

BAB II

LANDASAN TEORI

2.1 Penelitian Terkait

Tabel 2. 1 Tabel Artikel Penelitian Terkait

No.	Jurnal	Judul	Penulis	Metode	Hasil
1	International Journal of Intelligent Engineering & Systems, Vol. 14, No. 5, pp. 294-305, 2021	Aspect Based Sentiment Analysis for Explicit and Implicit Aspects in Restaurant Review using Grammatical Rules, Hybrid Approach, and SentiCircle	Rachmad Abdullah, Suhariyanto, dan Riyanarto Sarno [12]	TF-ICF, <i>SentiCircle</i> , <i>Opinion Lexicon</i> , dan <i>Implicit Lexicon</i>	ABSA1 menghasilkan F1-Measure sebesar 0.84, ABSA2 sebesar 0.86, dan ABSA3 sebesar 0.89. Akurasi tertinggi diraih oleh ABSA3 dengan menggabungkan keseluruhan metode.
2	Current Applied Science and Technology, Vol. 23, No. 2, pp. 1-16, 2023	A Hybrid Approach for Aspect-based Sentiment Analysis: A Case Study of Hotel Reviews	Khanista Namee, Jantima Polpinij, dan Bancha Luaphol [13]	SVM, CNN, dan RBF Kernel	Akurasi tertinggi diraih oleh SVM classifier dengan linear kernel, yaitu 3.66% Recall, 4.94% Precision, 4.94% F1, 4.88% Accuracy, dan 4.94% AUC.
3	Ultima InfoSys : Jurnal Ilmu Sistem Informasi, Vol. 13, No.	Aspect-Based Sentiment Analysis on Application Review using	Putri Arta Aritonang, Monika Evelin Johan, dan Iwan	<i>Convolutional Neural Network</i>	CNN menghasilkan performa yang baik dengan f1 score 92.23% pada klasifikasi aspek

No.	Jurnal	Judul	Penulis	Metode	Hasil
	1, pp. 54-61, 2022	Convolutional Neural Network	Prasetiawan [14]		dan 95.13% pada klasifikasi sentimen. Pada versi 4.0.5 hasil klasifikasi aspek naik sebesar 17.21%
4	Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi), Vol. 8, No. 2, pp. 393-404, 2024	Analisis Sentimen Film Dirty Vote Menggunakan BERT (Bidirectional Encoder Representations from Transformers)	Diah Fatma Sjoraida, Bucky Wibawa Karya Guna, dan Dudi Yudhakusuma [15]	<i>Bidirectional Encoder Representations from Transformers</i> (BERT)	Algoritma BERT menghasilkan tingkat kinerja yang tinggi dengan akurasi yang melebihi ambang batas 0.8 serta ditemukan aspek-aspek tertentu pada film.
5	International Journal of Intelligent Networks, Vol. 2, pp. 64-69, 2021	Improving the Performance of Aspect Based Sentiment Analysis using Fine-Tuned Bert Base Uncased Model	M.P. Geetha dan D. Karthika Renuka [16]	<i>Bidirectional Encoder Representations from Transformers</i> (BERT)	Algoritma BERT dan empat algoritma <i>machine learning</i> digunakan untuk melakukan prediksi sentiment terhadap ulasan pelanggan saat membeli produk dan didapatkan bahwa BERT memberikan prediksi yang lebih akurat.
6	Social Network Analysis and Mining Vol.	Sentiment Analysis on the Impact of Coronavirus in	Mrityunjay Singh, Amit Kumar Jakhar, dan	<i>Bidirectional Encoder Representations from</i>	Algoritma BERT digunakan untuk melakukan analisis sentimen terhadap

No.	Jurnal	Judul	Penulis	Metode	Hasil
	11, No. 1, pp. 1-11, 2021	Social Life using the BERT Model	Shivam Pandey [17]	<i>Transformers</i> (BERT)	kumpulan data pada Twitter. BERT menghasilkan akurasi sebesar 94%
7	G-Tech: Jurnal Teknologi Terapan Vol.8, No. 2, pp. 857-867, 2024	How Positive Sentiment of Telemedicine Applications using Naïve Bayes and Support Vector Machine?	Andreas dan Wella [18]	<i>Support Vector Machine</i> dan <i>Naïve Bayes</i>	Akurasi SVM lebih akurat 90,99% dibandingkan dengan Naive Bayes
8	International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 14, No. 8, pp. 113-117, 2023	Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach	Helmi Imaduddin, Fiddin Yusfida A'la, dan Yusuf Sulisty Nugroho [19]	IndoBERT	IndoBERT efektif dalam melakukan analisis sentimen teks Bahasa Indonesia dengan menghasilkan akurasi sebesar 96%.
9	IJCCS (Indonesian Journal of Computing and Cybernetics Systems), Vol.17, No.1, pp.79-90, 2023	Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews	Cynthia As Bahri dan Lya Hulliyyatus Suadaa [20]	SVM, Complement Naïve Bayes, Logistic Regression, pre-trained BERT, IndoBERT, dan mBERT	Model IndoBERT menghasilkan performa terbaik dan SVM memberikan hasil terbaik akurasi sebesar 89,16% pada analisis sentimen salah satu destinasi wisata di Indonesia.

No.	Jurnal	Judul	Penulis	Metode	Hasil
10	Jurnal Pendidikan Teknologi Informasi (JUKANTI), Vol.6, No.1, pp. 1-10, 2023	Analisis Sentimen berbasis Aspek pada Ulasan Aplikasi KAI Access menggunakan Metode Support Vector Machine	Gracia Radiena dan Adi Nugroho [21]	Support Vector Machine	Penggunaan model CRISP-DM dan algoritma SVM menghasilkan akurasi yang tinggi di tiap aspek pada sentiment.

Penelitian pertama melakukan analisis sentimen berbasis aspek terhadap ulasan restoran berdasarkan aspek *explicit* dan *implicit* menggunakan metode TF-ICF, *SentiCircle*, *Opinion Lexicon*, dan *Implicit Lexicon*. Setelah melakukan analisis sentimen, diketahui bahwa menggabungkan keseluruhan metode akan menghasilkan akurasi *F1-Measure* yang tinggi, yaitu sebesar 0.89 dibandingkan hanya memakai metode *SentiCircle* dan *Opinion Lexicon*. Dengan aspek yang diketahui melalui pendekatan 3 ABSA adalah makanan dan harga [12]. Pada penelitian kedua, melakukan analisis sentimen berbasis aspek terhadap ulasan customer pada hotel dengan aspek yang dipertimbangkan adalah pelayanan staf, kebersihan kamar, uang, serta kenyamanan menggunakan metode SVM dan CNN. Dapat dihasilkan bahwa akurasi tertinggi diraih oleh SVM *classifier* dengan linear kernel, yaitu 3.66% *Recall*, 4.94% *Precision*, 4.94% *F1-score*, 4.88% *Accuracy*, dan 4.94% *AUC*. Di sisi lain, klasifikasi aspek menghasilkan nilai *Recall* sebesar 0.9, *Precision* sebesar 0.89, dan *F1-score* sebesar 0.895 [13].

Penelitian ketiga melakukan analisis sentimen berbasis aspek terhadap ulasan aplikasi PeduliLindungi di berbagai versi menggunakan algoritma CNN dan dihasilkan bahwa CNN menghasilkan performa yang baik dengan *F1-score* 92.23% pada klasifikasi aspek dan 95.13% pada klasifikasi sentimen. Pada versi 4.0.5 hasil klasifikasi aspek naik sebesar 17.21%. Berdasarkan sentimen pengguna, ulasan dominan menunjukkan sentimen negatif pada aspek aplikasi, yaitu *Visual Experience*, *Scan-Check-in/out*, *Vaccine Certificate*, *eHac*, *COVID Test*, *Register/Login*, *Performance and Stability*, dan *Privacy, Data, and Security* [14].

Penelitian keempat melakukan analisis sentimen terhadap film ‘*Dirty Vote*’ dengan menggunakan algoritma BERT dan hasil yang didapatkan adalah algoritma BERT memiliki kinerja yang tinggi dengan akurasi yang melebihi ambang batas 0.8 dan terbukti bahwa algoritma BERT dapat melakukan prediksi terhadap sentimen dengan baik [15].

Penelitian kelima melakukan analisis sentiment dengan menggunakan empat algoritma machine learning dan hasil dari penelitian tersebut menunjukkan bahwa algoritma BERT memiliki kinerja yang lebih unggul dibandingkan dengan algoritma tradisional dengan akurasi yang dihasilkan adalah lebih tinggi serta dapat memahami konteks secara mendalam [16]. Penelitian keenam melakukan analisis sentimen terhadap dataset pada Twitter yang berlokasi di India pada saat terjadi COVID-19 dan didapatkan akurasi sebesar 94% dengan menggunakan algoritma BERT [17]. Penelitian ketujuh melakukan prediksi sentimen publik terhadap perangkat lunak telemedis menggunakan SVM dan Naïve Bayes dan didapatkan akurasi yang tinggi pada SVM, yaitu 90,99% [18]. Penelitian kedelapan melakukan analisis sentimen terhadap aplikasi Kesehatan di Indonesia menggunakan pendekatan IndoBERT dan menghasilkan akurasi sebesar 96% dengan memberikan wawasan terhadap kekuatan dan kelemahan aplikasi layanan Kesehatan di Google Play Store [19]. Penelitian kesembilan melakukan penelitian sentimen terhadap ulasan salah satu destinasi wisata di Indonesia pada Google Maps menggunakan beberapa model *machine learning* dan didapatkan hasil SVM menjadi model yang menghasilkan akurasi terbaik, yaitu 89,16% serta IndoBERT yang mencapai performa terbaik [20]. Penelitian kesepuluh melakukan analisis sentimen terhadap ulasan aplikasi KAI Access menggunakan algoritma SVM yang bertujuan untuk mengetahui tingkat kepuasan masyarakat terhadap aplikasi KAI Access. Penggunaan SVM menghasilkan akurasi yang tinggi di tiap aspek yang telah ditentukan [21].

Sepuluh penelitian di atas dijadikan referensi oleh peneliti untuk melakukan analisis sentimen berbasis aspek terhadap ulasan aplikasi karena algoritma SVM dan model BERT menghasilkan akurasi yang cukup baik pada proses analisis

ulasan atau sentimen pengguna. Tentunya, penelitian yang dilakukan oleh penulis memiliki perbedaan dengan penelitian sebelumnya karena beberapa hal. Pertama, analisis sentimen dilakukan berbasis aspek sehingga hasil yang didapatkan tidak hanya memisahkan ulasan positif dan negatif, tetapi aspek yang membuat pengguna menulis ulasan positif atau negatif diketahui secara objektif. Kedua, penelitian kali ini mengkombinasikan algoritma *Support Vector Machine* (SVM) dan *Bidirectional Encoder Representations from Transformers* (BERT) dengan model IndoBERT untuk mendapatkan hasil yang lebih akurat. Penggunaan algoritma SVM dan model IndoBERT dilakukan karena pada penelitian terdahulu belum ada yang melakukan penelitian analisis dengan menggunakan algoritma SVM untuk klasifikasi sentimen dengan model IndoBERT sebagai *feature extraction* representasi teks, sehingga penulis menggunakan algoritma dan model tersebut untuk memanfaatkan keunggulan kedua model dan mendapatkan akurasi yang optimal. Perbedaan ketiga, penelitian mengambil dataset ulasan aplikasi Duolingo pada Google Play Store sebanyak 30,000 data ulasan dari tahun 2020-2024 yang akan disortir kembali untuk wilayah Indonesia.

2.2 Teori Penelitian

2.2.1 Duolingo



Gambar 2. 1 Logo aplikasi Duolingo

Sumber: Google Play Store [22]

Duolingo merupakan sebuah aplikasi pembelajaran bahasa asing yang diciptakan oleh Luis von Ahn dan Severin Hacker pada tahun 2011. Luis von Ahn merupakan seorang ilmuwan komputer dan merupakan pendiri sebuah proyek Google yang membantu ratusan buku dan dokumen menjadi digital, yaitu reCAPTCHA. Sementara itu, Severin Hacker merupakan seorang insinyur

software di Austria dan menjadi salah satu pendiri reCAPTCHA bersama Luis. Duolingo memiliki perusahaan yang berpusat di Pittsburgh, Pennsylvania. Dalam pembuatannya, Luis dan Severin menciptakan aplikasi Duolingo agar pengguna mendapatkan pengalaman belajar bahasa yang mudah dan menyenangkan karena aplikasi Duolingo diciptakan berbentuk gamifikasi [23].

Di tahun 2012, Duolingo memiliki 5 kursus dan pada tahun berikutnya, Duolingo mulai menambahkan kursus sebanyak lebih dari 95 kursus dengan di awal aplikasi dirilis hanya tersedia dalam bahasa Inggris dan sekarang telah menawarkan 38 bahasa yang dapat digunakan. Saat ini, sudah lebih dari 500 juta pengguna mengunduh aplikasi Duolingo dengan 42 juta orang aktif satu kali dalam sebulan dan 10 juta orang lainnya merupakan pengguna aktif yang menggunakan aplikasi Duolingo setiap hari di tahun 2022. Adapun beberapa misi yang dijalankan oleh aplikasi Duolingo [24], antara lain:

- 1) Pendidikan yang dipersonalisasi: Karena setiap orang yang memiliki cara belajar yang berbeda, aplikasi Duolingo diciptakan untuk dapat menganalisis bagaimana jutaan orang belajar sehingga tercipta sistem yang efektif dan dapat menyesuaikan cara pembelajaran dengan setiap orang.
- 2) Pembelajaran yang menyenangkan: Duolingo menyadari bahwa banyak orang memiliki kesulitan ketika melakukan pembelajaran secara daring, sehingga aplikasi Duolingo diciptakan berbentuk gamifikasi yang dapat memberikan pengguna pengalaman belajar bahasa secara daring yang menyenangkan.
- 3) Diakses secara universal: Peluang besar akan terbuka ketika orang memiliki kemampuan dalam berbahasa asing. Tetapi, hambatan yang dialami ketika ingin belajar bahasa adalah biaya yang tidak murah dan tidak semua orang bisa memiliki akses untuk belajar bahasa. Oleh karena itu, aplikasi Duolingo diciptakan sebagai aplikasi belajar bahasa gratis sehingga semua orang memiliki kesempatan untuk melakukan pembelajaran bahasa asing.

Selain beberapa misi yang dijalankan Duolingo, terdapat pula pendekatan yang dilaksanakan, meliputi pendekatan dalam pengajaran yang tidak seperti buku teks, melainkan seperti bermain game sehingga memiliki pembelajaran yang menyenangkan dan tetap mengikuti kurikulum dengan standar internasional; aplikasi Duolingo mengedepankan pembelajaran seperti berkomunikasi di kehidupan nyata dengan berfokus pada pelatihan membaca, menulis, mendengarkan, dan berbicara dalam kehidupan sehari-hari; membuat konten yang menonjol, seperti menggunakan kalimat unik sehingga lebih mudah diingat dan menyenangkan untuk dipelajari; memungkinkan pengguna untuk belajar tanpa perlu berfokus pada aturan bahasa, sehingga pengetahuan dasar pengguna tentang bahasa dan aturan dapat berkembang; dengan menggunakan *machine learning*, aplikasi Duolingo menyajikan materi dengan tingkat kesulitan yang tepat bagi pengguna; menyajikan berbagai cara untuk belajar bahasa, seperti dengan cerita yang interaktif dan *podcast*; serta Duolingo berkomitmen untuk terus melakukan improvisasi sehingga materi-materi pembelajaran terus berkembang dan menjadi lebih baik.

2.2.2 Analisis Sentimen

Analisis sentimen merupakan proses melakukan identifikasi terhadap sentimen, evaluasi, dan opini dari data yang tertulis untuk menilai sebuah objek yang diteliti apakah cenderung memiliki sentimen yang positif atau negatif [25]. Analisis sentimen termasuk bagian dari *text mining* dan sub-bagian dari *natural language processing* (NLP). Dengan bantuan *natural language processing* (NLP), data tertulis yang di analisis akan melalui proses pembersihan dan pemahaman data, ekstrak fitur dari data tertulis, dan membangun model pada data tertulis secara otomatis agar informasi sentimen pada kalimat akan didapatkan dengan mengklasifikasikan sentimen pada data tertulis [26]. Pada saat ingin mengklasifikasikan data, terdapat tiga metode analisis sentimen yang dapat digunakan, antara lain:

- 1) Analisis Sentimen berbasis *Machine Learning*

Analisis sentimen ini membutuhkan pengimplementasian algoritma dalam menganalisis data agar kata-kata yang tertulis dapat diklasifikasikan ke dalam positif, negatif, dan netral. Fitur pada analisis sentimen berbasis pembelajaran mesin dapat mengetahui sentimen dalam bentuk teks dan semakin populer karena metode ini dinilai representatif. Beberapa algoritma yang dapat digunakan pada analisis sentimen berbasis pembelajaran mesin adalah SVM, NB, RF, dan DT [27].

2) Analisis Sentimen berbasis *Lexicon*

Cara kerja analisis sentimen berbasis *lexicon* adalah menggunakan diksi agar tanggapan konsumen/pengguna dapat diketahui dengan penilaian skor polaritas. Keunggulan yang dimiliki oleh metode ini adalah tidak memerlukan *data training*, tetapi diksi yang terdapat di leksikon belum termuat banyak [27].

3) Analisis Sentimen Campuran

Metode analisis sentimen merupakan gabungan dari metode *machine learning* dan *lexicon-based*. Metode ini dapat mengukur sentimen dengan hasil akurasi tinggi menggunakan algoritma *supervised learning* [27].

Dalam pengaplikasian analisis sentimen untuk mengidentifikasi sentimen pengguna, beberapa tipe analisis sentimen pada digunakan, yaitu *fine-grained sentiment analysis*, *intent sentiment analysis*, dan *aspect-based sentiment analysis*. *Aspect-based sentiment analysis* atau analisis sentimen berbasis aspek (ABSA) dalam beberapa tahun terakhir mendapatkan perhatian yang cukup besar karena merupakan pembaruan dari pendekatan orisinal yang berkembang dengan fokus utama adalah mengidentifikasi aspek-aspek pada sebuah teks, mengekstraksi emosi linguistik dalam sebuah teks yang merujuk pada aspek tersebut, dan mengidentifikasi polaritas opini untuk setiap aspek [28]. Dengan menggunakan analisis sentimen berbasis aspek, informasi akan didapatkan lebih detail terhadap opini pengguna/konsumen baik dalam sentimen maupun aspek kategori [29].

2.2.3 Text Mining

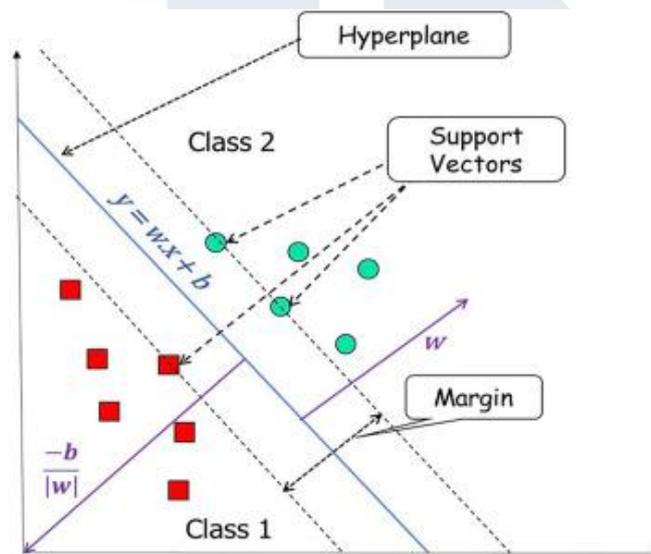
Text mining adalah bagian dari *data mining* dan merupakan suatu proses yang dilakukan untuk menganalisis data teks yang tidak terstruktur dengan cara mengekstraksi informasi berupa artikel, ulasan, media sosial, dan dokumen [30]. Dengan menggunakan *text mining*, data yang tidak diketahui dapat digali sehingga menghasilkan potensi dan pola atau pengetahuan yang dapat dimengerti berdasarkan data teks secara utuh dan tidak terstruktur. *Text mining* digunakan dalam berbagai tujuan, seperti memahami opini pengguna/konsumen dengan analisis sentimen, meningkatkan kualitas pelayanan, dan membuat keputusan yang tepat di berbagai bidang [31]. Terdapat beberapa tahapan pada *text mining* dalam melakukan proses penggalian, antara lain:

- 1) Tahap pra-proses: Pada tahap pra-proses, data teks dipersiapkan agar dapat dianalisis. Proses yang dilakukan pada tahap ini, meliputi:
 - a. *Data Collecting*: Data yang ingin dianalisis dikumpulkan dari berbagai sumber, seperti artikel, ulasan, media sosial, dan sebagainya [31], [32].
 - b. *Data Cleaning*: Data yang telah dikumpulkan kemudian akan dilakukan pembersihan sehingga bagian-bagian yang tidak relevan dan *noise* dapat dihilangkan [31], [32].
 - c. *Tokenization*: Pada tahap ini, data akan dipecah menjadi beberapa token [31], [32].
 - d. *Stemming*: Proses stemming akan menyeragamkan beberapa kata ke bentuk dasar dengan memisahkan kata dasar dari kata awalan dan kata akhiran. Proses ini akan mengelompokkan beberapa kata yang sudah dilakukan *stemming* menjadi satu.
 - e. *Stopword Removal*: Pada tahap ini, kata-kata umum pada sebuah kalimat dan tidak penting akan dihilangkan [31], [32].
 - f. *Case Folding*: Proses ini berfungsi untuk mengubah karakter kata pada sebuah kalimat menjadi huruf kecil atau huruf besar [31], [32].

- 2) Tahap analisis: Pada tahap analisis, data yang telah dilakukan pra-proses akan masuk ke proses analisis, meliputi analisis sentimen untuk mengidentifikasi ulasan/sosial media/artikel menjadi sentimen positif dan negatif, kategorisasi untuk mengklasifikasikan data teks ke dalam kategori, dan ekstraksi informasi agar informasi tertentu pada data teks dapat diidentifikasi.

2.3 Framework dan Algoritma Penelitian

2.3.1 Support Vector Machine (SVM)



Gambar 2. 2 Cara Kerja SVM

Sumber: ScienceDirect [33]

Support Vector Machine (SVM) merupakan salah satu algoritma dari *machine learning* yang dapat digunakan untuk regresi dan klasifikasi. Kehadiran algoritma SVM sebagai algoritma jenis baru dari *machine learning* pertama kali dikenalkan oleh Cortes dan Vapnik pada tahun 1992. Cortes dan Vapnik menerbitkan sebuah artikel sehingga SVM menjadi populer di kalangan peneliti [34]. Cara SVM bekerja adalah mencari *hyperplane* yang optimal agar dua kelas yang berbeda dapat dipisahkan dengan menggunakan margin maksimum sebagai pembatas keputusan antara *hyperplane* dan titik data terdekat. Pada klasifikasi biner, *hyperplane* membagi ke dalam dua kelas, yaitu positif dan negatif. Pada gambar 2.2, terdapat dua kelas data, *class 1*

sebagai kelas negatif dan *class 2* sebagai kelas positif. Titik ekstrim data positif dan data negatif dapat dijadikan sebagai *support vector* atau vektor pendukung yang berfungsi untuk menentukan posisi *hyperplane* yang akan membentuk garis *hyperplane* positif dan negatif. Titik data tersebut dapat diketahui setelah algoritma SVM melakukan *data training* terhadap distribusi data [35].

2.3.2 Bidirectional Encoder Representations from Transformers (BERT)

Bidirectional Encoder Representations from Transformers atau BERT merupakan model *deep learning* yang dikembangkan oleh Google pada tahun 2018 dan dirancang untuk melakukan pemahaman lebih mendalam terhadap konteks dari teks yang diberikan [36]. Penggunaan BERT adalah dengan membaca teks dari dua arah secara bersamaan. Hal tersebut memungkinkan model untuk melakukan pemahaman terhadap konteks dengan membedakan beberapa arti kata yang sama secara lebih menyeluruh dengan mempertimbangkan konteks disekitarnya. Lebih lanjut, BERT menggunakan arsitektur transformer untuk menangkap korelasi kata-kata pada teks tidak dengan kata per kata, melainkan memproses secara menyeluruh terhadap teks [9].

Dalam proses pelatihan, terdapat 2 tahap utama yang dimiliki oleh BERT, yaitu *pre-training* dan *fine-tuning*.

- 1) *Pre-training* merupakan tahap awal dalam melakukan proses pelatihan, dimana model akan dilatih dengan data dalam jumlah besar dan tidak berlabel. Hal tersebut bertujuan agar model dapat memahami struktur kata secara kontekstual [37].
- 2) *Fine-tuning* merupakan tahap yang dilakukan setelah *pre-training* selesai dilakukan. Pada tahap ini, BERT telah memiliki pemahaman kontekstual dan akan disesuaikan dengan data berlabel untuk melakukan tugas *Natural Language Processing* (NLP) [37].

2.3.3 IndoBERT

IndoBERT merupakan salah satu model *transformer* BERT yang memiliki 12 *hidden layers* dengan pemodelan yang dilatih dengan dataset

bahasa Indonesia, termasuk Wikipedia, artikel-artikel berita, hingga Korpus Web Indonesia [38]. IndoBERT memiliki beberapa jenis, diantaranya adalah IndoBERT-large, IndoBERT-base, dan IndoBERT-lite. IndoBERT-base-p2 adalah model dasar yang dimiliki oleh IndoBERT dengan pelatihan yang lebih kompleks pada dataset untuk mendapatkan output yang lebih akurat dan sesuai untuk melakukan analisis sentimen yang membutuhkan pemahaman bahasa lebih dalam [39].

2.3.4 Confusion Matrix

Confusion Matrix merupakan tabel dari metode evaluasi matriks yang digunakan sebagai pernyataan dari pengukuran kinerja model klasifikasi berdasarkan jumlah data uji yang benar dan salah [40]. Tabel evaluasi ini sangat populer dalam pembelajaran mesin dan banyak digunakan pada *supervised classification* atau model klasifikasi lainnya karena struktur dari tabel *confusion matrix* dipresentasikan dengan baris dan kolom [41].

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Gambar 2. 3 Tabel Confusion Matrix

Sumber: Towards Data Science [42]

Berdasarkan gambar 2.3, tabel *confusion matrix* memiliki empat sel, yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN) dimana kolom adalah *predicted class* dan baris adalah *actual class*. Sel *true positive* (TP) diisi dengan data yang diprediksi benar dan masuk ke dalam kelas yang benar, *true negative* (TN) diisi dengan data yang

diprediksi benar tidak termasuk ke dalam kelas tersebut, *false positive* (FP) diisi dengan data yang diprediksi benar tetapi tidak termasuk ke dalam kelas benar, dan *false negative* (FN) diisi dengan data yang diprediksi salah tetapi sebenarnya termasuk ke dalam kelas benar. *Confusion matrix* dapat digunakan untuk mengukur akurasi model dan menganalisis performa model dengan pengukuran *matrix* [41]. Terdapat empat pengukuran nilai pada *confusion matrix*, meliputi *accuracy*, *f1-score*, *recall*, dan *precision*.

1) *Accuracy*

Accuracy atau akurasi merupakan alat ukur pada *confusion matrix* yang biasa digunakan untuk mengukur kinerja model sehingga kelas dapat diprediksi dengan benar [43]. Rumus 2.1 merupakan rumus *accuracy*.

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

Rumus 2. 1 Rumus Accuracy

2) *F-1 score*

F1-score merupakan alat ukur pada *confusion matrix* yang biasa digunakan dalam pengevaluasian kinerja pada keseluruhan model klasifikasi. Dengan *F1-score*, *recall* dan *precision* dapat diukur keseimbangannya [43]. Rumus 2.2 merupakan rumus *f1-score*.

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Rumus 2. 2 Rumus F1-Score

3) *Recall*

Recall adalah alat ukur pada *confusion matrix* yang digunakan untuk mengukur efektivitas nilai yang diidentifikasi pada dataset terhadap model klasifikasi [43]. Rumus 2.3 merupakan rumus *recall*.

$$Recall = \frac{TP}{TP + FN}$$

Rumus 2. 3 Rumus Recall

4) *Precision*

Precision adalah alat ukur pada *confusion matrix* yang digunakan untuk mengukur keakuratan dari prediksi positif pada model klasifikasi [43]. Rumus 2.4 merupakan rumus *precision*.

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

Rumus 2. 4 Rumus Precision

2.3.5 Word2Vec

Word2vec adalah sebuah teknik pada *Natural Language Processing* (NLP) serta salah satu algoritma *word embedding* yang melakukan pemetaan pada setiap kata yang terdapat dalam teks ke dalam vector dengan Panjang N dengan membawa semantik atau makna dari kata tersebut. Word2vec menggunakan neural network dengan pendekatan *hidden layer* dan *fully connected layer* yang ukuran dari matriks bobot *layer* tersebut tergantung pada jumlah kata dalam korpus dan jumlah neuron di *hidden layer*. Sebagai contoh, kata “Indonesia” direpresentasikan menjadi sebuah vektor yang akan dilatih menggunakan korpus, sehingga vektor dari kata “Indonesia” akan berkesinambungan dengan vektor “Jakarta”. Sama halnya dengan vektor “Filipina” akan berdekatan dengan vektor “Manila”. Hal tersebut menandakan bahwa model word2vec dapat memahami jika kata “Indonesia” dan “Jakarta” adalah sama dengan kata “Filipina” dengan “Manila” yaitu negara dan ibukotanya [44]. Word2vec memiliki 2 pendekatan model yaitu *Continuous Bag-of-Word* (CBOW) dan *Skip-Gram*.

a. *Continuous Bag-of-Word* (CBOW)

Model CBOW bekerja dengan memprediksi suatu kata berdasarkan konteks di sekitarnya, yaitu kata-kata yang muncul baik sebelum ataupun setelah target kata. Pada kata-kata yang sering muncul, CBOW memiliki waktu latih yang lebih cepat dan akurasi yang sedikit lebih baik. Pada model CBOW, *projection layer* digunakan secara bersama untuk memproses semua kata dalam korpus sehingga dapat memaksimalkan kemungkinan kemunculan kata target berdasarkan konteks [45]

b. Skip-Gram

Model *Skip-Gram* merupakan model yang serupa dengan CBOW, namun memiliki pendekatan yang berlawanan dalam memprediksi output dan input target kata. *Skip-Gram* melakukan prediksi pada target konteks menggunakan satu kata tertentu. Model *Skip-Gram* memaksimalkan probabilitas munculnya kata-kata disekitarnya [45]. Dengan kata lain, model itu bertujuan untuk mempelajari hubungan antar kata secara mendalam, khususnya untuk kata-kata yang memiliki frekuensi kemunculan yang rendah.

2.3.6 Stratified K-Fold Cross-Validation (SKCV)

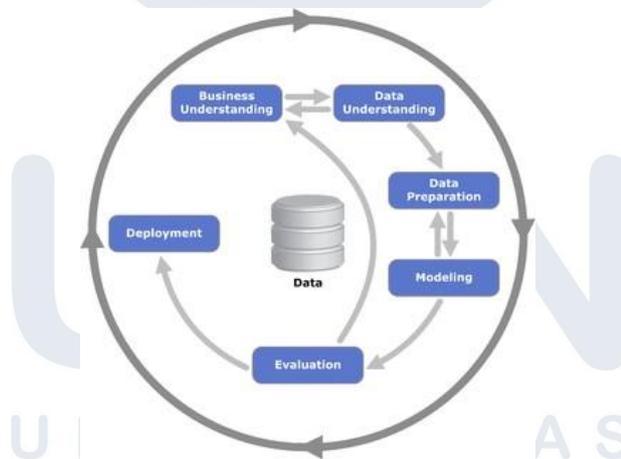
Stratified K-Fold Cross-Validation (SKCV) adalah teknik yang digunakan untuk melakukan evaluasi terhadap sebuah model dan merupakan bentuk lanjutan dari *Cross-Validation* (CV). SKCV bekerja dengan cara membagi distribusi kelas pada dataset secara merata di seluruh *k-fold* dan tidak mengganggu distribusi sampel tiap kelasnya [46]. Teknik SKCV berguna pada dataset yang memiliki kelas yang tidak seimbang karena menggunakan stratifikasi untuk menjaga distribusi asli dari setiap kelas di tiap lipatan. Sebagai contoh, jika dataset memiliki distribusi 85% kelas 0 dan 15% kelas 1, maka setiap lipatan akan dipertahankan pada proporsi yang serupa dalam data, sehingga menghasilkan akurasi model yang representatif [47].

2.3.7 Adaptive Synthetic Sampling (ADASYN)

ADASYN atau *Adaptive Synthetic Sampling* merupakan salah satu teknik *oversampling* yang digunakan untuk menangani kelas tidak seimbang

dalam dataset. ADASYN adalah varian atau pengembangan dari *Synthetic Minority Over-sampling Technique* (SMOTE), yang membedakan antara ADASYN dengan SMOTE adalah jumlah sampel kelas minoritas yang dihasilkan oleh ADASYN lebih banyak di area *k-nearest neighbor* dan berfokus pada bagian yang sulit diklasifikasikan [48]. Dengan kata lain, ADASYN memiliki kemampuan untuk memfokuskan pemrosesan duplikasi terhadap data pada area dengan tingkat sebaran sampel atau kepadatan data minoritas yang rendah dibandingkan area dengan tingkat sebaran tinggi untuk mencegah adanya ketidakseimbangan data dan dapat meningkatkan performa algoritma klasifikasi [49]. ADASYN memiliki keunggulan dalam menghasilkan data sintetis secara adaptif, sehingga data yang ditambah berdasar pada area yang sulit dipelajari model, sehingga dapat meningkatkan kemampuan klasifikasi model terhadap kelas minoritas, sedangkan SMOTE hanya melakukan penambahan data minoritas secara merata tanpa mempertimbangkan tingkat kesulitan tiap titik data yang dapat menyebabkan *overfitting*.

2.3.8 CRISP-DM



Gambar 2. 4 Siklus pada CRISP-DM

Sumber: DJKN Kementerian Keuangan Republik Indonesia [50]

Cross-Industry Standart Process for Data Mining (CRISP-DM) dirancang oleh 5 perusahaan pada tahun 1996 oleh Integral Solutions Ltd (ISL), Daimler AG, Teradata, NCR Corporation, dan OHRA yang kemudian dikembangkan oleh banyak organisasi dan perusahaan sehingga menjadikan CRISP-DM

sebagai *framework data mining* [51]. Pada umumnya, CRISP-DM merupakan kerangka kerja pada *data mining* dengan mengimplementasikan enam fase dalam mengidentifikasi masukan serta keluaran pada suatu proses. Berdasarkan gambar 2.4, enam fase pada CRISP-DM, meliputi *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*.

1) *Business Understanding*

Pada tahap ini, permasalahan suatu bisnis yang ingin dipecahkan, dipahami terlebih dahulu agar gambaran dari permasalahan dapat diidentifikasi serta menetapkan target *data mining* [52].

2) *Data Understanding*

Setelah memahami permasalahan suatu bisnis, langkah yang dapat dilakukan selanjutnya adalah mengumpulkan data yang diperlukan, melakukan eksplorasi data, mengevaluasi data, serta memahami korelasi antar atribut [52].

3) *Data Preparation*

Data yang telah dilakukan eksplorasi, selanjutnya dilakukan pra-proses data dengan melakukan *data cleaning* dengan menghilangkan data *null*, melakukan pengurangan dan penilaian terhadap dimensional data agar atribut yang penting dan tidak penting dapat diketahui, dan melakukan transformasi terhadap data [52].

4) *Data Modeling*

Pada tahap ini, algoritma pada *data mining* ditentukan untuk diaplikasikan pada objek penelitian. Tahap ini memiliki tujuan untuk membangun model pada *data mining* yang telah ditentukan. Dalam pengembangan model, *Google Colaboratory* akan digunakan sebagai *tools* pengembangan. Setelah menentukan algoritma yang digunakan, model dapat dilakukan *training* serta evaluasi [52].

5) *Evaluation*

Tahap evaluasi dapat dilakukan ketika model telah melalui proses pelatihan. Tahap ini akan dilakukan pengujian akurasi terhadap model yang dipilih dengan metode evaluasi, melakukan pengujian ketahanan model, serta interpretabilitas model [52].

6) *Deployment*

Tahap *deployment* memiliki tujuan untuk menerapkan model yang telah berhasil melewati tahap evaluasi. Tahap deployment ini melakukan integrasi terhadap model ke dalam sistem, melakukan pemantauan performa model, dan menjalankan model secara berkala pada sistem [52].

2.4 Tools dan Software Penelitian

2.4.1 Google Play Store



Gambar 2. 5 Logo Google Play Store

Sumber: katadata [53]

Google Play Store merupakan *marketplace* resmi yang dikembangkan dan dikelola langsung oleh Google sebagai toko yang dikhususkan bagi aplikasi Android, permainan, musik, buku, dan film. Pada tahun 2012, Android Market diubah nama menjadi Google Play yang berisi semua aplikasi yang dapat diunduh di Android, termasuk aplikasi Duolingo [54]. Pada Google Play Store, pengguna dapat melihat nama aplikasi, deskripsi singkat aplikasi, ulasan dan *rating* aplikasi yang ditulis oleh pengguna, jumlah unduhan aplikasi, versi aplikasi, dan sebagainya. Google Play Store menawarkan aplikasi yang lengkap dengan koleksi yang luas, keamanan yang terjamin karena tidak

sembarang aplikasi dapat dipublikasi ke Google Play Store, serta fitur pada aplikasi yang lengkap.

2.4.2 Python



Gambar 2. 6 Logo Python

Sumber: Python.org [55]

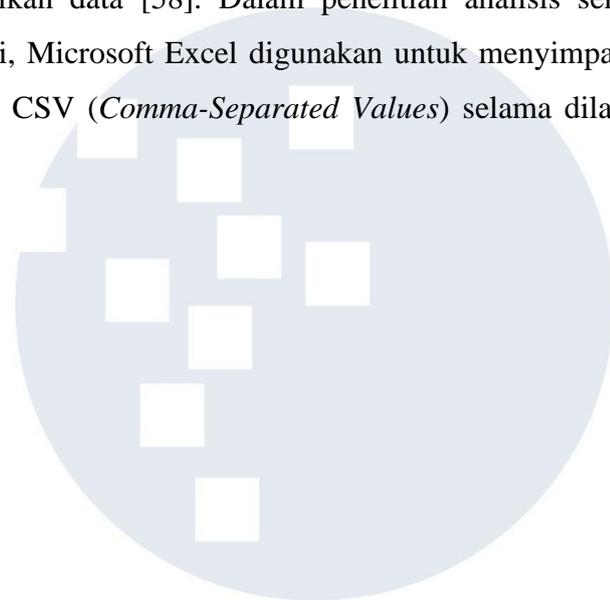
Python adalah salah satu bahasa pemrograman yang sangat populer digunakan oleh ratusan peneliti dan perusahaan. Sebagai bahasa pemrograman *open source*, Python dapat diimplementasikan di berbagai *platform*, seperti Windows, Mac OS, Linux, dan sebagainya. Python merupakan bahasa pemrograman serbaguna dan mudah dipelajari. Hal ini, disebabkan karena bahasa Python memiliki banyak *library code* dan alat pengembangan sehingga dapat terintegrasi dengan baik. Terdapat beberapa contoh dari penggunaan Python, seperti untuk mengembangkan web, mengembangkan aplikasi, sebagai ilmu data yang memiliki banyak kerangka kerja dan *library* untuk menganalisa data, dan sebagai bahasa pemrograman bagi pembelajaran mesin [56].

2.4.3 Google Colaboratory

Google Colaboratory merupakan sebuah *platform* yang digunakan untuk menulis, menjalankan, hingga berbagi kode Python kepada pengguna lain dan berbasis *cloud*. Google Colab dirancang untuk melakukan komputasi yang fleksibel dan mudah diakses oleh pengguna. Platform ini memiliki keunggulan yaitu mendukung GPU dan TPU secara gratis yang dapat membantu menjalankan model data atau pemrosesan data skala besar [57]. Penggunaan Google Colab dalam penelitian analisis sentiment berbasis aspek terhadap ulasan aplikasi adalah untuk menjalankan pemrosesan data serta model *machine learning*.

2.4.4 Microsoft Excel

Microsoft Excel merupakan program aplikasi bagian dari Microsoft Office berbentuk lembar kerja yang dirancang dan dikembangkan oleh Microsoft Corporation. Excel berfungsi untuk membantu pengguna menyelesaikan soal dalam bentuk matematika, menghitung aritmatika, mengolah data, dan membuat grafik dan diagram terhadap data. Dengan kata lain, Excel dapat melakukan perhitungan, analisis, mengolah data, serta mempresentasikan data [58]. Dalam penelitian analisis sentimen terhadap ulasan aplikasi, Microsoft Excel digunakan untuk menyimpan dataset dalam bentuk format CSV (*Comma-Separated Values*) selama dilakukan data pra-proses.



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA