan metodologis. Ringkasan penelitian-penelitian tersebut disajikan pada Tabel 2.1 sebagai landasan dalam pengembangan penelitian ini.

Tabel 2. 1 Perbandingan Penelitian Terdahulu

Penelitian Pertama	
Judul	Pre-trained language model for code-mixed text in Indonesian, Javanese, and English using transformer. [26]
Penulis	Ahmad Fathan Hidayatullah, Rosyzie Anna Apong, dkk,
Nama Journal	Social Network Analysis and Mining, Volume 15, 2025.
Objek Penelitian	<i>Teks code-mixed</i> (Bahasa Indonesia, Jawa, dan Inggris) pada media sosial
Metode	<ol style="list-style-type: none"> 1. <i>Pre-trained language model</i> berbasis transformer, Model: BERT, RoBERTa, IndoBERT, IndoBERTweet 2. Pendekatan training from scratch dan transfer learning. 3. Pre-training dengan <i>Masked Language Modeling (MLM)</i>.
Sumber Data	Data dari platform X, ±40 juta post Januari 2022 – Januari 2023.
<i>Findings</i>	<ol style="list-style-type: none"> 1. Transfer learning lebih efektif dan efisien dibanding training dari mentah. 2. Model IndoJavE outperform model multilingual dan monolingual. 3. Model khusus code-mixed sangat penting untuk bahasa low-resource.
Penelitian Dua	
Judul	A Multi-feature Fusion Approach Based on Domain Adaptive Pretraining for Aspect-based Sentiment Analysis [27]
Penulis	Yinglong Ma, Ming He, Yunhe Pang, dkk.
Nama Journal	Application of Soft Computing, Volume 28, 2024
Objek Penelitian	<i>Aspect-based sentiment classification (ASC)</i> dengan fokus meminimalisir informasi bising dari pohon dependensi sintaksis.
Metode	Menggunakan teknik <i>Multi-task Learning (MTL)</i> untuk <i>domain adaptive pretraining</i> yang menggabungkan <i>Biaffine Attention Model (BAM)</i> dan <i>Mask Language Model (MLM)</i> .
Sumber Data	Dataset <i>English Penn Treebank (PTB 3.3.0-SD)</i> untuk melatih <i>encoder/decoder</i> . Empat <i>benchmark dataset</i> ulasan dari SemEval: Laptop14, Rest14, Rest15, dan Rest16.
<i>Findings</i>	Model yang diusulkan (LD2G) berhasil mengungguli berbagai metode baseline <i>state-of-the-art</i> untuk ASC dalam hal Akurasi dan

	F1-score. Teknik <i>domain adaptive pretraining</i> terbukti krusial dan efisien dalam menghasilkan parsing dependensi yang mengurangi noise.
Penelitian Tiga	
Judul	Multi-Label Aspect-Sentiment Classification on Indonesian Cosmetic Product Reviews with IndoBERT Model. [25]
Penulis	Ng Chin Mei, Sabrina Tiun, Gita Sastria.
Nama Journal	International Journal of Advanced Computer Science Applications, Vol. 15, No. 11, 2024.
Objek Penelitian	Ulasan produk kecantikan berbahasa Indonesia dari platform Female Daily.
Metode	Exploratory Data Analysis (EDA) dengan uji Chi-square, preprocessing, word embedding menggunakan varian IndoBERT dan Word2Vec, klasifikasi multi-label menggunakan pendekatan problem transformation (Binary Relevance, Classifier Chain, Label Powerset, RAKEL D) dengan algoritma Gaussian Naive Bayes, SVM, Random Forest, dan Linear SGD, serta penggunaan IndoBERT sebagai model end-to-end. Evaluasi menggunakan akurasi per label, akurasi, hamming loss, dan micro f1-score.
Sumber Data	Data sekunder berupa 3.960 ulasan pelanggan tentang produk kecantikan berbahasa Indonesia yang dikumpulkan dari situs web Female Daily.
<i>Findings</i>	<ol style="list-style-type: none"> 1. Multi-label lebih efektif dibanding single-label 2. IndoBERT outperform Word2Vec dalam representasi teks 3. Hasil terbaik: End-to-end IndoBERT: 86.98% accuracy
Penelitian Empat	
Judul	Improving Cross-Domain Aspect-Based Sentiment Analysis using Bert-BiLSTM Model and Dual Attention Mechanism. [24]
Penulis	Yadi Xu, Noor Farizah Ibrahim
Nama Journal	Advances in Artificial Intelligence and Machine Learning, Volume 4, Issue 3, 2024.
Objek Penelitian	Aspect-Based Sentiment Analysis lintas domain.
Metode	Hybrid deep learning model seperti BERT (word embedding), BiLSTM (<i>sequence modeling</i>), <i>Dual Attention Mechanism</i> , Input dipisah menjadi konteks sebelum, aspek, dan setelah aspek
Sumber Data	Dataset <i>multi-domain</i> untuk evaluasi <i>cross-domain</i> .
<i>Findings</i>	Model BB-DAM menghasilkan performa terbaik dibandingkan model baseline seperti AHF, BERT, dan AD-SAL. Rata-rata Micro-F1 BB-DAM mencapai 39,41%, lebih tinggi dibanding AHF (37,67%), BERT (35,42%), dan AD-SAL (33,71%). Hasil ini menunjukkan bahwa kombinasi BERT, BiLSTM, dan Dual Attention efektif dalam mengekstraksi informasi kontekstual dan hubungan antara aspek dan konteks pada skenario lintas domain
Penelitian Lima	
Judul	Cross-Domain Aspect-Based Sentiment Classification with a Pre-Training and Fine-Tuning Strategy for Low-Resource Domains. [22]

Penulis	Chuanjun Zhao, Meiling Wu, dkk.
Nama Journal	ACM Transactions on Asian and Low-Resource Language Information Processing, Volume 23, Issue 4, 2024.
Objek Penelitian	<i>Aspect-Based Sentiment Classification</i> lintas, khususnya pada kondisi low-resource domain.
Metode	Pendekatan: Transfer learning (pre-training + fine-tuning). Base model: AEGCN (Attention-based Encoding Graph Convolutional Network). Strategi utama: A. Pre-training di source domain (data besar). B. Fine-tuning di target domain (data kecil).
Sumber Data	Dataset benchmark ABSA: 1. Laptop reviews (SemEval 2014) 2. Restaurant reviews (SemEval 2014, 2015, 2016)
<i>Findings</i>	1. Model dengan pre-training + fine-tuning outperform semua baseline dengan hasil Micro-F1 lebih tinggi. 2. Strategi terbaik: unfreezing first n layers saat fine-tuning. 3. Transfer learning efektif untuk mengatasi keterbatasan data. 4. Fine-tuning membantu model menangkap domain-specific features, sedangkan tanpa fine-tuning model bias terhadap domain sumber.
Penelitian Enam	
Judul	Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback. [28]
Penulis	Ahmad Jazuli, Widowati, dan Retno Kusumaningrum.
Nama Journal	Applied Sciences, Volume 15, Tahun 2025.
Objek Penelitian	Ulasan atau umpan balik dari mahasiswa mengenai evaluasi proses pembelajaran di perguruan tinggi Indonesia, yang berfokus pada aspek kualitas pengajaran, kurikulum, infrastruktur, dan layanan.
Metode	Pengumpulan data dengan scraping, pelabelan data gabungan secara manual oleh ahli bahasa dan otomatis oleh mesin, pra-pemrosesan teks seperti normalisasi dan tokenisasi, pembagian data latih dan uji, klasifikasi sentimen berbasis aspek menggunakan model IndoBERT, pencarian parameter paling optimal atau hyperparameter tuning, serta evaluasi menggunakan akurasi, presisi, recall, dan f1-score.
Sumber Data	Data primer berupa 10.000 ulasan mahasiswa berbahasa Indonesia yang dikumpulkan dari berbagai portal akademik kampus yang ada di Indonesia.
<i>Findings</i>	Proses klasifikasi diawali dengan pemahaman konteks bahasa pada teks ulasan menggunakan model bawaan IndoBERT agar sistem bisa mengenali struktur bahasa Indonesia dengan baik sejak awal. Selanjutnya, model melakukan ekstraksi aspek dan pengelompokan sentimen yang kinerjanya dioptimalkan melalui tahapan penyesuaian parameter, seperti mengatur ukuran kelompok data dan tingkat kecepatan pembelajaran mesinnya. Pada tahap akhir, model berhasil mengklasifikasikan sentimen untuk setiap aspek dengan sangat baik, mencapai tingkat akurasi 97,9% untuk klasifikasi sentimen dan 97,3% untuk ekstraksi aspek,

	yang membuktikan bahwa penyesuaian parameter secara sistematis sangat efektif untuk meningkatkan performa model
Penelitian Tujuh	
Judul	Multilabel Aspect-Based Sentiment Analysis with Diverse Embedding Techniques and Finetuned Transformers for Label Detection. [29]
Penulis	Yusril Falih Izzaddien , Abdullah Faqih Septiyanto, dkk.
Nama Journal	Beyond Technology Summit on Informatics International Conference (BTS-12C), 2024.
Objek Penelitian	Deteksi ujaran kebencian di media sosial dengan klasifikasi multilabel. Terdapat delapan kategori yang diteliti: <i>violence, directed vs generalized, gender, race, national origin, disability, religion, dan sexual orientation.</i>
Metode	Membandingkan lima pendekatan, yaitu Word2Vec + <i>cosine similarity</i> (AC1), GloVe + <i>cosine similarity</i> (AC2), <i>BERT embedding + cosine similarity</i> (AC3), <i>Fine-tuned BERT</i> (AC4), dan <i>Pretrained RoBERTa</i> dengan <i>fine-tuning</i> (AC5).
Sumber Data	433 data dengan delapan label aspek. Data bersumber dari video YouTube dan <i>thread</i> Reddit di Hugging Face.
Findings	Metode AC5 (RoBERTa + <i>Fine-Tuning</i>) menghasilkan performa terbaik dengan rata-rata <i>F1 score</i> 0.674, rata-rata presisi 0.862, dan rata-rata akurasi 0.890
Penelitian Delapan	
Judul	Cross-domain aspect-based sentiment analysis using domain adversarial training. [23]
Penulis	Joris Knoester, Flavius Frasinca, dkk.
Nama Journal	World Wide Web Journal, Volume 26, Tahun 2023
Objek Penelitian	<ol style="list-style-type: none"> 1. Task: Aspect-Based Sentiment Classification (ABSC) 2. Fokus: Cross-domain sentiment classification 3. Masalah utama: <ol style="list-style-type: none"> A. Keterbatasan data berlabel di target domain B. Domain shift (perbedaan distribusi antar domain) C. Membangun model yang bisa bekerja pada target domain tanpa labeled data
Metode	Transfer Learning + Domain Adversarial Training (DAT)
Sumber Data	<ol style="list-style-type: none"> 1. SemEval 2014 = Domain restaurant & laptop. 2. ALT (Amazon / LibraryThing) 2019 = Domain book.
Findings	Akurasi bervariasi 35% – 75%+ tergantung domain

Berdasarkan tabel 2.1 berbagai pendekatan telah digunakan untuk meningkatkan kinerja analisis sentimen dan klasifikasi aspek. Penelitian [26] menerapkan *pre-trained language model* dengan pendekatan *Masked Language Modeling* (MLM) pada data *code-mixed* dalam kondisi *low-resource*, yang menunjukkan bahwa *transfer learning* lebih efektif dan efisien dibandingkan *training* dari awal. Penelitian [25] mengimplementasikan IndoBERT secara *end-to-end* untuk klasifikasi *multi-label* pada ulasan kosmetik berbahasa Indonesia dan

menghasilkan akurasi sebesar 86,98%, tetapi belum menerapkan pendekatan *cross-domain*. Selanjutnya, penelitian [24] menggunakan metode *cross-domain* berbasis model hibrida BB-DAM (BERT, BiLSTM, *Dual Attention*) dan memperoleh nilai rata-rata *micro-F1* sebesar 39,41%. Pada penelitian [27], penerapan *Multi-task Learning* (MTL) untuk *domain adaptive pretraining* berhasil meminimalisir informasi bising dari pohon dependensi sintaksis dan secara konsisten mengungguli metode *baseline* dalam hal akurasi dan *F1-score*.

Penelitian [22] yang mengkaji strategi *pre-training* dan *fine-tuning* pada *cross-domain low-resource* menunjukkan bahwa *transfer learning* secara konsisten mampu menangkap fitur spesifik domain dan menghasilkan nilai *micro-F1* yang lebih tinggi dibandingkan seluruh model *baseline*. Sementara itu, penelitian [28] menggunakan IndoBERT dengan tahapan *hyperparameter tuning* untuk klasifikasi aspek dan sentimen pada ulasan mahasiswa, menghasilkan akurasi yang sangat tinggi yaitu 97,9% untuk sentimen dan 97,3% untuk aspek. Penelitian [29] menerapkan RoBERTa dengan proses *fine-tuning* pada klasifikasi *multilabel* dan memperoleh rata-rata akurasi sebesar 89%, sedangkan penelitian [23] yang menggunakan *Domain-Adversarial Training* (DAT) dengan *cross-domain* tanpa data berlabel mencapai akurasi yang bervariasi antara 35% hingga 75% bergantung pada jarak antar domainnya.

Meskipun berbagai penelitian tersebut menunjukkan performa yang baik, masih terdapat sejumlah keterbatasan, terutama pada integrasi antara pendekatan *multi-label*, *cross-domain learning*, strategi adaptasi domain, dan domain produk kecantikan berbahasa Indonesia. Sebagian penelitian hanya berfokus pada klasifikasi *single-label*, sebagian lainnya tidak memanfaatkan transfer antar-domain, atau belum menggunakan arsitektur yang mengoptimalkan ekstraksi fitur kontekstual secara bersamaan. Oleh karena itu, penelitian ini diusulkan untuk mengisi kesenjangan tersebut dengan mengembangkan model *Cross-Domain Multi-Label Aspect-Based Sentiment Analysis* pada ulasan produk kecantikan. Penelitian ini memanfaatkan *source domain* yang telah berlabel untuk mentransfer pengetahuan ke *target domain* yang belum berlabel, menggunakan strategi *Domain-Adaptive Pre-Training* berbasis IndoBERT, yang kemudian diintegrasikan

dengan arsitektur BiLSTM *Dual Attention* serta *head classification* untuk memprediksi aspek dan sentimen secara bersamaan.

2.2 Teori yang berkaitan

2.2.1 Sentiment Analysis (SA)

Sentiment Analysis (SA) merupakan metode yang berfokus pada pengolahan opini, penilaian, serta emosi seseorang terhadap berbagai subjek, mulai dari produk dan layanan hingga isu tertentu. Teknik ini umumnya diterapkan pada data berbasis teks, seperti komentar di media sosial, ulasan pelanggan, atau artikel berita. Fokus utamanya adalah untuk membedah persepsi maupun pandangan masyarakat terhadap suatu merek atau peristiwa secara sistematis. [30]

Pada tingkat analisis kalimat, tugas utama dari pendekatan ini adalah mengidentifikasi kecenderungan emosi atau polaritas dari sebuah pernyataan tertulis. Sistem komputasi akan secara otomatis mengklasifikasikan apakah setiap kalimat tersebut mengungkapkan opini yang bersifat positif, negatif, ataupun netral. Dalam praktiknya, sentimen netral umumnya didefinisikan sebagai bentuk pernyataan faktual dan objektif yang tidak mengandung unsur opini emosional sama sekali [31].

2.2.2 Aspect Based Sentiment Analysis (ABSA)

Aspect-Based Sentiment Analysis (ABSA) merupakan pendekatan analisis sentimen beraliran halus yang berfokus pada evaluasi opini publik terhadap suatu entitas beserta atribut-atribut spesifiknya. Metode komputasi ini bekerja dengan memproses sekumpulan data teks tidak terstruktur, seperti ulasan pelanggan atau unggahan di platform media sosial, yang membahas suatu subjek tertentu. Berbeda dengan klasifikasi sentimen pada tingkat dokumen atau kalimat utuh, ABSA mampu mengekstraksi informasi sentimen yang sangat mendetail dan spesifik dari berbagai komponen yang berbeda meskipun berada di dalam satu teks yang sama.

Fokus utama dari sistem ini adalah mendeteksi aspek atau fitur yang paling sering dibicarakan dalam teks tersebut [32]. Sebagai contoh praktis, sistem dapat secara otomatis mengenali dan memisahkan aspek seperti 'baterai' atau 'layar' dari sekumpulan teks ulasan mengenai sebuah perangkat ponsel pintar[33]. Setelah target aspek berhasil diidentifikasi, model komputasi akan mengevaluasi serta memperkirakan rata-rata sentimen yang diberikan oleh pengguna terhadap masing-masing aspek tersebut [34].

2.2.3 Multi-Label Aspect Based Sentiment Analysis (ML-ABSA)

Multi-Label Aspect-Based Sentiment Analysis (ML-ABSA) merupakan pengembangan dari ABSA yang memodelkan proses identifikasi aspek dan sentimen sebagai permasalahan *multi-label classification* [35]. Dalam pendekatan ini, satu ulasan dapat memiliki lebih dari satu pasangan aspek-sentimen secara bersamaan. Oleh karena itu, ML-ABSA memungkinkan model untuk menangkap berbagai sentimen terhadap aspek yang berbeda secara lebih menyeluruh dan merepresentasikan penilaian konsumen secara lebih independen [36].

Penerapan arsitektur ML-ABSA memungkinkan model untuk menangkap berbagai sentimen yang saling tumpang tindih secara jauh lebih menyeluruh. Melalui deteksi opini yang lebih komprehensif, metode ini mampu merepresentasikan penilaian aktual konsumen secara lebih realistis sesuai dengan kompleksitas bahasa natural manusia. Perkembangan riset terkini membuktikan bahwa klasifikasi multilabel pada ABSA menjadi semakin akurat ketika dikombinasikan dengan teknik ekstraksi opini berbasis jaringan saraf tiruan. Hasilnya, pendekatan komputasi ini secara signifikan membantu para peneliti maupun pemangku kepentingan dalam mengidentifikasi area operasional spesifik yang membutuhkan peningkatan kualitas secara presisi [37].

2.2.4 X Platform

Platform X, yang sebelumnya dikenal sebagai Twitter, merupakan media sosial berbasis *microblogging* yang memfasilitasi komunikasi singkat dan

interaksi secara *real-time*. Karakteristik utama platform ini terletak pada kemampuannya dalam mendistribusikan informasi secara cepat melalui fitur seperti *repost*, *mention*, dan *hashtag* [38]. Sifat teksnya yang memiliki batasan karakter mendorong pengguna untuk menghasilkan konten tulisan yang padat, ekspresif, dan langsung pada intinya, menjadikannya sumber data tekstual yang sangat ideal untuk riset pemrosesan bahasa alami.

Karena sifatnya yang terbuka, X sering digunakan oleh pengguna sebagai saluran publik untuk menyampaikan opini, dukungan, maupun kritik terhadap berbagai hal, termasuk produk dan layanan. Berkat tingginya volume partisipasi penggunanya, platform ini telah berevolusi menjadi salah satu sumber *dataset* yang paling esensial dan aktual untuk mengukur reaksi publik di dalam berbagai penelitian analisis sentimen [39]. Aliran konten buatan pengguna yang terjadi secara terus-menerus ini pada akhirnya menyediakan sekumpulan data besar bagi para peneliti untuk memantau, mengekstraksi, dan mengevaluasi pergeseran subjektivitas publik secara sistematis

2.2.5 Cross Domain Learning

Cross Domain Learning merupakan sebuah pembelajaran yang digunakan untuk menyelaraskan struktur keseluruhan dari dua atau lebih sistem representasi [40]. Pendekatan pembelajaran mesin yang bertujuan mentransfer pengetahuan yang diperoleh dari suatu domain sumber atau *source domain* ke domain target yang memiliki karakteristik data berbeda, tetapi masih berkaitan dengan tugas yang sama [41]. Meskipun kedua domain tersebut memiliki pola fitur turunan yang tidak identik, proses transfer pembelajaran ini tetap difokuskan secara konsisten pada penyelesaian jenis tugas analitik yang sama.

Dalam penerapannya, metode ini sering kali diintegrasikan dengan kerangka *semi-supervised learning* yang memanfaatkan kombinasi antara data berlabel dan data tidak berlabel secara bersamaan [42]. Skenario pembelajaran hibrida ini berfungsi untuk mengoptimalkan pemetaan pola fitur dari domain sumber agar dapat diaplikasikan pada domain target secara lebih akurat. Pada akhirnya, mekanisme adaptasi lintas domain ini terbukti mampu meningkatkan

performa model prediksi secara komprehensif di lingkungan domain yang baru. Strategi adaptif ini menjadi solusi analitik yang sangat krusial, terutama untuk mengatasi kendala utama ketika ketersediaan data berlabel pada domain target sangat terbatas atau sulit diperoleh [43].

2.2.6 Tweet Harvest

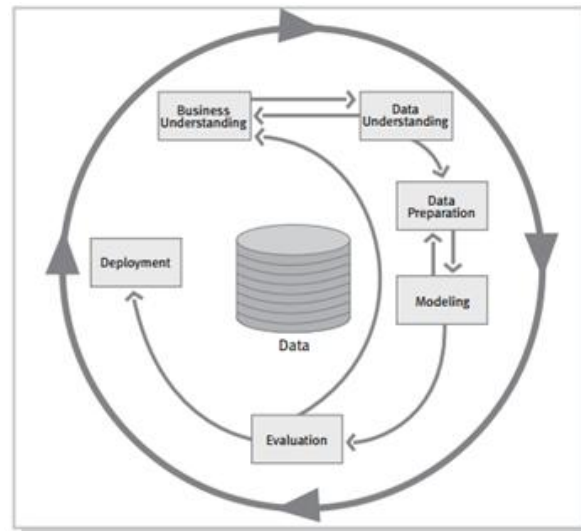
Tweet-Harvest adalah alat otomasi ekstraksi data pada platform X yang beroperasi melalui *Application Programming Interface* (API). Alat komputasi ini berfungsi untuk mengambil berbagai elemen informasi publik yang tersebar di dalam ekosistem media sosial tersebut ke dalam format yang lebih terstruktur. Melalui jalur integrasi API ini, program mampu menelusuri, memfilter, dan mengunduh unggahan spesifik secara otomatis berdasarkan parameter pencarian yang telah ditentukan [44].

Dalam proses penarikan datanya, alat ini mengandalkan *auth_token* unik sebagai kunci akses keamanan dan validasi identitas akun. token otorisasi tersebut bertindak sebagai protokol verifikasi utama yang memastikan bahwa seluruh aktivitas pengambilan data terhubung dengan sesi pengguna yang sah. [45]. Teknologi ini memungkinkan untuk mengumpulkan data teks dalam volume besar secara efisien untuk keperluan analisis lebih lanjut.

2.3 Framework atau Algoritma yang digunakan

2.3.1 CRISP-DM

Cross-Industry Standard Process for Data Mining atau CRISP-DM merupakan kerangka kerja standar yang digunakan secara luas dalam proyek *data mining* karena fleksibilitasnya dalam menyelesaikan berbagai persoalan industri. [46] Metodologi ini memastikan penelitian berjalan lebih terstruktur melalui siklus yang berulang seperti pada gambar 2.1.



Gambar 2. 1 Alur Framework CRISP-DM [46]

Berdasarkan standar CRISP-DM, terdapat enam tahapan utama dalam proses pengolahan data:

1. Business Understanding

Tahap awal yang berfokus pada pemahaman tujuan dan kebutuhan dari sudut pandang organisasi. Pada fase ini, masalah yang akan diselesaikan diidentifikasi secara jelas untuk menentukan tujuan teknis *data mining* yang relevan.

2. Data Understanding

Proses pengumpulan data awal guna mengenali karakteristik serta kualitas data tersebut. Langkah ini melibatkan eksplorasi data untuk mengidentifikasi potensi masalah pada data dan memahami hubungan antar variabel yang ada.

3. Data Preparation

Tahap ini meliputi seluruh kegiatan untuk menghasilkan kumpulan data akhir yang akan diproses lebih lanjut. Aktivasnya mencakup pembersihan data, integrasi sumber data, serta transformasi data ke dalam format yang sesuai dengan kebutuhan model.

4. Modeling

Pada tahap ini, berbagai teknik pemodelan atau algoritma dipilih dan diterapkan. Proses ini juga melibatkan pembagian data menjadi data latih (*training*) dan data uji (*testing*), serta penyesuaian parameter agar model memberikan hasil yang optimal.

5. Evaluation

Model yang telah dibangun dievaluasi untuk memastikan kesesuaiannya dengan tujuan bisnis yang telah ditetapkan di awal. Performa model diukur menggunakan metrik tertentu, seperti akurasi, *precision*, *recall*, dan *F1-score*, guna memvalidasi hasil sebelum masuk ke tahap penggunaan.

6. Deployment

Tahap akhir di mana hasil analisis disusun ke dalam laporan atau diintegrasikan ke dalam sistem produksi. Tujuan utamanya adalah agar temuan dari proses *data mining* dapat dikonversi menjadi strategi yang dapat diimplementasikan oleh pihak terkait. Namun, penelitian yang dilakukan saat ini hanya akan sampai pada tahap kelima atau evaluasi karena fokus pada pengujian efektivitas model klasifikasi dalam mengidentifikasi sentimen pada aspek produk.

Penentuan CRISP-DM sebagai framework yang digunakan dalam penelitian karena telah dilakukan perbandingan dengan framework lainnya seperti KDD dan SEMMA dengan perbandingan pada tabel 2.2

Tabel 2. 2 Perbandingan Framework

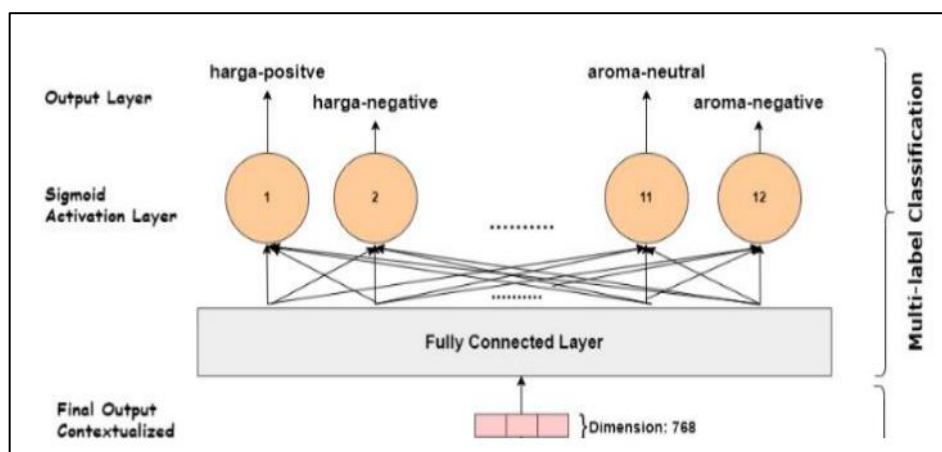
Aspek Pemanding	CRISP-DM[47]	KDD[48]	SEMMA[49]
Fokus & Struktur	Terdiri dari 6 tahapan yang berfokus pada penyesuaian pemahaman bisnis dengan penerapan teknisnya.	Terdiri dari 5 tahapan yang berfokus murni pada penggalian dan pemahaman informasi dari sekumpulan data.	Terdiri dari 5 tahapan terpisah yang saling mendukung, dengan fokus utama pada proses pengembangan model.
Fleksibilitas Alur	Sangat adaptif; memungkinkan peneliti bergerak maju-mundur dan melakukan iterasi	Tidak terikat aturan ketat; peneliti bebas mengulangi atau bergerak maju-mundur di setiap langkahnya.	Memiliki fleksibilitas yang sangat bebas pada setiap tahapannya selama siklus proyek berlangsung.

Aspek Perbandingan	CRISP-DM[47]	KDD[48]	SEMMA[49]
	antar tahap sesuai kebutuhan.		
Fokus Akhir	Memiliki tahap <i>deployment</i> secara eksplisit untuk mengimplementasikan model final ke dunia nyata.	Tidak memiliki tahap implementasi sistem; hasil akhirnya lebih difokuskan pada penerapan pengetahuan.	Lebih memprioritaskan evaluasi tingkat keandalan model serta penentuan keputusan kegunaannya.
Rincian Tahapan	<ol style="list-style-type: none"> 1. Pemahaman Bisnis 2. Pemahaman Data 3. Persiapan Data 4. <i>Modeling</i> 5. Evaluasi 6. <i>Deployment</i> 	<ol style="list-style-type: none"> 1. <i>Pre-KDD</i> 2. Seleksi 3. <i>Pre-processing</i> 4. Transformasi 5. <i>Data Mining</i> 6. Interpretasi atau Evaluasi 	<ol style="list-style-type: none"> 1. Sampel 2. Jelajah 3. Modifikasi 4. <i>Model</i> 5. Evaluasi

Berdasarkan perbandingan ketiga framework pada tabel 2.2 memiliki tujuan yang sama, yaitu menyediakan kerangka kerja sistematis untuk proyek data mining, namun dengan fokus yang berbeda. KDD menitikberatkan pada proses penemuan pengetahuan dari data, SEMMA lebih berorientasi pada tahapan teknis pengembangan model, sedangkan CRISP-DM mengintegrasikan pemahaman bisnis, pengolahan data, pemodelan, evaluasi, hingga implementasi hasil. Dari sisi struktur, CRISP-DM memiliki tahapan yang paling lengkap karena mencakup enam fase yang mencerminkan seluruh siklus proyek analitik. Selain itu, CRISP-DM bersifat berulang dan menyesuaikan, sehingga memungkinkan peneliti untuk kembali ke tahap sebelumnya ketika ditemukan permasalahan pada data maupun model. Keunggulan utama CRISP-DM dibandingkan KDD dan SEMMA terletak pada keberadaan fase *Business Understanding* dan *Deployment*, yang memastikan bahwa penelitian tidak hanya menghasilkan model yang akurat, tetapi juga relevan terhadap tujuan penelitian dan siap diterapkan. Oleh karena itu, CRISP-DM dinilai lebih unggul dan lebih sesuai digunakan dalam penelitian ini karena mampu menyediakan alur kerja yang terstruktur dan menyesuaikan dari tahap perumusan masalah hingga evaluasi hasil model.

2.3.2 Indo-BERT

Pendekatan IndoBERT merupakan strategi klasifikasi di mana model berfungsi sebagai representasi teks sekaligus pengklasifikasi dalam satu arsitektur. Dalam mekanisme ini, teks ulasan produk kecantikan diproses langsung oleh IndoBERT untuk menghasilkan prediksi sentimen pada aspek-aspek yang telah ditentukan seperti pada gambar 2.2.



Gambar 2. 2 Arsitektur BERT[25]

Beberapa komponen teknis utama dalam klasifikasi *end-to-end* IndoBERT adalah sebagai berikut:

1. Arsitektur Transformer

IndoBERT menggunakan arsitektur *contextual embedding* yang mampu menangkap makna kata berdasarkan kata-kata di sekitarnya, berbeda dengan model statis seperti Word2Vec.

2. Lapisan Klasifikasi (*Classifier Layer*)

Sebuah lapisan tambahan ditempatkan di atas output *mean pooling* dari model IndoBERT untuk memetakan representasi vektor ke label sentimen.

3. Fungsi Aktivasi Sigmoid

Digunakan untuk menangani tugas klasifikasi multi-label, digunakan fungsi aktivasi Sigmoid dibandingkan Softmax. Sigmoid memungkinkan setiap label aspek-sentimen (misalnya: Harga-Positif, Aroma-Negatif) memiliki probabilitas independen antara 0 hingga 1.

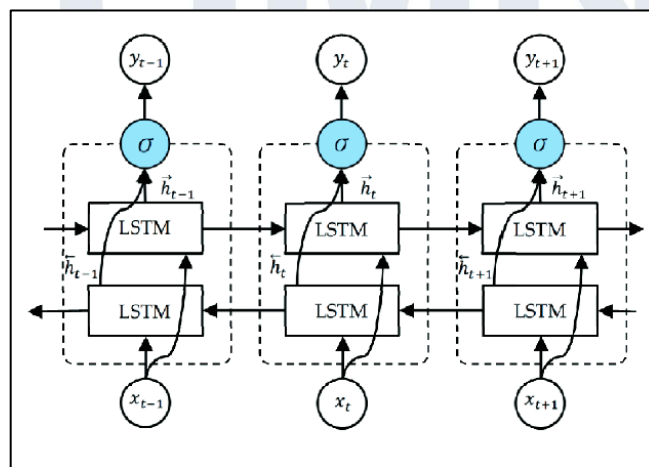
4. Ambang Batas (*Threshold*)

Prediksi akhir ditentukan berdasarkan nilai ambang batas, di mana nilai probabilitas di atas 0,5 dianggap sebagai label yang muncul dalam ulasan tersebut. Kelebihan utama dari model *end-to-end* ini adalah kemampuannya dalam menangkap korelasi atau hubungan antar aspek yang muncul dalam satu ulasan. Hasil penelitian menunjukkan bahwa model *end-to-end* IndoBERT (khususnya varian *base*) memiliki performa yang signifikan dengan tingkat akurasi mencapai 86,98% pada ulasan produk kecantikan Indonesia [25].

2.3.3 Bidirectional Long Short-Term Memory

Bidirectional Long Short-Term Memory (BiLSTM) merupakan pengembangan dari arsitektur *Long Short-Term Memory* (LSTM) yang dirancang untuk memahami informasi sekuensial dari dua arah, yaitu *forward* dan *backward*. Berbeda dengan LSTM biasa yang hanya memproses data dari urutan awal ke akhir [50], BiLSTM memproses data dari awal ke akhir dan dari akhir ke awal secara bersamaan sehingga mampu memahami konteks kalimat secara lebih menyeluruh.

Pada NLP, BiLSTM banyak digunakan karena mampu menangkap hubungan antar kata dalam suatu kalimat dengan lebih baik. Hal ini sangat penting dalam tugas analisis sentimen dan ABSA, karena makna suatu kata sering dipengaruhi oleh kata sebelum maupun sesudahnya.



Gambar 2.3 Arsitektur BiLSTM [51]

Arsitektur BiLSTM terdiri dari dua lapisan LSTM, yaitu forward LSTM dan backward LSTM. Forward LSTM memproses urutan data dari kiri ke kanan, sedangkan backward LSTM memproses data dari kanan ke kiri. Hasil dari kedua proses tersebut kemudian digabungkan untuk menghasilkan representasi konteks yang lebih lengkap. BiLSTM memiliki keunggulan dalam memahami dependensi jangka panjang pada teks serta mampu meningkatkan performa klasifikasi sentimen dibandingkan metode sequential tradisional. Oleh karena itu, model ini banyak diterapkan pada penelitian sentiment analysis, text classification, machine translation, hingga named entity recognition.

2.3.4 Sentence-Pair Injection

Pada pendekatan ini, input tidak lagi berupa satu kalimat tunggal, melainkan pasangan kalimat yang terdiri dari teks ulasan dan kalimat tambahan (*auxiliary sentence*) yang merepresentasikan aspek tertentu. Dengan demikian, proses klasifikasi tidak hanya bergantung pada representasi teks secara keseluruhan, tetapi juga mempertimbangkan hubungan antara ulasan dan aspek yang dianalisis. Dalam mekanisme ini, setiap teks ulasan akan dipasangkan dengan beberapa *auxiliary sentence* yang dibentuk berdasarkan kategori aspek. Artinya, satu ulasan dapat diubah menjadi beberapa pasangan data, di mana setiap pasangan merepresentasikan satu aspek tertentu. [52]

$$S = ([CLS], C, [SEP], A, [SEP])$$

Rumus 2. 1 Rumus Sentence-Pair Injection

- C = Teks ulasan pelanggan.
- A = Kalimat tambahan.
- $[CLS]$ = Representasi keseluruhan input.
- $[SEP]$ = Pemisah antara kalimat.

Penggunaan *auxiliary sentence* memungkinkan model untuk memanfaatkan mekanisme *next sentence prediction* (NSP) yang telah dipelajari pada tahap *pre-training*. Dengan adanya pasangan kalimat tersebut, model dapat memahami hubungan semantik antara teks ulasan dan aspek yang diberikan, sehingga *attention* dapat difokuskan pada bagian teks yang relevan terhadap

aspek tertentu. Hal ini menjadikan pendekatan *sentence-pair classification* lebih efektif dalam menangkap keterkaitan antara aspek dan konteks dibandingkan pendekatan *single-sentence classification*.

2.3.5 Dual Attention Mechanism

Dual Attention Mechanism merupakan pengembangan dari mekanisme *attention* yang dirancang untuk meningkatkan kemampuan model dalam menangkap hubungan antara aspek dan konteks. Pendekatan *dual attention* yang terdiri dari perhatian pada tingkat kalimat untuk memahami relevansi aspek secara global, serta perhatian pada tingkat konteks untuk menangkap urutan dan hubungan antar kata dalam kalimat. [53] Pada implementasinya dalam model BB-DAM, urutan input teks dibagi menjadi tiga bagian utama, yaitu *above*, *aspectual words*, dan *below*. Segmen *above* merepresentasikan bagian kalimat sebelum aspek, *aspectual words* merupakan kata atau frasa aspek yang menjadi fokus analisis, sedangkan *below* adalah bagian kalimat setelah aspek. Pembagian ini bertujuan untuk memisahkan konteks berdasarkan posisi relatif terhadap aspek sehingga interaksi antara aspek dan konteks dapat dianalisis secara lebih terstruktur. [24]

Selanjutnya, mekanisme *dual attention* bekerja dengan mengevaluasi interaksi antara kata aspek dengan ketiga segmen tersebut. *Sentence-level attention* digunakan untuk mengidentifikasi bagian kalimat yang paling relevan terhadap aspek secara keseluruhan, sedangkan *context-level attention* berperan dalam menangkap dependensi dan hubungan antar kata dalam masing-masing segmen konteks. Kombinasi kedua mekanisme ini memungkinkan model untuk mengekstraksi informasi interaksi secara lebih menyeluruh, baik dari sisi global maupun lokal.

2.3.6 Evaluasi Matriks

Confusion Matrix merupakan alat evaluasi yang digunakan dalam masalah klasifikasi pembelajaran mesin untuk mengukur kinerja model dengan cara membandingkan hasil prediksi terhadap data aktual [54]. Matriks ini menyajikan

sebaran kebenaran prediksi dalam bentuk tabel yang terdiri dari empat kombinasi nilai utama [55]:

1. True Positive (TP): Jumlah data positif yang diprediksi dengan benar sebagai positif.
2. True Negative (TN): Jumlah data negatif yang diprediksi dengan benar sebagai negatif.
3. False Positive (FP): Jumlah data negatif yang salah diprediksi sebagai positif.
4. False Negative (FN): Jumlah data positif yang salah diprediksi sebagai negatif.

Berdasarkan nilai-nilai dalam confusion matrix tersebut, dapat dihitung beberapa metrik utama untuk memberikan gambaran performa model secara lebih mendalam, yaitu:

2.3.5.1 Accuracy

Mengukur tingkat kebenaran model dalam mengklasifikasikan seluruh data secara keseluruhan yang didefinisikan pada rumus 2.2.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Rumus 2. 2 Rumus Akurasi

2.3.5.2 Precision

Menunjukkan tingkat keakuratan antara data yang diprediksi positif dengan realitas hasil positifnya yang didefinisikan pada rumus 2.3.

$$Presisi = \frac{TP}{TP + FP}$$

Rumus 2. 3 Rumus Precision

2.3.5.3 Recall

Mengukur kemampuan model dalam mengidentifikasi kembali seluruh data yang termasuk dalam kategori positif yang didefinisikan pada rumus 2.4.

$$\text{Sensitivitas} = \frac{TP}{TP + FN}$$

Rumus 2. 4 Rumus Recall

2.3.5.4 F1-Score

Merupakan nilai rata-rata harmonik yang mengombinasikan precision dan recall untuk memberikan keseimbangan penilaian performa model yang didefinisikan pada rumus 2.5.

$$F1 = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Rumus 2. 5 Rumus F1-Score

2.4 Tools atau software yang digunakan

2.4.1 Google Collaboration

Google Colaboratory, atau yang lebih dikenal sebagai Colab, merupakan sebuah lingkungan *notebook* Jupyter berbasis *cloud* yang disediakan oleh Google. Alat ini digunakan untuk memfasilitasi kebutuhan pengolahan data dan pengembangan model secara gratis. Platform ini karena tidak memerlukan konfigurasi lokal, sehingga pengguna dapat langsung membuat serta menjalankan kode pemrograman Python melalui antarmuka yang ramah pengguna.

Sebagai *tools*, Colaboratory menawarkan integrasi mendalam dengan layanan *Google Cloud Platform* (GCP), termasuk kemampuan untuk mengimpor, membaca, maupun menulis data secara langsung ke Google Drive menggunakan pustaka *google.colab*. Secara teknis, platform ini memungkinkan fleksibilitas tinggi dalam pemilihan jenis *runtime*, di mana pengguna dapat memanfaatkan dukungan komputasi tingkat tinggi seperti *Graphics Processing Unit* (GPU) dan *Tensor Processing Unit* (TPU) untuk mempercepat proses pelatihan model pembelajaran mesin yang kompleks [56].

2.4.2 Python

Python adalah bahasa pemrograman tingkat tinggi yang banyak digunakan dalam analisis data dan machine learning. Keunggulan utama dari teknologi perangkat lunak ini terletak pada struktur sintaksisnya yang ringkas, intuitif, dan

sangat mudah dipahami oleh para pengembang. Desain pemrograman yang mengedepankan tingkat keterbacaan tinggi ini sangat membantu para peneliti dalam merancang algoritma komputasi yang rumit dengan penulisan baris kode yang sangat efisien.

Selain kemudahan sintaksisnya, daya tarik utama Python juga berasal dari ketersediaan ekosistem pustaka sumber terbuka yang sangat kaya dan lengkap. Pada tahapan pengolahan informasi, bahasa ini menyediakan modul Pandas yang dirancang secara khusus untuk memfasilitasi manipulasi serta transformasi data berstruktur secara sistematis. Selanjutnya, terdapat pustaka Scikit-learn yang menawarkan ragam algoritma *machine learning* mutakhir, serta Matplotlib yang berfungsi untuk menyajikan visualisasi analitik secara komprehensif. Kombinasi dari seluruh pustaka fungsional tersebut menjadikan Python sebagai lingkungan pengembangan yang paling ideal untuk membangun arsitektur pemrosesan bahasa alami [57].

